Proceedings of the 2004 International Conference on Development and Learning

Editors: Jochen Triesch and Tony Jebara

Publisher: UCSD Institute for Neural Computation

Location: The Salk Institute for Biological Studies
La Jolla California, USA

ISBN: 0-615-12704-5

Organizing Committee

General Chair: Javier R. Movellan

Co-Chairs: Andrea Chiba, Gedeon Deák, Jochen Triesch, Marian Stewart Bartlett

Program Chairs: Jochen Triesch, Tony Jebara

Program Co-Chairs: Gwen Ford Littlewort, Marian Stewart Bartlett

Communications Chair: Andrea Chiba

Publications: Gedeon Deák

Finance: Kate Kern, Gris Arellano Ramirez

InformationTechnology / Audio Visual: Luis Palacios

Program Committee Members

Minoru Asada, Dana Ballard, Simon Baron-Cohen, Luis Baumela, Mark Baxter, Jeff Cohn, Kerstin Dautenhahn, Kenji Doya, Martha Farah, Teresa Farroni, Masahiro Fujita, Ann Graybiel, William Greenough, Michael Hasselmo, Hiroshi Ishiguro, Shoji Itakura, Robert Jacobs, David Kleinfeld, Brian Knutson, Mark Konishi, Christoph von der Malsburg, Denis Mareshal, Risto Miikkulainen, Douglas Nitz, Roz Picard, Steven Quartz, Rajesh Rao, Matthew Schlesinger, Geoffrey Schoenbaum, Gregor Schoener, Linda Smith, Olaf Sporns, Luc Steels, Valerie Stone, Manuela Veloso, Paul Verschure, Hiroyuki Yano

TABLE OF CONTENTS

Section 1: Attention and Learning in Social Systems

Papers

Abstracts

Section 2: Social Robots

Papers

Abstracts

Section 3: Social Systems

Papers

Abstracts

Section 4: Reinforcement and Neuromodulation

Papers

Section 7: Developmental Disorders

Abstracts

Section 8: Learning and Development, Modeling, Algorithms and Architectures

Papers

Abstracts

Section 9: Brain, Emotion, and Social Dynamics

Papers

Abstracts

Section 10: Perceptual Learning and Development

Papers

Abstracts

Section 11: Binding and Modularity

Papers

Cumulative Learning of Hierarchical Skills

Pat Langley
Seth Rogers
Computational Learning Laboratory
Center for the Study of Language and Information
Stanford University, Stanford, CA 94305 USA

Abstract

In this paper, we review ICARUS, a cognitive architecture that utilizes hierarchical skills and concepts for reactive execution in physical environments. In addition, we present two extensions to the framework. The first involves the incorporation of means-ends analysis, which lets the system compose known skills to solve novel problems. The second involves the storage of new skills and concepts that are based on successful means-ends traces. We report experimental studies of this mechanism in the blocks world, which show that learning operates in a cumulative manner that reduces the effort required to handle new tasks. We conclude with a discussion of related work on learning and prospects for additional research.

1. Introduction and Motivation

Research on cognitive architectures (Newell, 1990) attempts to understand the computational infrastructures that support intelligent behavior. A specific architecture specifies the aspects of a cognitive agent that remain the same across time and over different domains, and typically makes strong commitments about the representation of knowledge structures and the processes that operate them. Learning has been a central concern in most architectural research, with a variety of mechanisms having been proposed to model the acquisition of knowledge from experience.

In this paper we review ICARUS, a candidate architecture that diverges from its predecessors on a number of dimensions. One important difference is that most frameworks focus on production systems, which encode knowledge as a 'flat' set of condition-action rules, whereas ICARUS provides explicit support for hierarchies of both concepts and skills. In addition, most cognitive architectures evolved from theories of human problem solving, and thus focus on mental phenomena. In contrast, ICARUS is mainly an execution architecture that perceives and reacts to external environments.

However, ICARUS' reliance on hierarchical structures raises key questions about their origin. Moreover, the architecture's emphasis on execution does not mean that mental activities like problem solving are unimportant, since they can let an agent handle novel tasks for which stored knowledge is unavailable. The central hypothesis of this paper is that hierarchical skills and concepts arise, at least in many cases, from problem-solving behavior, and that, once learned, the agent can use these structures to support reactive execution in the environment.

In the sections that follow, we review ICARUS' representation and organization of concepts and skills, along with the categorization and execution processes that utilize them. After this, we present a new module that interleaves means-ends problem solving with execution when known skills are insufficient to solve a task. Next we describe a mechanism for creating generalized skills and concepts from traces of successful problem solving that supports both incremental and cumulative learning. We report experiments with this learning mechanism that demonstrate its ability to reduce effort on new problems and that examine effects of training order. In closing, we discuss earlier research on learning problem-solving knowledge and cumulative learning, along with some directions for future work.

2. Representation and Organization

Like other cognitive architectures, ICARUS makes commitments to its representation of knowledge, the manner in which it is organized, and the memories in which it resides. Here we discuss the framework's long-term and short-term memories, including formalisms used to encode their contents. We will take our examples from the blocks world, since many readers should find this domain familiar.

One of ICARUS' long-term memories stores Boolean concepts that describe situations in the environment. These may involve isolated objects, such as individual blocks, but they can also characterize physical relations among objects, such as the relative positions of blocks. Long-term conceptual memory contains the definitions of these logical categories. Each element specifies the concept's name and arguments, along with five optional fields – `:percepts`, which describes perceptual entities that must be present; `:positives`, which indicates lower-level concepts that must match; `:negatives`, which specifies lower-level concepts that must not match; `:tests`, which states numeric relations that must be satisfied; and `:excludes`, which indicates literals whose negation is entailed when the concept holds.

Table 1 presents some concepts from the blocks world. For example, *on* describes a perceived situation in which two blocks have the same x position and the bottom of one has the same y position as the top of the other. The concept *clear* refers to a single block, but one that cannot hold the relation *on* to any other, as specified in its `:negatives` field.

Definitions of this sort organize ICARUS categories into a conceptual hierarchy. Primitive concepts are defined entirely in terms of perceptual conditions and numeric tests, but many incorporate other concepts in their definitions. This imposes a lattice structure on the memory, with more basic concepts at the bottom and more complex concepts at higher levels. The resulting hierarchy is similar in spirit to early models of human memory like EPAM (Feigenbaum, 1963), as well as to frameworks like description logics.

ICARUS also incorporates a second long-term memory that stores knowledge about skills it can execute in the environment, including their conditions for application and their expected effects. Each skill has a name, arguments, and a set of optional fields. The `:start` field specifies the concepts that must hold to initiate the skill, whereas the `:requires` field indicates conditions that must hold throughout its execution, which may require multiple cycles to complete. The `:effects` field specifies a conjunction of concepts that, taken together, describe the situation the skill produces when done. For example, Table 2 shows the skill *pickup*, which must satisfy the single start condition, *(pickupable ?block ?from)*, defined in Table 1. The skill's only stated effect is to make *(holding ?block)* true.[1]

Each ICARUS skill also includes a field that specifies how to decompose it further. Two example skills in the table utilize the `:actions` field, which refers

[1]Note that the use of `:excludes` fields in concepts avoids the need for explicit delete lists in skills, as in most planning systems.

Table 1. Some ICARUS concepts for the blocks world, with variables indicated by question marks.

```
(on (?block1 ?block2)
 :percepts  ((block ?block1 xpos ?x1 ypos ?y1)
             (block ?block2 xpos ?x2 ypos ?y2
                    height ?h2))
 :tests     ((equal ?x1 ?x2) (>= ?y1 ?y2)
             (<= ?y1 (+ ?y2 ?h2)))
 :excludes  ((clear ?block2)))
(clear (?block)
 :percepts  ((block ?block))
 :negatives ((on ?other ?block))
 :excludes  ((on ?other ?block)))
(pickupable (?block ?from)
 :percepts  ((block ?block)(table ?from))
 :positives ((ontable ?block ?from)
             (clear ?block)
             (hand-empty)))
(pickup-stackable (?block ?from ?to)
 :percepts  ((block ?block)(table ?from)(block ?to))
 :positives ((pickupable ?block ?from)
             (clear ?to)))
```

to opaque actions the agent can execute directly in the environment. For instance, *unstack* invokes both **grasp*, which grasps a block, and **vertical-move*, which moves the hand in the vertical direction. However, the nonprimitive skill *pickup-stack* instead includes an `:ordered` field, which specifies the subskills of which it is composed, in this case the primitive skills *pickup* and *stack*.[2]

In fact, ICARUS lets one specify multiple ways to decompose a given concept or skill, much as a Prolog program can include more than one Horn clause with the same head. In addition, each skill decomposition can include a value function that encodes the utility expected if the agent executes the skill with that decomposition. Neither capability plays an important role in this paper, but we have described them in some detail elsewhere (Choi et al., 2004).

In addition to long-term memories, which encode stable knowledge about a domain, ICARUS includes short-term stores that change more rapidly. These make contact with long-term concepts and skills, but they represent temporary beliefs about the environment and intended activities. In particular, the *perceptual buffer* contains descriptions of physical entities that correspond to the output of sensors. For the blocks world, this includes literals like *(block B xpos 10 ypos*

[2]ICARUS also supports an unordered field for subskills that can be executed in any order, but they play no role here.

Table 2. Some ICARUS skills for the blocks world.

```
(pickup (?block ?from)
 :percepts ((block ?block)
            (table ?from height ?h))
 :start    ((pickupable ?block ?from))
 :actions  ((*grasp ?block)
            (*vertical-move ?block (+ ?h 10)))
 :effects  ((holding ?block)))
(stack (?block ?to)
 :percepts ((block ?block)
            (block ?to xpos ?x ypos ?y height ?h))
 :start    ((stackable ?block ?to))
 :actions  ((*horizontal-move ?block ?x)
            (*vertical-move ?block (+ ?y ?h))
            (*ungrasp ?block))
 :effects  ((on ?block ?to)(hand-empty)))
(pickup-stack (?block ?from ?to)
 :percepts ((block ?block)(block ?from)(table ?to))
 :start    ((pickup-stackable ?block ?from ?to))
 :ordered  ((pickup ?block ?from)
            (stack ?block ?to))
 :effects  ((on ?block ?to)))
```

2 width 2 height 2), which specify the position and size of individual blocks. Moreover, the *short-term conceptual memory* contains beliefs about the environment that the agent infers from items present in its perceptual buffer and its long-term concept memory. For instance, this might contain the instance *(on B C)*, which is an instance of the *on* concept in Table 1. Finally, a *short-term skill memory* contains the agent's intentions about skill instances it plans to execute, which lets the system engage in behavior that persists over time. Each literal specifies the skill's name and its arguments, as in *(stack B C)*.

3. Categorization and Execution

Like most cognitive architectures, ICARUS operates in distinct cycles. On each such iteration, the system updates its perceptual buffer by sensing objects in its field of view. This produces perceptual elements that initiate matching against long-term concepts. The matcher checks to see which concepts are satisfied, adds each matched instance to conceptual short-term memory, and repeats the process on the expanded set. In this way, ICARUS infers all instances of concepts that are implied by its conceptual definitions and the contents of the perceptual buffer. In the blocks world, the agent would first update its descriptions of the blocks and the table, then infer primitive concepts like *on*, and finally infer complex concepts like *unstackable*.

On each cycle, the architecture examines the intentions in short-term skill memory to determine which, if any, apply to the current situation and which one has the highest utility or value. For each skill instance, ICARUS accesses all expansions of the general skill to see if they are applicable. A skill is applicable if, for its current variable bindings, its `:effects` field does not match, the `:requires` field matches, and, if the system has not yet started executing it, the `:start` field matches the current situation. Also, for higher-level skills, at least one subskill must be applicable. Because this test is recursive, a skill is applicable only when ICARUS can find at least one acceptable path downward to an executable action. ICARUS considers all acceptable paths downward through the skill hierarchy, returning the path with the highest value.

When the values are equal, ICARUS selects one of the skill paths at random, as we assume in this paper. For example, suppose the agent has the intention *(pickup-stack A table B)* in a situation where the concept instances *(pickup-stackable A table B)* and *(pickupable A table)* hold. This means the path *((pickup-stack A table B), (pickup A table))* is applicable and would be considered for execution. If selected, the *pickup* skill would alter the environment, making acceptable the path *((pickup-stack A table B), (stack A B))* on the next cycle. This would produce a state that satisfies the effects of *(pickup-stack A table B)*, making any path in which it occurs unacceptable.

The architecture handles a skill differently depending on how it is decomposed. For primitive skills that include an `:actions` field, ICARUS executes each of the physical actions, one after another, on a single cycle. For higher-level skills that have an `:ordered` field, it treats the list as a reactive program that considers subskills in reverse order. If the final subskill is applicable, then it considers only paths which include that subskill. Otherwise, it considers the penultimate skill, the one before that, and so forth. Presumably, the subskills are ordered because later ones are closer to the parent skill's objective and thus are preferred when applicable.

4. Means-Ends Problem Solving

As explained above, the previous version of ICARUS could execute complex hierarchical skills in a reactive manner, but it assumed that these skills were already present in long-term memory. Although much human behavior involves the application of such routine skills, people can solve novel problems that require the combination of existing knowledge elements.

To model this capability in ICARUS, we have introduced a variant of means-ends analysis (Newell, Shaw, & Simon, 1960) that operates over the architecture's knowledge structures. Traditional means-ends problem solving selects some unsatisfied aspect of the goal state to achieve, then selects an operator that would achieve it. If that operator's preconditions match the current state, it is applied; otherwise, the method selects an unsatisfied precondition to achieve, selects an operator that would achieve it, and so on. Once a condition is met, the process is repeated until the original goal description is satisfied. This may require search, which is typically pursued in a depth-first manner. Means-ends analysis has been implicated repeatedly in human problem solving on novel tasks.

Icarus implements a variant of this mechanism with a stack that contains goal elements in an ordered list. Each goal element specifies an objective (a desired goal literal) and whether it involves backward chaining off a concept definition or a skill. If the latter, then the element may specify a skill that achieves the objective. Also, a goal element may have a 'failed' field for skills or concepts that it has tried and rejected. On each cycle, Icarus takes one of six steps:

- If the stack's top entry E has objective O but has no associated skill, it retrieves skills with O in their effects that have not failed before and do not clobber any achieved goals earlier in the stack, selects an instance S from this set, and associates S with E.
- If Icarus retrieves no skills that would achieve objective O, it determines which instantiated subconcepts of O are not met, selects one (C) at random, and adds a goal element with C as its objective.
- If the top entry E on the stack has an associated skill instance S that is applicable, in the sense described above, then Icarus selects a skill path for S and executes it in the environment.
- If the stack's top entry E includes an associated skill instance S that is not applicable, then Icarus adds a new entry on top of the stack with the start condition of skill S as its objective.
- If the objective O for the top entry E on the stack is satisfied by the current environmental state, then Icarus pops E from the stack.
- Otherwise, if the system cannot find a skill instance that does not appear in the entry E's failed field, and if chaining off the unmatched elements of E's objective O has failed, then it pops E from the stack and stores O in the failed field of E's parent.

Each of these activities takes a single cycle of the architecture, with the initial situation being a special case of the first item that triggers the process. Because reasoning about how to achieve an objective can require many manipulations of the goal stack, it takes more cycles than executing a hierarchical skill for that objective, even when the agent does not have to backtrack.

Search enters into this formulation in two places. One involves backward chaining off the unmatched elements of a concept definition. Here Icarus selects a literal randomly from those not yet tried and keeps track of literals it has failed to achieve. The other involves backward chaining off skills that, if executed, would achieve the objective of the current stack entry. Here Icarus considers only skill instances that have not yet failed and prefers ones that have the fewest expanded `:start` conditions unmet by the current environmental state, with fully matched conditions being most desirable. If candidates tie on this criterion, it prefers skill instances that have a shorter expected duration, and if ties remain, it selects a candidate at random.

Taken together, these biases produce a heuristic version of means-ends analysis. However, this problem-solving method is tightly integrated with the execution process. Icarus backward chains off concept or skill definitions when necessary, but it executes the skill associated with the top stack entry as soon as it becomes applicable. Moreover, because the architecture can chain over hierarchical reactive skills, their execution may continue for many cycles before problem solving is resumed. In contrast, most models of human problem solving and most AI planning systems focus on the generation or the execution of plans, rather than interleaving the two processes.

Of course, executing a component skill before constructing a complete plan can lead an agent into difficulties, since it is harder to backtrack in the world than in one's head. This strategy may well lead to suboptimal behaviors, but human intelligence is more about satisficing than optimizing, and interleaving problem solving with executing requires far less memory than constructing a full plan before executing it. However, it can produce situations from which the agent cannot recover. Thus, if Icarus has not achieved the top-level objective in a goal stack within N cycles, it resets the environment in the original situation and tries again, with no memory of its earlier attempts.

5. Learning from Problem Solving

In the previous pages, we described two facets of Icarus: its execution of hierarchical skills on familiar tasks and its use of problem solving to handle novel ones. The first lets the system operate efficiently, but skills are tedious to construct manually, whereas the second gives the system flexibility but requires reasoning and means-ends search. We believe that humans also have both capabilities, but that they use learning to transform the results of successful problem solving into hierarchical skills. We would like to incorporate a similar capability into Icarus.

However, we want our learning mechanisms to reflect certain properties that appear to hold for human skill acquisition. One is that learning should take advantage of existing knowledge, such as the definitions of current skills and concepts. In addition, acquisition should be incremental and interleaved with the problem-solving process. Taken together, these imply that learning should be *cumulative* in that it builds directly on the results of previous learning. The literature on computational learning contains remarkably few cases of such cumulative knowledge acquisition.

Our extension of Icarus achieves this effect through a form of impasse-driven learning that is tied closely to its problem-solving and execution processes. As in Soar (Laird et al., 1986), the purpose of skill learning is to avoid such impasses in the future. Thus, whenever the architecture achieves an objective that is associated with an entry in the goal stack, this provides an opportunity for learning. The system acquires three distinct forms of skill, which we describe in turn.

The first category results from situations in which Icarus has attempted to execute a skill instance S to achieve an objective O, but found its start conditions unsatisfied and selected another skill instance, P, to achieve them. Once both skills have been executed successfully and the objective reached, the system constructs a new skill N that has P and S as ordered subskills. The objective of N is the original objective, O, and the start condition is a new concept, C, that includes the conditions of O that were satisfied initially, the preconditions of S that were satisfied initially, and the start conditions of P. The definitions have their arguments replaced by variables in a consistent manner. For example, the skill *pickup-stack* in Table 1 might be learned from executing *(pickup A table)* followed by *(stack A B)* to achieve the goal *(on A B)*.

The other types of skills result from situations in which the problem solver could not find a skill to achieve an objective O, and thus created as subgoals the literals $\{O_1, O_2, \ldots, O_n\}$ from the unsatisfied conditions of O's conceptual definition. Suppose these subgoals have each been achieved in turn by executing the skill instances $\{S_1, S_2, \ldots, S_n\}$, respectively, thus satisfying the parent goal O. When this occurs, Icarus constructs a new skill N with ordered subskills $\{S_1, G_2, \ldots, G_n\}$. Each G_k is a "guard" skill with S_k as a single subskill, with no effects, and with $\{O_1, \ldots, O_{k-1}\}$ as its start conditions, which ensure that S_k is invoked only after these objectives have been met. Their parent skill N has O as its effect and, as its start condition, a new concept C that includes both the elements of O that were satisfied initially and the analogous elements of S_1, ..., and S_n. Again, specific arguments are replaced consistently by variables.

We have emphasized the construction of hierarchical skills, but, as noted above, Icarus also acquires new concepts in the process. These play the role of start conditions for the new skills and ensure they are executed only when appropriate. Thus, one can view these concepts as functionally motivated, even though their definitions are purely structural. For example, the concept *(pickup-stackable ?block ?from ?to)* created as the start condition of skill *pickup-stack* above is defined as the conjunction of *(pickupable ?block ?from)* and *(clear ?to)*, which is the situation in which executing *(pickup ?block ?from)* followed by *(stack ?block ?to)* will achieve the effect *(on ?block ?to)*.

These learning mechanisms are fully incremental, in that each learning event draws on a single problem-solving experience and thus requires no memory of previous ones. They support within-trial learning, since skills acquired on one subproblem may be used to handle later subproblems. The processes also build on existing knowledge, since the construction of new skills and concepts involves the composition of those used in a training problem's solution. Taken together, these support a form of cumulative learning, in which Icarus learns skills and concepts on one problem, uses them to solve a later problem, and incorporates them into still higher-level skills and concepts.

6. Experiments with Hierarchy Learning

Initial studies with the blocks world and the Tower of Hanoi confirmed that the extended version of Icarus learns hierarchical skills and concepts in the manner described. Moreover, they revealed that, when given the same task to solve a second time, the system utilizes this knowledge to handle it without problem solving, although this does not mean it completes the problem in a single cycle. Recall that, unlike traditional cognitive architectures, Icarus resorts to problem solving only to enable execution, and it must still execute its acquired skills to reach an objective. Thus, for a problem that requires four primitive steps, the system takes six cycles on the second encounter, with one to retrieve the hierarchical skill and one to realize it has finished.

For the blocks world, Icarus learns skills for achieving particular configurations from different initial configurations, along with concepts for the start conditions of each skill and subskill. Yet because the system generalizes its learned structures beyond the specific instances on which they are based, it can handle without problem solving any task that is isomorphic to one it has already solved. This isomorphism must involve the same objective and have the same subconcepts satisfied or unsatisfied in the initial environment.

However, we desired more than anecdotal demonstrations that the new mechanisms supported cumulative learning of skills and concepts. We also wanted evidence from systematic experiments that this learned knowledge produces more effective behavior. To this end, we examined the state space for blocks-world problems that involve three blocks. If one ignores isomorphisms, then there are five problems that can be solved in two primitive steps, eight tasks solvable in four steps, nine six-step problems, and four eight-step problems.[3] These 26 tasks constituted both the training and test problems for the study.

We provided the system with four primitive skills and ten concepts, including one for the desired state, that were sufficient, in principle, to solve these blocks-world problems. We then presented it with these problems in sequence, using each task as a training problem but also recording the number of cycles required to complete it. Because misguided search combined with execution can lead the problem solver into undesirable physical states, we told it to halt if it had not finished a run within 50 cycles and start over from the initial state. However, it could attempt a given problem only ten times, and thus spend at most 500 cycles before giving up entirely. We also limited the stack depth to six goal elements. We enforced these constraints for reasons of practicality and because we think they reflect the manner in which humans tackle novel problems.

We ran Icarus on the 26 blocks-world problems, ordering them by difficulty class (two-step tasks first and eight-step tasks last) but randomly within a class. The intuition was that the system would learn more effectively if we presented it first with simpler problems, which it could then use in solving more difficult ones. To this end, Icarus retained skills and concepts acquired on successful runs for use in later tasks. However, if the system failed on a given run, it removed any skills and concepts created during that run, to prevent influence on later attempts. We ran Icarus over 200 randomly generated problem orders and averaged the number of cycles needed at each level of experience. As a control, we also ran the architecture with its learning mechanisms off for another 200 random problem sets.

Figure 1 shows the result of this experiment, including 95 percent confidence intervals around each mean. The two curves show clearly that learning reduces the total cycles required to solve problems in the blocks world. Both curves are step functions that increase with problem difficulty, as one would expect. Remember that none of the problems are isomorphic, although

[3]We ignored tasks with an odd number of steps, since these start or end with a block in the air. Also, we considered only problems in which the objective was a fully specified state.

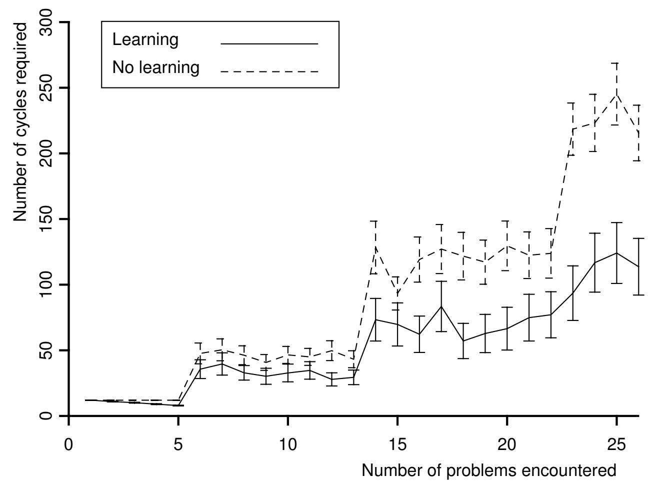


Figure 1. Number of cycles required by Icarus to solve a blocks-world task as a function of the number of training problems, averaged over 200 runs, with order randomized within each difficulty class.

they may involve isomorphic subtasks. The results suggest that Icarus takes advantage of that similar substructure to reduce its effort on later problems. In a typical run on 26 problems, the system constructed 9 new concepts and 74 skills, including 9 skill-chaining skills, 34 concept-chaining skills, and 31 guard skills.

We presented Icarus with problems in increasing order of difficulty because we believed this training regimen would lead to better learning. Our intuition was that, because the system would be more likely to solve simpler problems, it would more readily acquire skills and concepts that would prove useful in more complex ones encountered later. However, this hypothesis seemed worth testing experimentally, so we carried out another study with this in mind. In this case, we held back the four eight-step tasks for testing, and let Icarus learn only from the 22 simpler problems.

We examined three conditions, one in which (as before) problems were ordered randomly within their difficulty class, one in which they were ordered randomly without this consideration, and one in which no learning occurred. Again we averaged the required number of cycles over 200 different runs and, in this case, over the four test problems. As expected, the condition with no learning fared far worst, taking 236.78 ± 11.43 cycles. However, the skills acquired from problems ordered by difficulty took 113.88 ± 11.37, whereas those learned from randomly ordered tasks took 99.06 ± 10.16. Thus, presenting simpler problems earlier did not appear to help Icarus learn any more effectively.

To understand better the factors at work, we repeated the random order condition with fewer training problems, again testing on the four eight-step tasks. When trained only on the five two-step problems, the average over 200 runs was 262.66 ± 12.72 cycles, and when the system learned from these and the eight four-step problems, the average was 160.38 ± 12.48 cycles, while the difficulty ordering produced nearly the same results. Thus, ICARUS shows steady improvement with experience, apparently acquiring useful skills and concepts even from relatively complex training problems.

7. Related Research

The use of background knowledge to support learning has a long history within both AI and cognitive science. Research on explanation-based learning often aimed to improve efficiency on problem-solving tasks and combined experience with a domain theory to create new cognitive structures. Some techniques acquired search-control rules to guide problem solving, but others instead constructed macro-operators from primitive operators (e.g., Iba, 1988; Mooney, 1989). Our approach to skill learning comes closer to the second tradition, since both involve composing knowledge elements into larger structures. However, ICARUS adapts this idea for the creation of skill hierarchies, whereas earlier methods produced flat macro-operators that contained less structure than the original knowledge base.[4]

ICARUS also has similarities to other cognitive architectures that incorporate varieties of explanation-based learning. For example, Laird, Rosenbloom, and Newell's (1986) SOAR revolves around a problem solver that proceeds until the system encounters an impasse, in which case it carries out search to resolve it. Once SOAR has handled the impasse, it creates a *chunk* that encodes a generalized explanation of the result in terms of the original goal context. Anderson's (1993) ACT-R employs another mechanism, *compilation*, that creates a new production rule from ones that are involved in the same reasoning chain. This scheme produces very specific rules that replace variables with the declarative elements against which they matched, rather than forming generalized structures. In fact, our approach is much closer to the composition process that played a role in much earlier versions of ACT.

The architecture most akin to ICARUS is PRODIGY (Minton, 1990), which invokes means-ends analysis to solve problems and uses an analytical method to learn either search-control roles or macro-operators from problem-solving traces. Veloso and Carbonell (1993) also describe an extension that records these traces in memory and solves new problems by analogy with earlier ones. None of these mechanisms generates explicit hierarchical structures, but because the latter stores cases of ever-increasing size, it can produce effects similar to the cumulative learning found in ICARUS.

A few researchers have built systems that support cumulative learning outside the context of problem-solving tasks. One early example was Sammut and Banerji's (1986) Marvin, which learns logical concepts that are composed of other concepts. A human trainer presents the system with examples of increasingly complex concepts, ensuring it has mastered each one before moving to the next. Pfleger (in press) describes another system that acquires hierarchical patterns in an unsupervised on-line setting. Like Marvin, it learns conceptual structures from the bottom up, so that more complex patterns are apparent after simpler ones have been acquired. Stracuzzi and Utgoff (2002) report a third system that learns in a cumulative fashion.

Ruby and Kibler's (1991) SteppingStone also learns to solve more difficult problems based on solutions generalized from simpler ones, which it obtains through a mixture of problem reduction and exhaustive search. Benson's (1995) TRAIL incorporates a reactive control module that invokes learning when it reaches an execution impasse. Observation and experimentation drive learning rather than problem solving, and the system acquires models for primitive actions rather than hierarchical structures, but its later learning depends on earlier experience. Ilghami et al. (2002) present another system that organizes plan knowledge in a hierarchical task network, but learns conditions for method selection rather than the network itself. A closer relative to ICARUS is Reddy and Tadepalli's (1997) X-Learn, which acquires goal-decomposition rules from a sequence of training exercises. Their system does not include an execution engine, but it generates hierarchical plans and learns structures in a cumulative manner.

8. Concluding Remarks

In the preceding pages, we presented ICARUS, a cognitive architecture for physical agents that uses stored concepts and skills, both organized in hierarchies, to recognize familiar situations and control its behavior. We described a new module that supports means-ends problem solving on novel tasks, along with a learning mechanism that produces new skills and concepts from traces of problem solutions. This method operates in an incremental and cumulative manner, creating hierarchical structures that refer to others learned earlier. In addition, we reported experiments with the blocks world that showed such learning enables more effective behavior on unfamiliar problems.

Despite these advances, our work on cumulative learning in ICARUS is still in its early stages. For in-

[4] However, we have adopted Mooney's key idea that one should not chain off the preconditions of learned skills.

stance, we should show its ability to learn hierarchical structures on other problem-solving tasks besides the blocks world and the Tower of Hanoi. More important, we should study Icarus' behavior in dynamic domains that require integration of problem solving with reactive control. A prime candidate is the driving environment we have used to evaluate the architecture's categorization and execution modules (Choi et al., 2004).

In addition, Icarus' methods for problem solving and hierarchical learning would benefit from new capabilities. The current system selects subgoals randomly when chaining off a concept definition, which means that it must often backtrack even when it has skills for component subproblems. Extending the problem solver to select subgoals heuristically would let it take better advantage of learned subskills. Nor can Icarus acquire recursive skills for tasks that involve regular structure, such as building towers in the blocks world. Analyzing relations among learned skills may provide this ability, which should let the system transfer learned knowledge to problems with more objects.

We should also address the *utility problem*, which can actually produce slower behavior in systems that learn problem-solving skills. Our plans here involve storing an expected duration and success probability with each skill, which would then be used in execution and problem solving. Initial estimates would come from a skill's components but would be revised as the agent utlizes the skill. Combined with other extensions, this should give Icarus a more robust and effective approach to cumulative learning that, in its own right, builds on our experience with the current architecture.

Acknowledgements

This research was funded in part by Grant HR0011-04-1-0008 from DARPA IPTO and by Grant IIS-0335353 from the National Science Foundation. Discussions with Glenn Iba, David Nicholas, Stephanie Sage, and Dan Shapiro contributed to many ideas presented here.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Benson, S. (1995). Induction learning of reactive action models. *Proceedings of the Twelfth International Conference on Machine Learning* (pp. 47–54). San Francisco: Morgan Kaufmann.

Choi, D., Kaufman, M., Langley, P., Nejati, N., & Shapiro, D. (2004). An architecture for persistent reactive behavior. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi Agent Systems* (pp. 988–995). ACM Press.

Feigenbaum, E. A. (1963). The simulation of verbal learning behavior. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought*. New York, NY: McGraw-Hill.

Iba, G.A. (1989). A heuristic approach to the discovery of macro-operators. *Machine Learning*, *3*, 285–317.

Ilghami, O., Nau, D. S., Muñoz-Avila, H., & Aha, D. W. (2002). CaMeL: Learning method preconditions for HTN planning. *Proceedings of the Sixth International Conference on AI Planning and Scheduling* (pp. 131–14). Toulouse, France.

Laird, J. E., Rosenbloom, P. S., & Newell, A. (1986). Chunking in Soar: The anatomy of a general learning mechanism. *Machine Learning*, *1*, 11–46.

Minton, S. N. (1990). Quantitative results concerning the utility of explanation-based learning. *Artificial Intelligence*, *42*, 363–391.

Mooney, R. J. (1989). The effect of rule use on the utility of explanation-based learning. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence* (pp. 725–730). Morgan Kaufmann.

Newell, A., Shaw, J. C., & Simon, H. A. (1960). Report on a general problem-solving program for a computer. *Information Processing: Proceedings of the International Conference on Information Processing* (pp. 256–264). UNESCO House, Paris.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Pfleger, K. (in press). On-line cumulative learning of hierarchical sparse n-grams. *Proceedings of the Third International Conference on Development and Learning*. San Diego, CA: IEEE Press.

Reddy, C., & Tadepalli, P. (1997). Learning goal-decomposition rules using exercises. *Proceedings of the Fourteenth International Conference on Machine Learning* (pp. 278–286). Morgan Kaufmann.

Ruby, D., & Kibler, D. (1991). SteppingStone: An empirical and analytical evaluation. *Proceedings of the Tenth National Conference on Artificial Intelligence* (pp. 527–532). Menlo Park, CA: AAAI Press.

Sammut, C., & Banerji, R. B. (1986). Learning concepts by asking questions. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2). Los Altos, CA: Morgan Kaufmann.

Utgoff, P., & Stracuzzi, D. (2002). Many-layered learning. *Proceedings of the Second International Conference on Development and Learning* (pp. 141–146).

Veloso, M. M., & Carbonell, J. G. (1993). Derivational analogy in Prodigy: Automating case acquisition, storage, and utilization. *Machine Learning*, *10*, 249–278.

To Care or Not to Care: Analyzing the Caregiver in a Computational Gaze Following Framework

Christof Teuscher and Jochen Triesch
University of California, San Diego, Department of Cognitive Science
9500 Gilman Drive, La Jolla, CA 92093–0515, USA
christof@teuscher.ch, triesch@ucsd.edu

Abstract

We first present a computational framework of the emergence of gaze following that is based on a generic basic set of mechanisms. Whereas much attention has been focused so far on the study of the infant's behavior, we systematically analyze the caregiver and show that he plays a crucial role in the development of gaze following in our model, especially for virtual infants with "developmental disorders".

We first create two nearly optimal infant parameter sets by means of an evolutionary algorithm and test their behavior with a simple standard caregiver. Based on these findings we then propose new caregiver models and evaluate them on normally developing and on infants with "developmental disorders".

1 Introduction

Humans and many animals live in social groups, which confers a number of benefits and costs to its members. But living in groups also requires different cognitive abilities for social interactions, which often rely on visual inputs since they are less ambiguous than auditory and olfactory signals and allow for a much richer and more complex communication. One such cognitive skill—with an immediate benefit to the members of a group—is the capacity to imitate other members and to learn from them, reason why humans share with a number of non-human primates the ability to use the eyes, the head, and the body of others to orient to important objects and events in their environment [18, 20]. The ability to follow the direction of conspecifics' visual gaze does not only help to localize interesting or dangerous entities in the environment, but also provides rich information about the group mates. For example, gaze following can be used to determine the position of an individual in the dominance hierarchy of large groups, where each individual receives attention as a function of its social rank [7].

Compared to other mammals, human newborns are nearly helpless, but rather quickly begin to show burgeoning social responsiveness [8], and by 3–6 months, infant and caregiver typically engage in complex patterns of reciprocal interaction [14]. By their first birthday, normally developing infants show a robust gaze following [9]. The child then gets more and more adept at recognizing diverse social cues, such as eye direction and pointing gestures. These skills serve as a developmental basis for more complex social communication skills such as the development of the infant's theory of mind [1] and the development of language, which starts around 13 months and is largely based on joint attentional interactions with adults and objects [4, 24]. Note that in the literature, a distinction is sometimes made between *shared attention* and *joint attention* [10]: in joint attention, two individuals are attending the same object only, whereas shared attention requires each having knowledge of the directions of the other's attention. In this paper, we only deal with joint attention.

Gaze following has been addressed in a number of computational and embodied models (see Section 1.1) where most attention has been focused on the infant. The goal of this paper is to systematically investigate the role of the caregiver in a computational gaze following framework.

We first present a computational framework (Section 2) of the emergence of gaze following that is based on a generic basic set of mechanisms. In order to have a nearly optimal infant opposite to the fixed caregiver, we optimized the infant's parameters by means of an evolutionary algorithm in Section 2.5. The outcome are two infants which perform almost equally well, but have different "personalities". Section 2.6 analyzes the caregiver's parameters of the original computational model as presented in [6]. Section 3 describes and analyzes four new caregiver models with respect to our two reference infants, including versions with developmental disorders. Section 4 concludes the paper.

1.1 Related Work

One of the main questions is how shared attention develops and what the necessary and sufficient conditions are. There are basically two alternative hypotheses at the extremes, whereas a combination of both could also be imagined: (1) it is hard-wired into the infant's brain or (2) it emerges through a learning process while the infant interacts with its environment. The most prominent exponent of the hard-wired hypothesis is Baron-Cohen's theory of social-cognitive modules [2], which consists of four modules. This nativist and modularist description provides little explanatory and predictive power and does not explain how the modules develop.

In 2001, Matsuda and Omori [16] proposed a reinforcement learning model for acquiring joint visual attention in infants. In their model, the infant's behavior is supervised and rewarded when it follows the mother's gaze. Further, they concluded that three abilities are required for joint visual attention to emerge: (1) the recognition of the mother's face, (2) detection of the eye gaze direction, and (3) a reward system.

Recently Carlson *et al.* [6] proposed a new and more realistic dynamical systems approach which takes into account that complex behaviors can emerge from simple learning mechanisms and which is based on a different and more complete basic set of hypotheses (see Section 2). It basically relies on Moore's suggestion [17] that gaze following might emerge because infants learn that the caregiver's direction of gaze is a reliable predictor of where interesting things are located. The proposed approach is the outcome of a more general framework seeking to combine embodied models and empirical research [11]. In contrast to the work of Matsuda and Omori, the infant learns without supervision and only receives a reward when he sees an interesting object or looks at the caregiver's face (see Section 2.3), i.e., the model therefore learns how to use the caregiver's face as a predictor for where interesting things are.

Nagai *et al.* [19] present a constructive model by which a real robot acquires the ability of joint attention with a human caregiver without supervised feedback. Although the model develops according to the three developmental stages as proposed by Butterworth and Jarret [5], it cannot deal with ambiguous object situations, therefore Lau and Triesch [15] recently proposed a new approach which uses the infant's depth perception to resolve such ambiguities.

A number of other researchers (see for Example [3, 21]) proposed and implemented mechanisms for joint attention for their robots, however, the behaviors were usually fully programmed in advance.

2 A Simple Computational Model

The computational model proposed in [6] relies on a *basic set* of plausible mechanisms [11], which were shown to be sufficient for gaze following to emerge:

- **Perceptual preferences:** Normally developing infants enjoy looking at faces in general and their caregivers in particular. In contrast, many children with autism do not show such a social preference.
- **Habituation:** Infants tend to shift gaze from one target to another after some time.
- **Reward driven learning mechanisms:** The infant shifts gaze between social stimuli and interesting targets in order to maximize internal rewards resulting from visual stimuli.
- **Contingent interactions and a structured environment:** There must be a correlation between where the caregiver is looking and where interesting things are.

The basic set of mechanisms is then formalized within the framework of a biologically plausible temporal difference reinforcement learning algorithm [23]. The framework shall be briefly summarized in the following.

2.1 Environment and Object

The environment is represented as a set of N distinct spatially unattributed regions where exactly one interesting object is present at any discrete time step t in one of the N locations. After an initial fixation time T_{fix}, the object has a relocation probability of p_{shift} at each upcoming time step t. Once relocated to a new location it will remain fixed again for at least T_{fix} time steps before the cycle restarts. Carlson *et al.* [6] used an environment with $N = 10$ regions, $T_{\text{fix}} = 4$, and $p_{\text{shift}} = 0.5$.

2.2 The Caregiver (CG)

The caregiver can either look at the object, at one of the $N - 1$ remaining empty regions, or at the infant. Each time the object is relocated to a new location the caregiver decides to look at either the infant with probability $\frac{1}{N+1}$, at the object with probability $\frac{N}{N+1}p_{\text{valid}}$, or at an empty location with probability $\frac{N}{N+1}(1-p_{\text{valid}})$. He then fixates the location or the infant until the target is relocated again as described in the previous section. Carlson *et al.* [6] set p_{valid} to 0.75, which means that the CG spends about 10% of its time looking at the infant, 68% at the object and 22% at an empty location in an $N = 10$ region environment.

2.3 The Infant (INF)

The infant is modeled as a pleasure-driven temporal difference (TD) reinforcement learning agent [23], which tries to maximize the rewards it receives for looking at interesting things. At each time step, the infant can look at one of the N regions or at the caregiver, whereby, the infant can see four possible things: (1) the object, (2) no object, (3) the caregiver's frontal view, or (4) the caregiver's profile. Associated with these views are four base reward values $R_{\text{fix}} \in \{R_{\text{object}}, R_{\text{nothing}}, R_{\text{frontal}}, R_{\text{profile}}\}$. The infant can only tell where the caregiver is looking when he directly gazes at him. When the infant is looking at something he habituates to it. For each location N, the caregiver, and the object the infant has a habituation value. As the infant continues to fixate on something, this value decreases, likewise, the infant dishabituates to all other possible looking locations. The habituation change at each time step is given by $h(t+1) = \frac{h(t)}{e^{\beta}}$, the dishabituation by $h(t+1) = 1 - \frac{1-h(t)}{e^{\beta}}$. In our experiments, all habituation values are set to 1 at initialization and $\beta = 1$. The actual or instantaneous reward received by the infant at time t is given by $r_{inst}(t) = R_{fix}(t)h(t)$.

Based on the work of Findlay and Walker [12], the decision of when and where to shift gaze is implemented in two separate agents in our infant model. The *when-agent* decides whether to continue to fixate on the same location or two shift gaze. It's state space has two dimensions: (1) the time the infant has been fixated on the current location and (2) the instantaneous reward received by the infant. The *where-agent* provides the new gaze location if the when-agent decides to shift gaze. It has one dimension only, namely, the caregiver's gaze direction, which can be one out of the N locations, the infant $(N+1)$, or unknown $(N+2)$ in the case the infant is not looking at the caregiver. Both agents make use of a standard TD learning algorithm with tabular SARSA [23] to estimate the state-action values $Q_t(s,a)$:

$$\begin{aligned} Q_{t+1}(s,a) &= Q_t(s,a) \\ &\quad + \alpha[r_t + \gamma Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t)] \end{aligned} \tag{1}$$

The action selection is made with a standard *softmax* decision rule, where action a is chosen with probability

$$p_t(a|s) = \frac{e^{\tilde{Q}_t(s,a)/\tau}}{\sum_{a'=1}^{N} e^{\tilde{Q}_t(s,a)/\tau}}, \tag{2}$$

where $\tilde{Q}_t(s,a) = \frac{Q_t(s,a)}{max_{a'}|Q_t(s,a')|}$.

The following parameters are important for the model:

1. α, *learning rate*: A small learning rate induces slow learning, a large learning rate oscillations.

2. γ, *discount factor*: Specifies how far in the future rewards should be taken into account. For small γ the agent is only interested in immediate rewards and does not consider long-term consequences of its actions.

3. τ, *temperature*: The lower τ, the more likely it is for the model to chose the action with the highest Q-value (exploitation). For $\tau \to \infty$, all actions will be chosen with equal probability (exploration).

In many applications these parameters are hand-tuned and fixed, but ideally they should be dynamically adapted to the environment and the agent's performance. Schweigerhofer and Doya [22], for example, propose a meta-reinforcement learning algorithm to tune α, γ, and τ. For our purpose, the parameters were fixed in order to allow for comparisons between the various experiments.

2.4 Performance Measure

In order to measure the gaze following performance, we will only use the *gaze following index* (GFI) in this paper:

$$\text{GFI} = \frac{\#\text{ gaze shifts from CG to location looked at}}{\#\text{ gaze shifts}}$$

The gaze following index measures the frequency of gaze shifts that lead from the location of the caregiver—where the infant can determine the caregiver's gaze location—to where the caregiver is looking. During our experiments, the learning was suspended every 1000 time steps and the gaze following performance was tested for 1000 time steps before the learning process resumed.

Note that the maximum GFI with the original caregiver as described in Section 2.2 is $\text{GFI}_{\max} = \frac{N}{2(N+1)}$ as the caregiver spends on the average $\frac{1}{N+1}$ of his time looking at the infant. For $N = 10$ regions, $\mathrm{GFI}_{\max}$ becomes 0.45 on the average. For $N = 2$, the maximum GFI drops to 0.33.

2.5 Finding Optimal Infant Parameters by Means of an Evolutionary Algorithm

Carlson *et al.* [6] used an experimentally determined set of parameters which consisted of the following values: $\alpha = 0.0025$, $\gamma = 0.8$, and $\tau = 0.095$. For our caregiver experiments, however, we were interested in using a nearly "optimal" infant as a vis-à-vis and we therefore decided to use an evolutionary algorithm (EA) [13] to find optimal parameters for α, γ, and τ with regards to the speed of learning and the gaze following index at the end of the learning process.

The optimization was performed by means of a standard elitist genetic algorithm with a generation gap of 0.9, a cross-over probability of 0.7, a mutation rate of 0.01, fitness-based reinsertion, and a single-point cross-over operator. The variables α, γ, and τ use a 20 bit representation for the interval $[0.0001, 1]$ on a binary genotype. The fitness function was defined as the average GFI over T simulation time steps: $\text{fitness} = \frac{1}{T}\sum_{t=1}^{T}\mathrm{GFI}(t)$.

We ran the algorithm over ten trials for (1) a population size of $\text{ind} = 15$ individuals, $\text{gen} = 50$ generations, $T = 300,000$ simulation steps (5 trials), and (2) a population size of $\text{ind} = 20$ individuals, $\text{gen} = 100$ generations, $T = 50,000$ simulation steps (5 trials).

The results suggest that the algorithm gets easily stuck in a local optimum (independently of the two trial setups) around $\alpha_2 = 0.5$, $\gamma_2 = 0.03$, $\tau_2 = 0.05$, whereas a global optimum seems to be located around $\alpha_1 = 0.075$, $\gamma_1 = 0.5$, $\tau_1 = 0.007$. We decided to simply retain these two showcase candidates without further investigating the EA's underlying fitness landscape.

Figure 1 shows the evolution of the GFI as a function of four different parameter sets. *Origus* represents the original infant used by Carlson *et al.* [6], *Optimiss* and *Optimuse* are the two parameter sets found by the evolutionary algorithm, whereas *Mediocrus* is an example of a suboptimal solution for comparison (his parameters were experimentally determined).

What can we learn from these results? As one can see, *Optimuse* and *Optimiss* both perform almost equally well and are much faster learners than *Origus* and *Mediocrus*. *Optimuse* learns gaze following slightly faster than *Optimiss*, but both have approximatively the same final GFI. Due to the higher learning rate, his performance slightly oscillates. *Origus* ultimately also reaches almost the same GFI as a the two evolved infants.

Interestingly, although *Optimuse* and *Optimiss* show very similar performance, they have very different parameters. Let us try to characterize their different behaviors in order to understand why they perform almost equally well. *Optimiss* could be described as fast-learning, risky, nearsighted, and exploratory because she has a high learning rate (fast-learning) and a higher temperature (more exploratory) than *Optimuse*. She also takes more "risk" because she prefers immediate rewards (low discount factor). Her smartness, however, allows her to make up for the risk due to the exploratory short-term oriented behavior. *Optimuse*, on the other hand, might be described as tenacious, perspicacious, and exploitive because of her low learning rate and the low temperature. She is interested in getting long-term rewards and therefore closely stands by her policy without taking too much risk and without being seduced by tempting decoys. Interestingly, both of these characteristic personalities perform almost equally well, but with a very different strategy.

Why, one might ask, do *Mediocrus* and *Origus* perform less good? What do they do wrong? According to his parameters, *Mediocrus* is a quick learner, but all the same sticks to his conservative policy (exploitive). One might say that he uses a wrong combination of his capacities and therefore only reaches a GFI of about 0.3. Finally, *Origus* is a really slow learner, but makes otherwise good use of his exploitive behavior and interest in long-term rewards, which ultimately also brings him to the top.

In the remainder of this paper we shall use our two virtual toddlers *Optimuse* and *Optimiss* and their different behavioral strategies to evaluate and compare the performance of various caregivers.

2.6 Analyzing the Parameters T_{fix}, p_{shift}, and p_{valid}

In order to analyze the influence of the caregiver and the environment on the infant's gaze following behavior we ran several simulations with different parameter sets for the caregiver as described in Section 2.2—who shall be baptized *Ancestrus*—while the infant's parameters were that of *Optimuse* and *Optimiss* (see previous section).

Figure 2 shows *Ancestrus'* average gaze following index as a function of each of the three parameters p_{valid}, p_{shift}, and T_{fix}. The parameter values (two were fixed during each run) were as following: $p_{\mathrm{shift}} = 0.5$, $p_{\mathrm{valid}} = 0.75$, and $T_{\mathrm{fix}} = 4$.

As one can see, the predictiveness of the caregiver's gaze p_{valid} is the sole parameter which significantly affects the infant's GFI, i.e., for $p_{\mathrm{valid}} < 0.5$, the GFI begins to dramatically drop for both toddlers because

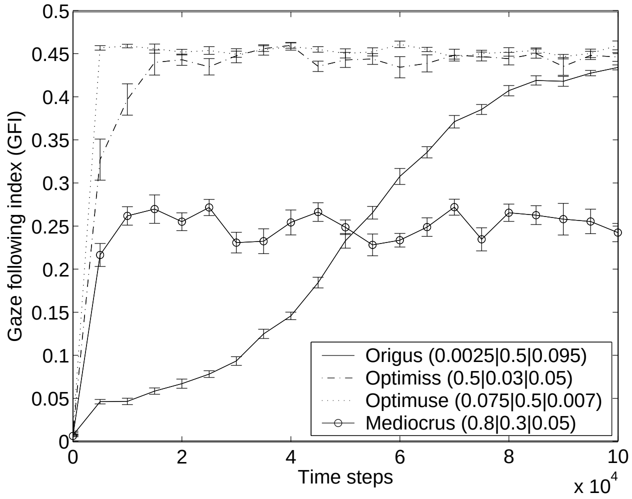


Figure 1. Learning curves for different α, γ, and τ. Average GFI over 10 runs with standard error. The evolved infants *Optimiss* ($\alpha_2 = 0.5$, $\gamma_2 = 0.03$, $\tau_2 = 0.05$) and *Optimuse* ($\alpha_1 = 0.075$, $\gamma_1 = 0.5$, $\tau_1 = 0.007$) perform almost equally well and better than *Origus*. Infant *Mediocrus* is suboptimal.

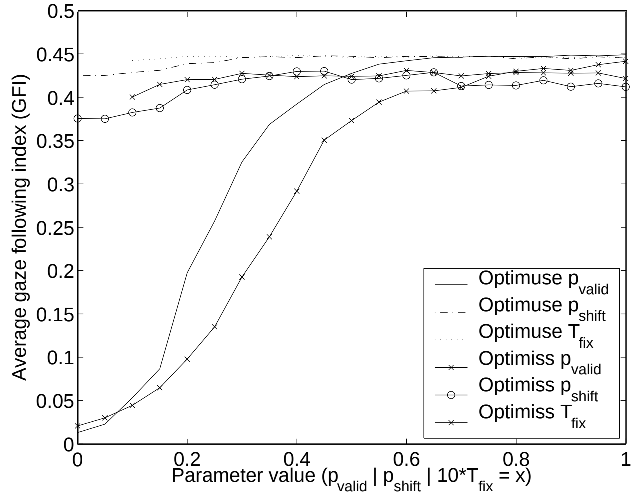


Figure 2. Average GFI as a function of p_{shift}, p_{valid}, and T_{fix} over $T = 100,000$ time steps for *Optimuse* and *Optimiss*. The fixed values were: $p_{\text{shift}} = 0.5$, $p_{\text{valid}} = 0.75$, and $T_{\text{fix}} = 4$. Average GFI over 4 runs without standard error for better legibility.

the caregiver gradually becomes less predictive. On the other hand, T_{fix} and p_{shift} almost have no influence on the GFI, at least not when modified individually[1]. *Optimiss* has a slightly lower GFI for all three caregiver parameter sweeps and the GFI begins to drop earlier than for *Optimuse*, which suggests that she is more sensitive to the caregiver's behavior.

Figure 3 shows four plots for characteristic parameter sets. With *Ancestrus* as a vis-à-vis, *Optimuse* quickly learns to follow the caregiver's gaze. Caregiver *Ancestrus I* performs less good with $p_{\text{valid}} = 0.2$ because he less often looks at the object. *Ancestrus II* has the same p_{valid} as *Ancestrus* but a relocation probability of $p_{\text{shift}} = 1$ and an object fixation time of $T_{\text{fix}} = 1$, which prevents the infant from learning gaze following because the object is relocated et every time step and the caregiver is "nervously" shifting gaze. Finally, although *Ancestrus III* has a fixation time of $T_{\text{fix}} = 1$ only and a low p_{shift}, the infant learns gaze following fairly well due to $p_{\text{valid}} = 0.3$ (compare also with the average GFI of Figure 2).

Simulations for *Optimiss* show a different picture for *Ancestrus I* and *Ancestrus III*: they have a much lower GFI (average of about 0.1 and 0.15 respectively) than for *Optimuse*, which confirms the foregoing finding that she is more sensitive to the caregiver's behavior and well adapted to *Ancestrus* only.

[1]Note that the minimal T_{fix} is 1, i.e., the object moves at each time step if $p_{\text{shift}} = 1$.

It can be concluded—what seems intuitively obvious for this setting and these toddlers—that the more predictive and structured the environment is (i.e., the higher p_{valid}), the better and faster the infant learns gaze following. Thereby, the predictiveness of the caregiver plays a crucial role, whereas the object's behavior is less important.

In the next section we will analyze four new caregivers and compare them with our two toddlers in combination with two developmental disorders.

3 New Caregiver Models

The caregiver used so far did not adapt to the infant's behavior, which certainly represents a gross oversimplification of a real mother-infant interaction. One might hypothesize that an adaptive caregiver would allow the infant to learn gaze following faster and more reliably. In order to test this hypothesis, we implemented several caregiver models and environments and evaluated the gaze following performance of normally developing and autistic infants as well as of children with the Williams syndrome. The caregivers used are as following:

- *Ancestrus:* Original caregiver as described in Section 2.2.

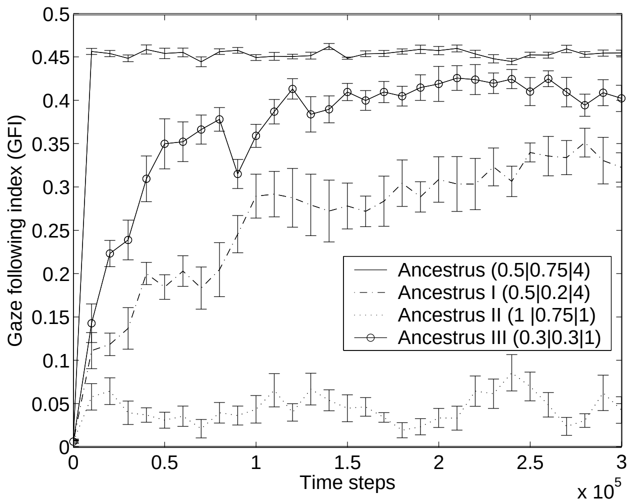


Figure 3. *Optimuse's* **learning curves for four different caregiver parameter sets** p_{shift}**,** p_{valid}**, and** T_{fix}**. Average GFI over 10 runs with standard error,** $\alpha = 0.075$**,** $\gamma = 0.5$**,** $\tau = 0.007$**.**

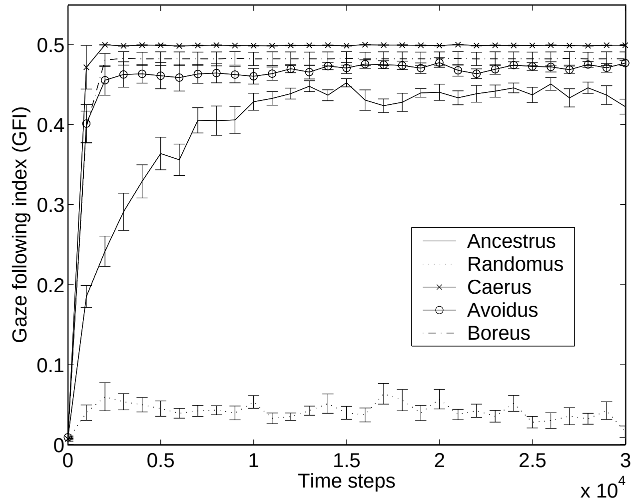


Figure 4. *Optimuse's* **GFI for the new caregivers. Average over 10 runs with standard error,** $\alpha = 0.075$**,** $\gamma = 0.5$**,** $\tau = 0.007$**.**

	Optimuse			*Optimiss*		
avg GFI	N	A	W	N	A	W
Ancestrus	0.44	0.004	0.05	0.41	0	0.38
Randomus	0.04	0	0.05	0.04	0	0.04
Careus	0.5	0	0.5	0.5	0	0.5
Avoidus	0.48	0.27	0.42	0.43	0.05	0.5
Boreus	0.47	0.45	0.24	0.49	0	0.5

Table 1. Summary of the average GFI for the two toddlers facing the new caregivers. Average values over $T = 300,000$ **time steps and over two runs. N=Normal, A=Autist, W=Williams syndrome.**

- *Randomus:* Random caregiver and object, i.e., at each time step the object and the caregiver's gaze are individually moved to a random location. The caregiver therefore looks at the infant with probability $\frac{1}{N+1}$ and even less rarely looks at the infant.
- *Careus:* This caregiver waits until he can establish mutual gaze contact with the infant. He then moves the object to a random location and directs his gaze to the very same location where he waits until the infant looks at the object. The caregiver then returns his gaze to the infant and waits again for mutual gaze contact.
- *Avoidus:* Same as *Ancestrus*, but he never looks at the infant.
- *Boreus:* The object and the caregiver's gaze both move together, stepwise, and indefinitely from location 1 to N and back, one step each $T_{fix} = 4$ time steps.

Figure 4 shows *Optimuse's* gaze following performance vis-à-vis of the five above described caregivers. One can see that the random caregiver *Randomus* does not provide a sufficiently structured environment to the infant, whereas *Ancestrus* only allows the infant to slowly acquire gaze following. *Careus* performs best as he adapts to the infant and "guides" his gaze. *Avoidus* still allows the infant to learn gaze following very well, although he never establishes mutual gaze contact. Similarly, *Boreus* provides a highly structured environment because of the deterministic object trajectory and his fully predictive gaze direction. Columns "N" (normal) of Table 1 summarize the simulation results for this experiment for both infants. As one can see, the results are very similar for *Optimiss*.

3.1 Developmental Disorders

Some autistic children show little or no eye contact and tend to avoid looking at faces whereas children with the Williams syndrome have an abnormally high preference for faces. Note that there seems to exist controversial evidence whether autistics perceive direct gaze as aversive or not. Carlson *et al.* [6] have demonstrated that simple changes in the infant's reward structure can lead to behaviors reminiscent of autism and the Williams syndrome. In this paper we used the following reward structures:

- **Autist:** $R_{\text{frontal}} = -1$, $R_{\text{profile}} = 0$;
- **Williams syndrome:** $R_{\text{frontal}} = 2$, $R_{\text{profile}} = 2$.

Columns "A" (autist) and "W" (Williams syndrome) of Table 1 summarize the simulation results for this experiment for both infants. In addition, Figure 5 illustrates the GFI of *Optimuse* as an autist and as a toddler with Williams syndrome. She faces *Ancestrus*, *Avoidus* for the autist, and *Careus* for the Williams toddler. As one can see, gaze following does not emerge with *Ancestrus* and the autist, whereas the Williams toddler only learns it badly. *Avoidus*, however, helps the autist to successfully learn gaze following because he avoids mutual gaze contact (i.e., avoids negative rewards for the infant), whereas *Careus* succeeds in "guiding away" the Williams toddler from staring at his face to the object.

From Table 1 we can also see that the random caregiver *Randomus* is unsuccessful in all situations. Looking at *Optimuse*, we find that *Boreus* performs even better than *Avoidus* for the autist because he too never looks at the infant, but provides an even more deterministic behavior. For the same infant with Williams syndrome, *Avoidus* performs also very well because he avoids mutual gaze.

The situation is a little different for *Optimiss*. In her autistic version, she prevents any gaze following to evolve whereas she does a much better job as a "Williams infant": all caregivers, except *Randomus*, succeed in "teaching" her gaze following. This is somehow surprising as she was the one who was more sensitive to the environment. However, this might exactly be the explanation: she is more likely to be disturbed and drifted away, which is beneficial for a Williams toddler, i.e., to shift attention away from the highly attracting face.

4 Conclusions

We presented a computational gaze following framework as first introduced by Carlson *et al.* [6] and optimized the infant's parameters by means of an evolutionary algorithm. The outcome were two infants which performed almost equally well, but used different strategies. We then analyzed the original and several new caregiver models and showed that the caregiver behavior plays a crucial role in the development of gaze following, especially for virtual infants with "developmental disorders". A second finding is that no "universally" optimal caregiver exists: every virtual toddler—whether with a developmental disorder or not—has its own needs and requires a particular caregiver. This immediately suggests that the ideal caregiver should itself be a learning agent that constantly adapts to the developing infant.

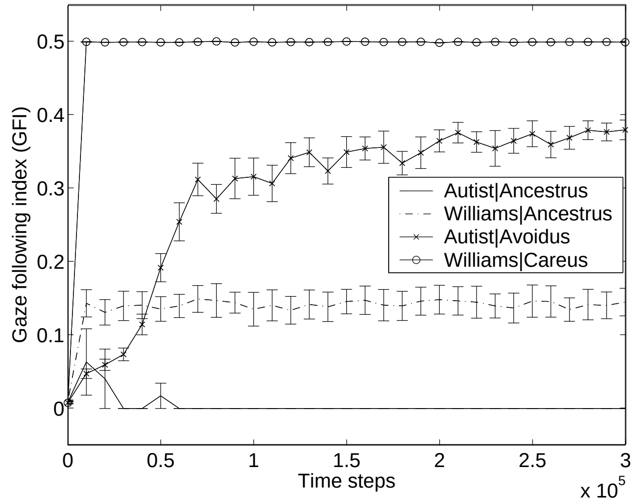


Figure 5. GFI of disordered *Optimuse* **who faces caregivers** *Ancestrus*, *Avoidus*, **and** *Careus*. **Average over 10 runs with standard error,** $\alpha = 0.075$, $\gamma = 0.5$, $\tau = 0.007$.

This finding might seem intuitively obvious and simply goes into the direction of developing more realistic computational models of the emergence of gaze following, and eventually shared attention. Nevertheless, the current simplicity of the model can also be considered as a strength since it brings the computational essence of the underlying mechanisms into focus.

Acknowledgments

This work described here is part of the UC San Diego MESA project (`http://mesa.ucsd.edu`) and is supported by the National Alliance for Autism Research (NAAR). Christof Teuscher was supported by the Swiss National Science Foundation under grant PBEL2-104420. The authors are grateful to Gedeon Deák, Hector Jasso, Boris Lau, and Ursina Teuscher for their helpful comments and discussions.

References

[1] S. Baron-Cohen. Development of a theory of mind: Where would we be without the intentional

stance? In M. Rutter and D. Hay, editors, *Development Through Life: A Handbook for Clinicians*, pages 303–308. Blackwell Scientific, Oxford, UK, 1994.

[2] S. Baron-Cohen. The eye direction detector (EDD) and the shared attention mechanism (SAM: Two cases for evolutionary psychology). In C. Moore and P. J. Dunham, editors, *Joint Attention: Its Origins and Role in Development*, chapter 6, pages 41–59. Erlenbaum, Hillsdale, NJ, 1995.

[3] C. Breazeal and B. Scassellati. Infant-like social interactions between a robot and a human caregiver. *Adaptive Behavior*, 8(1):49–74, 2000.

[4] J. Bruner. *Child's Talk*. Norton, New York, 1983.

[5] G. Butterworth and N. Jarrett. What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9:55–72, 1991.

[6] E. Carlson and J. Triesch. A computational model of the emergence of gaze following. In H. Bowman and C. Labiouse, editors, *Proceedings of the 8^{th} Neural Computation and Psychology Workshop*, Progress in Neural Processing. World Scientific, 2003.

[7] M. R. A. Chance. Attention structure as the basis for primate rank orders. *Man*, 2(503–518), 1967.

[8] M. Cole and S. Cole. *The Development of Children*. Freeman, New York, 3 edition, 1996.

[9] G. O. Deák, R. Flom, and A. D. Pick. Perceptual and motivational factor affecting joint visual attention in 12- and 18-month-olds. *Developmental Psychology*, 36(511–523), 2000.

[10] N. J. Emery. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 24:581–604, 2000.

[11] I. Fasel, G. O. Deák, Jochen Triesch, and J. Movellan. Combining embodied models and empirical research for understanding the development of shared attention. In *Proceedings of the 2^{nd} Internantional Conference on Development and Learning (ICDL02)*, pages 21–27, Los Alamitos, CA, 2002. IEEE Computer Society Press.

[12] J. M. Findlay and R. Walker. A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, 22:661–674, 1999.

[13] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Redwood City, CA, 1989.

[14] K. Kaye. *The Mental and Social Life of Babies*. University of Chicago Press, Chicago, 1982.

[15] B. Lau and J. Triesch. Learning gaze following in space: A computational model. Third International Conference on Development and Learning (ICDL'04), La Jolla, California, October 20–22, 2004.

[16] G. Matsuda and T. Omori. Learning of joint visual attention by reinforcement learning. In E. M. Altmann and A. Cleeremans, editors, *Proceedings of the Fourth International Conference on Cognitive Modeling*, pages 157–162, Mahwah, NJ, US, 2001. Lawrence Erlbaum Associates.

[17] C. Moore. Theories of mind in infancy. *British Journal of Developmental Psychology*, 14(19–40), 1996.

[18] C. Moore and P. J. Dunham, editors. *Joint Attention: Its Origins and Role in Development*. Erlenbaum, Hillsdale, NJ, 1995.

[19] Y. Nagai, K. Hosoda, A. Morita, and M. Asada. A constructive model for the development of joint attention. *Connection Science*, 15(4):211–229, December 2003.

[20] M. Scaife and J. S. Bruner. The capacity of joint visual attention in the infant. *Nature*, 253:265–266, 1975.

[21] B. Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, 12:13–24, 2002.

[22] N. Schweighofer and K. Doya. Meta-learning in reinforcement learning. *Neural Networks*, 16:5–9, 2003.

[23] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, MIT Press, Cambridge, MA, 1998.

[24] M. Tomasello and J. Todd. Joint attention and early lexical acquisition style. *First Language*, 4:197–212, 1983.

Joint attention between a humanoid robot and users in imitation game

Masato Ito
Sony Corporation
6-7-35 Kitashinagawa, Shinagawa-ku
Tokyo, 141-0001, Japan
masato@pdp.crl.sony.co.jp

Jun Tani
Brain Science Institute, RIKEN
2-1 Hirosawa, Wako-shi
Saitama, 351-0198, Japan
E-mail tani@brain.riken.go.jp

Abstract

This paper studies a synthetic model for joint attentions and turn taking by conducting experiments of robot-user imitative interactions. A Sony humanoid robot learns multiple cyclic movement patterns as embedded in a neural network model proposed by us and each of memorized movement patterns can be retrieved and generated by means of entrainment by external inputs in terms of users' movement patterns perceived. Our experiments on a simple imitation game showed that multiple movement patterns are generated as synchronized between the robot and users while the shared patterns shift spontaneously from one to another. Based on the experimental results, we show possible dynamical systems accounts for the underlying mechanisms for joint attentions and turn taking.

1 Introduction

In entertainment robotics, achieving natural interactions between robots and their users is one of the most essential issues to be solved. The ultimate goal for this is to develop communicative abilities on robots like humans. Human communications involve dynamic processes such as joint attention and turn taking with others. Joint attention is to share behaviors, events, interests and contexts in the world among agents from time to time. It requires mutual awareness of companion's attentions. On the other hand, turn taking is to switch the initiatives in interactions among agents spontaneously. Turn taking is considered to be prerequisite for joint attention. Speculating that such dynamic processes might appear as a consequence of mutual adaptation among agents, we develop a synthetic model for experiments of interactions between robots and their users.

Recent research on robotics have implemented a model of joint visual attention [3] between robots and humans [9, 11]. In such models, the robot guess the human's attentional target by detecting their gazing and pointing, and also pays attention to it. And then joint attention can be archived by the recognition of the robot's attention by human. However, in human communications, it seems that there are more complex situations of joint attention that can never be achieved by simply using such static and explicit cues [8]. For example, to share topics in dialogues and to share goals in collaborative works. It seems that the targets of such joint attention are determined in the flow of ongoing interactions in contextual ways. We speculate that such context dependent communicative interactions could emerge in terms of a class of dynamical structures appeared in the mutual adaptation processes between robots and humans.

In this study, we assume simple movement imitation game between a robot and human subjects where the problems of imitations are simplified from the reality. The imitation in our robot platform is not yet goal-oriented ones as have been discussed by [17]. Also the correspondence problems [4] between the perceptual space for others and motor space of own in learning are simplified. Rather, our focus is to observe dynamic interaction processes which take place in the coupling between robots and human in the simplified imitation game.

Firstly in this paper, we will introduce our neural network model: recurrent neural network with parametric biases (RNNPB) [15, 16]. The robot learns multiple cyclic movement patterns as embedded distributedly in self-organized dynamic structures of the RNNPB. Then, each of memorized movement patterns can be regenerated by means of entrainment by a corresponding users' movement pattern perceived. This is done by the on-line adaptation scheme of the parametric biases (PB). Then, two types of imitative interactions will be demonstrated using this model. In the first experiment, the on-line adaptation only in the robot side is considered in the imitation game. In the second experiment, the on-line adaptation in both of the robot and the users is performed. By going through the close examinations of different aspects in these two experiments, a novel theory for joint attentions and turn taking will be elucidated.

2 Task setting and neural network modeling

2.1 Task setting

In the current study, the Sony humanoid robot QRIO (SDR-4X II) [5] was used as the experimental platform (see Figure 1). In this experiment, only movement patterns of both arms were considered. Other movements were frozen.

Figure 1. A user is interacting with the Sony humanoid robot QRIO SDR-4XII.

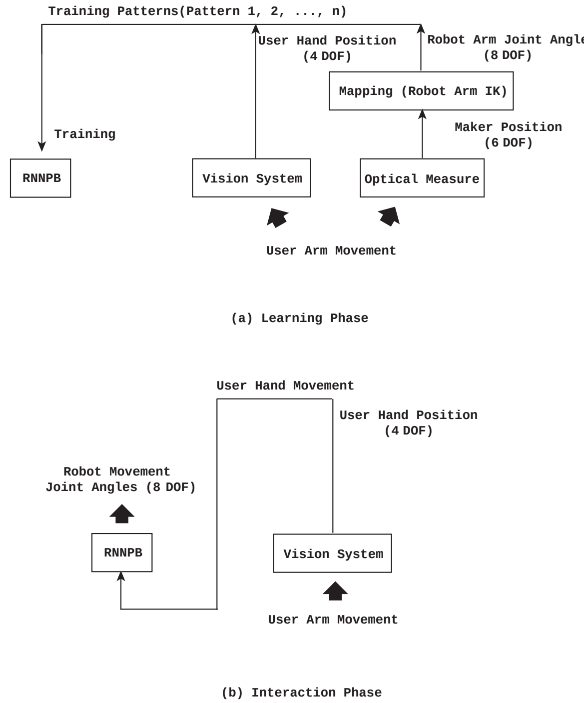


Figure 2. System configurations in learning phase (a) and interaction phase (b).

Before interaction experiments with users, the robot learns a set of robot movement patterns with different profile with associated with the corresponding user's visually perceived hand movement patterns as off-line in the learning phase. It is actually done in the form of sequence prediction learning for sensory-motor flow as will be detailed in later.

In the learning phase, the target trajectories of the robot are obtained by mapping the user's arm position to the robot joint angles. This mapping was conducted using the following engineering scheme. First, the user's hands' spatial coordinates were optically measured by tracking colored balls on the user's hands. (The depth information was obtained by using the measured size of the ball.) The obtained spatial coordinates of the user's hands are simply mapped to the robot's hand's 3-D positions in robot centered cartesian coordinates. Next, they are mapped to the robot joint angles (shoulder roll, pitch, yaw, and elbow pitch for each arm) by solving the inverse kinematics of the robot arm, assuming the constraint that elbow pitch is dependent on shoulder pitch. Note that this 3-D measuring is utilized only for generating the motor trajectories for the training data, and not used in the interaction phase.

As summarized in Figure 2(a), the learning process utilizes the paired trajectories of the robot joint angles, obtained by the mapping, and the user's hand positions, as visually perceived by the robot. The training of the employed neural network model (RNNPB) is conducted by using a set of training patterns, corresponding to multiple robot and user movement patterns.

In the interaction phase, the robot attempts to follow synchronously the user's hand movement patterns with predicting their sequences. As shown in Figure2(b), the robot perceives the user's hand movement patterns visually and generates its corresponding movement patterns in robot joint angles. The robot's ability to follow the user depend on the degree to which the user patterns are familiar to the robot, based on prior learning.

It is important to note that the robot learns not just a static mapping from user hand positions to robot joint angles. Instead, the robot learns the intrinsic dynamics hidden in the target movement patterns. Actually, the robot can generate its movement patterns autonomously without perceiving the user's hand movements, but by imagining it by means of a prediction mechanism, as will be described later. The perception of the hand movement patterns just triggers regeneration of the corresponding dynamic patterns of the robot movement. The underlying mechanism of how the perceptual patterns can trigger the generation of the motor patterns will be the focus of the current study.

Next, the employed neural network model (RNNPB) is explained.

2.2 RNNPB modeling

RNNPB is a version of the Jordan-type RNN [7] where the PB units allocated in the input layer play the roles of mirror neurons since their values encode both of generating and recognizing the same movement patterns. In generating patterns, the PB values function as control parameters for modulating the forward dynamics of the RNN. On the other hand in recognizing patterns, the corresponding PB values for currently perceiving patterns can be dynamically obtained by using the inverse dynamics of the RNN. It is, however, important to note that these recognition and generation processes are conducted simultaneously in the interaction phase i.e.– the robot generates corresponding patterns while recognizing the user's movement patterns. These ideas are detailed in the following associated with descriptions of the learning scheme.

A set of movement patterns is learned, in terms of the forward dynamics of the RNNPB, by self-determining both the PB values, that are differently assigned for each movement pattern, and a synaptic weight matrix that is common for all patterns. The information flow of the RNNPB in the learning phase is shown in Figure 3(a). This learning is conducted using both target sequences of the robot joint angles r_t and the user's hand positions u_t. With given r_t and u_t in the input layer, the network predicts their values at the next time step in the output layer as $\hat{r_{t+1}}$ and $\hat{u_{t+1}}$. The outputs are compared with their target values r_{t+1} and u_{t+1} and the error generated is back-propagated [12] for the purpose of updating both the synaptic weights and PB values. Note that the determined synaptic weights are common to all learning patterns, but the PB values are differently determined for each pattern. This scheme will be described in more detail later. c_t represents the context units where the self-feedback loop is established from c_{t+1} in the output layer to c_t in the input layer. The context unit activations represent the internal state of the network.

In the interaction phase, the pre-learned network is utilized without updating the synaptic weights. While the forward dynamics of the RNNPB generates the prediction of the sensory-motor sequences, the PB values are inversely computed by utilizing the error information obtained between the sensory prediction and the outcome. See Figure 3(b) for the information flow of the network in the interaction phase. The visually perceived hand positions are fed into the RNNPB as the target sequences. The RNNPB, when receiving u_t, attempts to predict its next value $\hat{u_{t+1}}$ in the outputs. The generated prediction error from the target value u_{t+1} in the outputs is back-propagated to the PB units and the PB values are updated in the direction of minimizing the error. Note that although the PB plays the role of the inputs for the forward computation, its values are slowly modulated in order to adapt to the current target sequence patterns. If pre-learned hand movement patterns are perceived, the PB values tend to converge to the values that have been determined in the learning phase while minimizing the prediction error. It is guaranteed that by minimizing the prediction error to zero the forward dynamics does not modulate anymore since the PB values converge. Then, the network becomes able to generate the associated motor patterns $\hat{r_{t+1}}$ as previously learned. The robot movement patterns are generated based on the PB values while these values are adapted by perceiving the hand movement patterns. An interesting feature of this model is that generation and perception are performed simultaneously in one neural dynamic system.

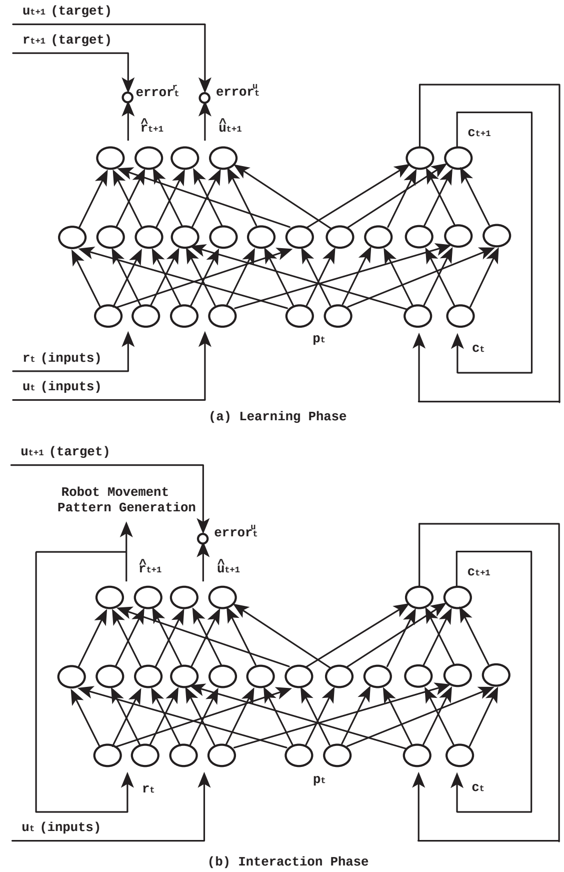


Figure 3. The system flow of RNNPB in learning phase (a) and interaction phase (b).

In the next section, the computational algorithm for modifying the PB values is reviewed.

2.3 Computational algorithm

The PB values are determined through regression of the past sequence pattern. In the interaction phase, the regression is applied for the immediate past window steps L and the temporal profile of p_t from L steps before to the current step ct is updated. Then, the current time motor outputs r_{ct} are generated by using the p_{ct-1} determined by this regression process. The window for the regression shifts as time goes by while p_t is updated through the iterations. In the learning phase the regression is conducted for all steps of the training sequence patterns. (This means that the window contains the whole sequence and it does not shift.)

The temporal profile of p_t in the sequence is computed via the back-propagation through time (BPTT) algorithm [12]. In this computation ρ_t, the internal value of the parametric bias, is obtained first. The internal value ρ_t changes due to the update computed by means of the error back-propagated to this parametric bias unit, which is integrated for a specific step length in the sequence. Then the parametric bias, p_t, is obtained by a sigmoid function of the output of the internal value. The utilization of the sigmoid function is just a computational device to bound the value of the parametric bias to a range of 0.0 to 1.0. In this way, the parametric bias is updated to minimize the error between the target and the output sequence.

For each iteration in the regression of the window, L steps of look-ahead prediction, starting from the onset step of the window, are computed by the forward dynamics of the RNN. Once the L steps of the prediction sequence are generated, the errors between the targets and the prediction outputs are computed and then back-propagated through time. The error back-propagation updates both the values of the parametric bias at each step and the synaptic weights. The update equations for the ith unit of the parametric bias at time t in the sequence are:

$$\delta{\rho_t}^i = k_{bp} \cdot \sum_{step=t-l/2}^{t+l/2} {\delta_{step}^{bp}}^i + k_{nb}(\rho_{t+1}^i - 2\rho_t^i + \rho_{t-1}^i) \quad (1)$$

$$\triangle\rho_t^i = \epsilon \cdot \delta{\rho_t}^i + \eta \cdot \triangle\rho_{t-1} \quad (2)$$

$$p_t^i = sigmoid(\rho_t) \quad (3)$$

In Eq. (1), $\delta\rho_t$, the delta component of the internal value of the parametric bias unit, is obtained from the summation of two terms. The first term represents the summation of the delta error, ${\delta_t^{bp}}^i$, in the parametric bias units for a fixed time duration l. ${\delta_t^{bp}}^i$, which is the error back-propagated from the output units to the ith parametric bias unit, is summed over the period from $t - l/2$ to $t + l/2$ time step. By summing the delta error, the local fluctuations of the output errors will not affect the temporal profile of the parametric bias significantly. The parametric bias should vary only with structural changes in the target sequence. Otherwise it should become flat, or constant, over time. The integration period, l, is taken as 20 steps in the experiment which is close to the time constant of the movement patterns in the training set.

The second term plays the role of a low pass filter through which frequent rapid changes of the parametric bias are inhibited. k_{nb} is the coefficient for this filtering effect. ρ_t is updated based on $\delta\rho_t$ obtained in Eq. (1). The actual update $\triangle\rho_t$ is computed by utilizing a momentum term to accelerate convergence as shown in Eq. (2). Then, the current parametric bias p_t is obtained by means of the sigmoidal outputs of the internal values ρ_t in Eq. (3).

In the interaction phase, the window step length for the regression L is taken as 30 steps. The regression, by means of the forward computation, and the error back-propagation iterates about 100 times in the real-time computation while the window shifts one step ahead. In the learning phase, the regression is iterated 50000 times for the fixed window containing the whole training sequence.

3 Experiments on imitation game

Two types of imitation game experiments are conducted using the proposed model. In the Experiment-1, the imitation game is set such that the on-line adaptation is conducted only in the robot side. In the Experiment-2, mutual adaptations between the robot and human subjects are conducted in the imitation game.

3.1 Experiment-1: robot adaptation only

In the Experiment-1, the robot learns three movement patterns shown by user's hand movements in the learning phase. In the interaction phase, we examined how the robot could follow target patterns while the user switched to demonstrate among various learned patterns.

The results of the experiment are plotted in Figure 4. It is observed that when the user hand movement pattern is switched from one pattern to another, the patterns in the sensory prediction and the motor outputs are also switched correspondingly by accompanying substantial shifts in the PB vector. Although the synchronization between the user hand movement pattern and the robot movement pattern is lost once during the transitions, the robot movement pattern is re-synchronized to the user hand movement pattern within several steps.

3.2 Experiment-2: mutual imitation game

The previous experiments focused mainly on the adaptation in the robot side. We conducted another experiment which focus on bi-directional adaptation in mutual interaction between the robot and users. In this new experimental set-up, after the robot learns 4 movement patterns in the same way as described previously, subjects who are ignorant of what the robot learned are faced with the robot. The subjects are then asked to find as many movement patterns as possible for which they and the robot can synchronize together by going through exploratory interactions. Five subjects participated in the experiments. The settings of the network and the robot were exactly the same as those in the previous interaction experiments. Each subject was allowed to explore the interactions with the robot for one hour, including four 5 minute breaks. Although most of the subjects could find all movement patterns by the end, the exploration processes were not trivial for the subjects. If the subjects merely attempted to follow the robot movement patterns, they could not converge in most situations since the PB values fluctuated when receiving unpredictable subject hand movement patterns as the inputs. If the subjects attempted to execute their desired movement patterns regardless of the robot movements, the robot could not follow them unless the movement patterns of the subjects corresponded with the ones learned by the robot.

One example of the interaction in imitation game is plotted in Figure 5.

It is observed that joint attention to a certain movement pattern between the robot and the subject as synchronization is achieved after some exploratory phase. It is also observed that this joint attentional state is break down once but joint attention to another pattern is achieved again.

There are interesting points in this new experiment as compared to the previous one. First, the master-slave relation, which was fixed between the subjects and the robot in the previous experiments, is no longer fixed but is instead spontaneously switched between the two sides. (Recall that the subjects initiated new movement patterns while also switching among multiple learned patterns in the previous experiments.) When the subjects feel that the robot movement patterns become close to theirs, they just keep following the robot movement patterns passively in order to stabilize the patterns. However, when the subjects feel that they and the robot cannot match each other's movements, they often initiate new patterns, hoping that the robot will start to follow them and become synchronized. Second, there are autonomous shifts between the coherent phase and the incoherent phase after the subjects become familiar with the robot responses to some extent. When the subjects happen to find synchronized movement patterns, they tend to keep the achieved synchronization for a moment in order to memorize the patterns. However, this coherence can break down after a while through various uncertainties in the mutual interactions. Even small perturbations in the synchronization could confuse the subjects if they are not yet fully confident of the repertoire of the robot's movement patterns. Also, the subjects' explorations of new movement patterns makes it difficult for the robot to predict and follow their movements.

4 Discussion

The authors speculate that appropriate analysis of these observed phenomena might shed a ray of light on the mechanism of joint attention [2, 10] as well as turn taking behaviors [18]. In our new experiment, when movement patterns of the robot and human are synchronized, joint attention is assumed to have been achieved for the pattern. However, the current joint attention can break down and another joint attention (attending to another movement pattern) can emerge after a while. Although joint attention itself might be explained simply by synchronization [1], a more interesting question is how a joint attention can break down and flip to another one spontaneously. This sort of spontaneity is also essential in turn taking behaviors. It was observed that the initiatives leading to synchronization switch spontaneously between the robot and the subjects. The essential question here is how the spontaneous shifts in turn taking behaviors can emerge.

Although extensive analysis of the observed data is required for further reasoning of the underlying mechanisms, the authors speculate that they might be closely related to the so-called open dynamic structure [14]. It was argued that the system state tends to flip between the coherent and the incoherent phases if stability, in terms of rational goal-directedness, and instability, caused by unpredictability of the open environment, coexist in cognitive systems. Tani [14] proposed one possible explanation of the spontaneous breakdown of self-consciousness through dynamic system characteristics. A more theoretical framework of this idea has been explained by the chaotic itinerary [19]. Furthermore, Ikegami and Iizuka [6] recently showed that spontaneous turn taking behaviors can emerge by evolving the coupled-dynamics for a simulated pair of agents. Their analysis indicated that both stable and unstable manifolds are generated in the evolved coupled dynamics. In our experiments of mutual interactions, the stability originated from the synchronization mechanisms for shared memories of movement patterns between the robot and the subjects. The instability arose from the potential uncertainty in predicting each other's movements. It is likely that the coexistence of stable and unstable characteristics in the system dynamics might be the main cause for the spontaneous shifts. Recently, Sato [13] related this characteristics to the undecidability of the turing test in the theoretical analysis of imitation game, although further examination is required in this part of the analysis. Future collaborative research among developmental psychology, synthetic modeling studies, and theoretical nonlinear dynamics studies would gain further understanding of the essential mechanisms in joint attention and turn taking behaviors.

In the mutual interaction experiments, most of the subjects reported that they occasionally felt as if the robot had its own “will” because of the spontaneity in the generated interactions. It is speculated that the spontaneity originated from the total system dynamics including the users in the loop might play an important role in attracting people to play with entertainment robots.

5 Summary

Our experiments with a humanoid robot have shown that diverse dynamic interactions can emerge in the form of either coherence or incoherence between the robot and the user. The robot can follow the learned user movement patterns synchronously by generating coherent dynamic states. It can be said that joint attention is accomplished for the current movement pattern shared in both the memories of the robot and the user. Our experiments of the mutual adaptation suggest that the essential mechanism for autonomous shifts in joint attention and turn taking behavior could be explained by the open dynamic structures in which stability, in terms of rational goal-directedness, and instability, caused by unpredictability of others coexist.

References

[1] P. Andry, P. Gaussier, S. Moga, J. Banquet, and J. Nadel. Learning and communication in imitation: An autnomous robot perspective. *IEEE Transaction on Systems, Man and Cybernetics. Part A : Systems and Humans*, 31(5):431–444, 2001.

[2] S. Baron-Cohen. *Mindblindness: An essay on autism and theory of mind*. MIT Press., 1996.

[3] G. E. Butterworth and N. L. M. Jarrett. What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9:55–72, 1991.

[4] K. Dautenhahn and C. L. Nehaniv, editors. *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, 2002.

[5] M. Fujita, Y. Kuroki, T. Ishida, and T. Doi. A small humonoid robot sdr-4x for entertainment applications. In *Proceedings of International Conference on Advanced Intelligent Mechatronics*, 2003.

[6] T. Ikegami and H. Iizuka. Joint attention and dynamics repertoire in coupled dynamical recognizers. In *the AISB 03: the Second International Symposium on Imitation in Animals and Artifacts*, pages 125–130, 2003.

[7] M.I. Jordan and D.E. Rumelhart. Forward models: supervised learning with a distal teacher. *Cognitive Science*, 16:307–354, 1992.

[8] F. Kaplan and V.V. Hafner. The challenges of joint attention. In *Proceedings of 4th International Workshop on Epigenetic Robotics*, 2004.

[9] H. Kozima and H. Yano. A robot that learns to communicate with human caregivers. In *First International Workshop on Epigenetic Robotics*, 2001.

[10] C. Moore and V. Corkum. Social understanding at the end of the first year of life. *Developmental Review*, 14(4):349–450, 1994.

[11] Y. Nagai. *Understanding the Development of Joint Attention from a Viewpoint of Cognitive Developmental Robotics*. PhD thesis, Osaka University, Japan, 2004.

[12] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning internal representations by error propagation. In *Parallel Distributed Processing*. MIT Press, 1986.

[13] Y. Sato and T. Ikegami. Undecidability in the imitation game. *Minds and machines*, 14:133–143, 2004.

[14] J. Tani. An interpretation of the ”self” from the dynamical systems perspective: a constructivist approach. *Journal of Consciousness Studies*, 5(5-6):516–42, 1998.

[15] J. Tani. Learning to generate articulated behavior through the bottom-up and the top-down interaction process. *Neural Networks*, 16:11–23, 2003.

[16] J. Tani and M. Ito. Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment. *IEEE Transactions on System, Man and Cybernetics Part A*, 33(4):481–488, 2003.

[17] M. Tomasello. *The caltural origins of human cognition*. Harvard University Press., 1999.

[18] C. Trevarthen. Descriptive analyses of infant communicative behaviour. In H.R. Schaffer, editor, *Studies in Mother-Infant Interaction*. Academic Press, London, 1977.

[19] I. Tsuda and T. Umemura. Chaotic itinerancy generated by coupling of milnor attractors. *Chaos*, 13(3):937–946, 2003.

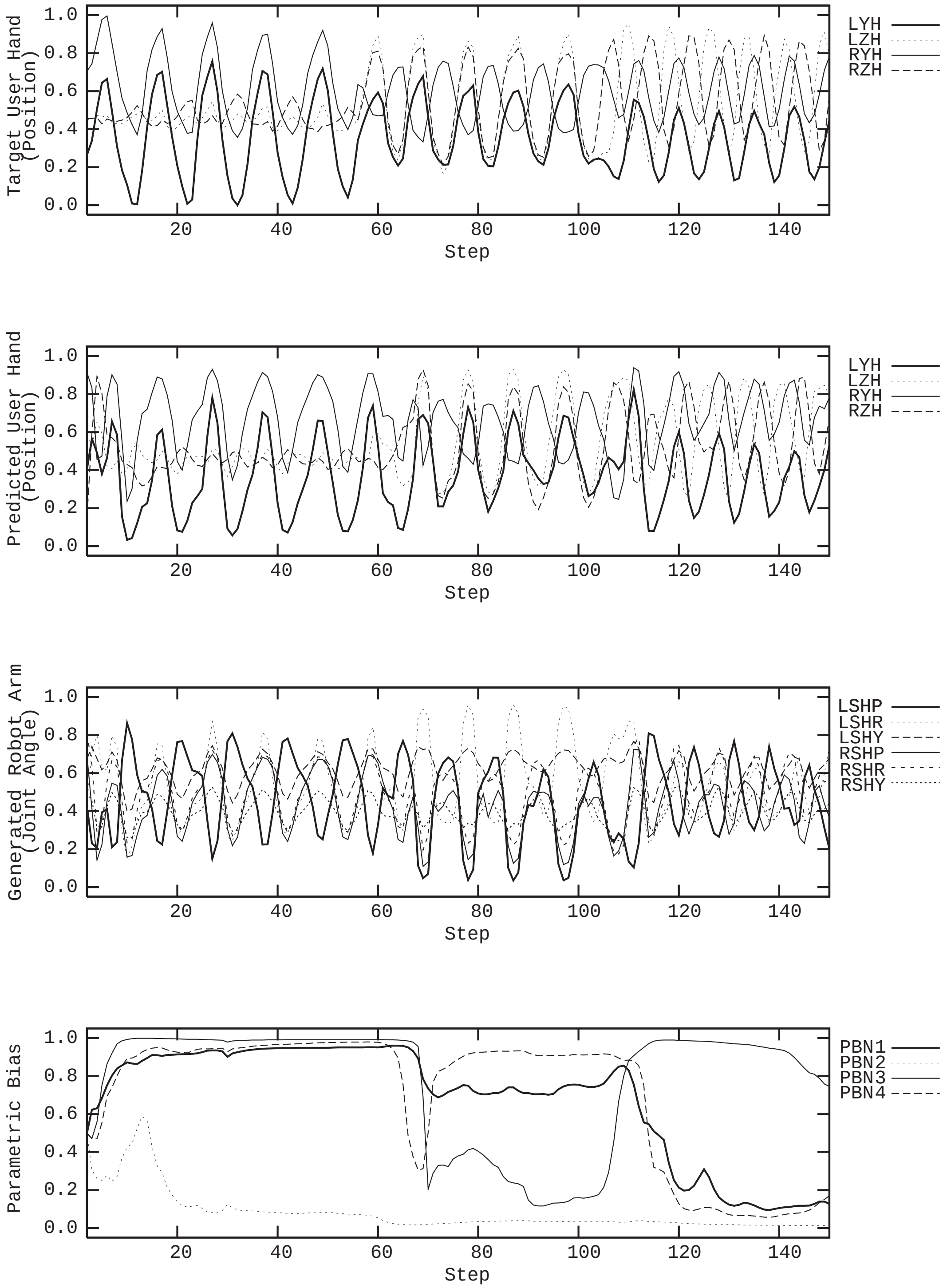


Figure 4. Switching of the robot movement pattern among three learned patterns as initiated by switching of user hand movement. User hand position and its prediction by the robot are shown in the first and the second row, respectively. The third row an the fourth row show motor outputs and PB vector, respectively.

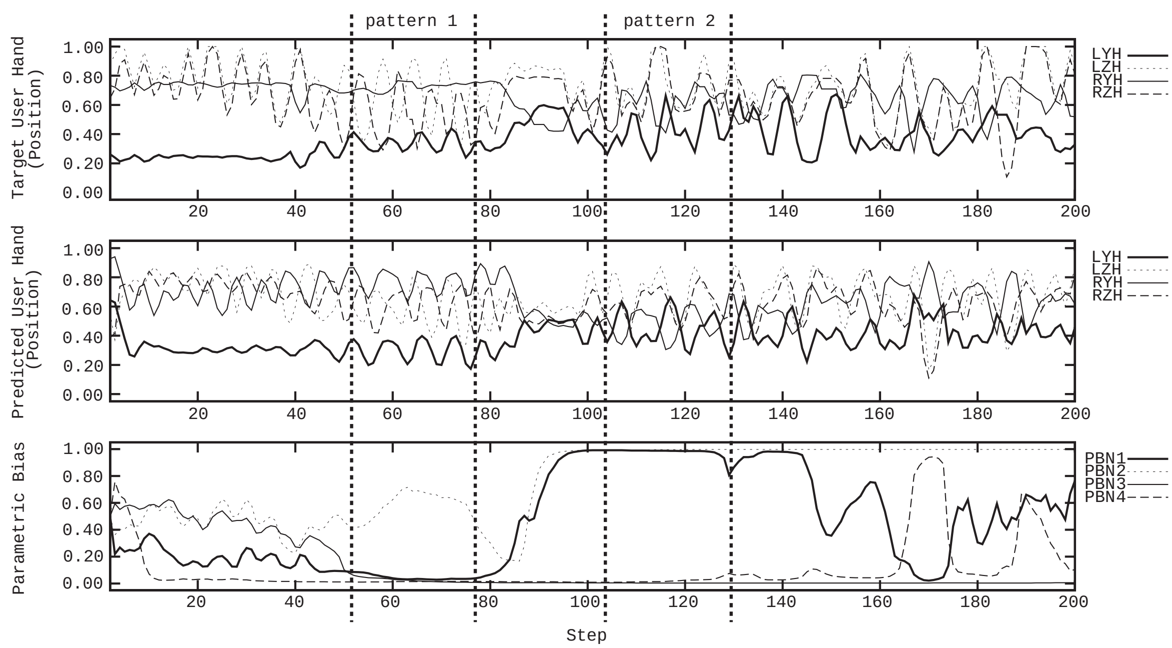


Figure 5. Joint attention as synchronization between the robot and the subject in imitation game. User hand position and its prediction by the robot are shown in the first and the second row, respectively. The third row shows PB vectors of the RNNPB.

Explaining Eye Movements During Learning as an Active Sampling Process

Jonathan D. Nelson^, Garrison W. Cottrell*, Javier R. Movellan+
^Cognitive Science Dept. 0515, UCSD, La Jolla, CA 92093-0515 , jnelson@cogsci.ucsd.edu
*Computer Science Dept. 0114, UCSD, La Jolla, CA 92093-0114, gary@cs.ucsd.edu
+Inst. for Neural Computation, 0523, UCSD, La Jolla, CA 92093-0523, movellan@mplab.ucsd.edu

Abstract

Savage [1] proposed analyzing active sampling problems as decision problems in which the goal is to maximize expected utility, relative to a probability distribution describing one's beliefs. In the past 20 years this framework has been applied to several psychological tasks [2]. We use this framework to model eye movements in a concept formation task [3], [4].

Introduction

In Shepard's classic concept learning task [3], participants gradually learn which of eight objects are consistent with a category unknown to them. The objects comprise each of 2^3 possible combinations of three binary stimulus dimensions. Objects include large black circle, small black triangle, small white circle, and so on. In each trial, the subject is shown a random object, guesses whether the object is consistent with the true concept, and receives feedback. This continues until the subject achieves near-perfect classification accuracy or a maximal number of trials is reached. Several theories make specific claims about how selective attention to different stimulus dimensions, for example shape or color, is deployed throughout learning. Rehder & Hoffman [4] devised a new version of Shepard's concept learning task to provide direct evidence of selective attention. Rehder & Hoffman separated the stimulus dimensions spatially, representing each binary dimension as a character that could take one of two values ($ or ¢, x or o, ? or !), at each vertex of a large triangle on a computer screen. Three primary findings were reported:

1. Early in learning, all stimulus dimensions are fixated.
2. There is gradual improvement in classification accuracy throughout learning.
3. After the concept is mastered, eye movements become efficient, restricted to only the dimensions needed to classify objects given the true concept.

Rehder & Hoffman suggested that RULEX [5], a prominent rule-based model of category learning, and ALCOVE [6], a prominent similarity-based model, each appeared to be contradicted by different features of their data, but did not specify a model to account for their data.

We show that a concise probabilistic model can account for the different amounts of learning required to master concepts in the classic task. Our generative Bayesian model gives higher probability to concepts that a priori criteria judge to be less complex, and that human subjects find easier to learn. Information obtained by fixating particular stimulus dimensions is assimilated in an optimal Bayesian manner. To calculate the usefulness of each possible eye movement we use a principled utility function, based on information theory [7], taking into account all learning to date. Results show that the eye movement model accounts for eye movement patterns observed both early and late in learning in the eye movement task. We further propose that this task exemplifies Helmholtz' idea [8] of vision as unconscious inference.

References

[1] Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.

[2] Nelson, J. D. (submitted). Finding useful tests: on Bayesian diagnosticity, probability, impact and information gain.

[3] Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied, 75(13)*, 1-42.

[4] Rehder, B., & Hoffman, A. B. (2003). Eyetracking and selective attention in category learning. In R. Alternman & D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Boston, MA: Cognitive Science Society.

[5] Nosofsky, R. M., Palmieri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101*, 53-79.

[6] Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review. 99*, 22-44.

[7] Cover, T. M., & Thomas, J. A. (1991). *Elements of Information Theory*. New York: Wiley.

[8] Helmholtz, H. L. F. von (1962). *Helmholtz's Treatise on Physiological Optics* (James P. C. Southall, Trans.). New York: Dover. (Original work, *Handbuch der physiologischen Optik*, published 1867)

Neural Correlates of Social Referencing

Leslie J. Carver
University of California, San Diego, ljcarver@ucsd.edu

Abstract

Social referencing refers to the ability to look at other people for information in novel situations. By the end of the first year of life, infants reliably look to adults when confronted with new situations, and regulate their behavior according to the adult's emotional response [1, 2]. The brain basis for this regulation of infants' emotional behavior is unclear. Because methods for measuring neural function in infants are highly constrained, there has been little research on the brain basis of social cognition. We report the results of a new method designed to evaluate the neural correlates of social referencing. We exposed infants to three novel, ambiguous stimuli and trained their caregivers to provide positive, negative, or neutral information about each of the stimuli. The association between emotion and individual stimuli was counterbalanced, such that no emotion was related to any one stimulus more frequently than any other emotion. Infants' behavior in this situation was rated by coders blind to the main hypothesis of the study. We then measured infants' brain activity in response to pictures of the stimuli used in behavioral testing. Infants show increased brain electrical activity in response to objects for which their caregivers provided negative emotional information (see Figure1).

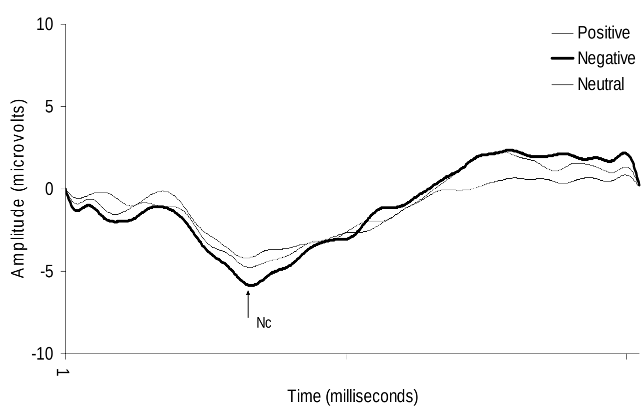


Figure 1. ERP response to picutres of objects associated with positive, negative, and neutral adult emotion.

This brain activity response is thought to reflect increased allocation of attention, and parallels behavioral results that suggest infants look more at toys associated with negative adult emotion. These results suggest that infants form associations between the emotional information and the stimulus with which it is associated. This method provides insight into the brain basis of infants' use of emotional information provided by a caregiver, and may prove useful in future studies of the brain basis and development of social cognition. We discuss the significance of these results for understanding the neural basis for the development of social referencing.
.

References

1. Gunnar, M.R. and C.S. Stone, *The effects of positive maternal affect on infant responses to pleasant, ambiguous, and fear-provoking toys.* Child Development, 1984. **55**: p. 1231-1236.
2. Mumme, D.L., A. Fernald, and C. Herrera, *Infants' responses to facial and vocal emotional signals in a social referencing paradigm.* Child Dev, 1996. **67**(6): p. 3219-37.

Learning to Manipulate Objects: A Quantitative Evaluation of Motionese

Katharina Rohlfing
Communication Science and Disorders
Northwestern University
k-rohlfing@northwestern.edu

Jannik Fritsch and Britta Wrede
Applied Computer Science
Technical Faculty, Bielefeld University
{jannik,bwrede}@techfak.uni-bielefeld.de

It is a well-known phenomenon that adults are speaking differently when addressing children. The purpose and the perceptual parameters of the child directed speech or Motherese have been discussed over the last three decades. Recently, this particular behavior has been characterized as directing the child's attention to relevant aspects of the speech signal [2]. What if - not limited to the speech signal - children are learning from input that is specifically designed for them? This is an idea that has barely been considered in the classical view on learning in pattern recognition but is gaining popularity in developmental psychology. Brand et al. [1] characterize a modification in mothers' infant-directed action as Motionese. In their studies, they observed that mothers' infant-directed actions reveal distinctive characteristics that amplify or exaggerate meaning and structure within their bodily motions. These characteristics were identified using 8 intuitive categories: range of motion, rate, repetitiveness, proximity to partner, enthusiasm, interactiveness, punctuation and simplification. Human coders gave each object demonstration a rating (0-4) for each category. However, the definition of the categories was tailored for human coders and often a single feature like, e.g., velocity was contained in several categories.

We present an algorithmic solution to identify visually observable modifications in actions in a quantitative manner requiring only the video stream. Our approach is based on a system for recognizing manipulative gestures [3]. The human hands are tracked based on skin color and the acting hand is related to objects in its vicinity based on their distance. Based on these processing steps, a variety of parameters like, e.g., hand velocity can be measured objectively. While our findings are in accordance with Brand's observations, we also provide insights into which specific aspects of the demonstration are relevant for motionese.

In our case study, we looked at an adult addressing either a child or another adult in his actions. Our approach allows us to identify the differences with respect to infant-directed action. In addition to the features extracted from visual data we performed an analysis of the sound signal to achieve a quantification of the multi-modal information directed towards the child. We hypothesize that cues in the acoustic signal can help to guide attention to relevant parts of actions. For example, the end of many actions like placing a cup on a table is signaled by the noise of the cup hitting the table top thus indicating the goal position of both the hand and the object. Analyses of the acoustic signal allow to highlight the action structure by (1) aligning noise from actions to the trajectories of gestures and (2) identifying especially emphasized regions in the speech signal.

Our findings suggest that techniques such as tracking manipulative gestures are an important step towards the analysis of qualitative and quantitative differences in the multi-modal input. They also indicate that multi-modal cues can help in understanding (and learning) of actions and their meaning. This will help research in robotics to develop ways of how to reduce the complexity of the acquisition problem, i.e. to reduce the requirements on built-in architectural constraints of learning mechanisms [4]. We will also discuss the potential application possibilities (such as plug-ins to transcription programs) to make these techniques available to developmental psychologists.

[1] This work has been supported by the German Research Foundation within the Emmy Noether-Program, the Collaborative Research Center 'Situated Artificial Communicators' as well as the Graduate Program 'Task Oriented Communication'.

References

[1] R. J. Brand, D. A. Baldwin, and L. A. Ashburn. Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science*, 5(1):72–83, 2002.

[2] P. Dominey and C. Dodane. Indeterminacy in language acquisition: the role of child directed speech and joint attention. *Journal of Neurolinguistics*, 17:121–145, 2004.

[3] J. Fritsch. *Vision-based Recognition of Gestures with Context*. Dissertation, Bielefeld University, Technical Faculty, 2003.

[4] G. Gergely. What should a robot learn from an infant? Mechanisms of action interpretation and observational learning in infancy. In *Proc. Third Int. Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, pages 13–24, Boston, MA, USA, 2003.

Can Robotic Brains be Social? Scientists Caught Back-peddling

Colin T. SCHMIDT
Sorbonne University
&
LIUM Computer Science Laboratory
Le Mans University

52 rue des Docteurs Calmette et Guérin,
BP 2045
53020 Laval Cedex 09
Colin.Schmidt@univ-lemans.fr

Abstract

The year 2001 was a "fast year" for research in Natural Computation and Robotics. In that year, the author of an article in Minds & Machines *asks two highly pertinent questions for robotics: 1) If a robot is able to participate in simple language games as adequately as a child, should we concede true meaning and intelligence to it? and 2) How would we go about developing a robot which could possibly live up to a positive answer to the first question? My approach is straightforward: a) refute the first question, so as to b) forget the last. I then argue in favour of supporting another well-known sub-domain of AI/HCI/Robotics thought in order to stimulate research in the artificial sciences.*

I. PREAMBLE ON ROBOTIC BRAINS

I herein address an issue that has a 50-year and more history in the Sciences of the Artificial. Important research being carried out at top-notch scientific institutions like MIT, Carnegie Mellon University and still yet many others seem to be having difficulty with the mind-body problem in creating robots that think. Weng, McClelland, Pentland, Sporns, Stockman, Sur & Thelen teamed up to confirm this in their *Science* article a few years back (2001) with discussion on "automous mental development" that was limited to brain and body building[1]. Whether their intention included outright occultation of the mind or not, *reductionism* cannot account for mind as it cuts this latter off from its socio-communicative dimension (i.e. relations with other minds), the very features that *make a mind a mind and not a brain* to state things in a 'folkish' manner. A few months later in that same year, Brian Scassellati (at MIT AI Lab. at the time) used the following citation from Turing's famous article presumably in order to sum up his Doctoral Dissertation (first citation, placed top centre-page, Chapter 1).

> Instead of trying to produce a program to simulate the adult mind, why not rather try to produce one which simulates the child's? (A. Turing 1950, p. 456)[2].

I do not have the impression that exponential progress in the area or "humanoid robotics" has since overcome this philosophical hurdle to capture the *dialogical* essence of mind.

With his "embodied theory of mind" Scassellati may have been referring to —or taking inspiration from— Jordan Zlatev's 1997 work on *Situated Embodiment.*

Whatever the relation, academics working in Robotics and related fields like Human-Machine Interaction and Artificial Intelligence often seem to undergo an out-of-proportion positivistic enthusiasm for their 'babies'. Why is this? Do not any of them have the liberty to *really* express their doubts? There surely must be some conceptual hesitation in their mind when the action implied by their work constitutes replacing human beings. It is a good thing that when they do replace a human being with a machine, it is quite often in the context of repetitive task handling that the human being no longer likes to do. But there are a few academics that work on technological challenges that remain purely technological in nature (i.e. not that useful since man does not want to give up the action concerned —examples involving speaking come to mind). Their technological audacity does not stem from usability reports or interviews with users. Simply defying physical laws is what they seek to do.

[1] WENG J., MCCLELLAND J., PENTLAND A., SPORNS O., STOCKMAN I., SUR M. & THELEN E. (2001), "Automous Mental Development by Robots and Animals", *Science Magazine,* Vol. 291 N° 5504, The American Association for the Advancement of Science (AAAS), pp. 599-600.

[2] *Cf.* SCASSELLATI B. (2001). *Foundations for a Theory of Mind for a Humanoid Robot*, Ph.D. Dissertation, May 6: Massachussetts Institute of Technology.

Scassellati gets his expectations about *machine intentionality* the wrong way around when he writes about the "Implications to Social Robotics" of his work: "Rather than requiring users to learn some esoteric and exact programming language or interface, more and more systems are beginning to use the natural social interfaces that people use with each other. People continuously use this extremely rich and complex communication mechanism with seemingly little effort. The desire to have technologies that are responsive to these same social cues will continue to drive the development of systems [...] Theory of mind skills will be central to any technology that interacts with people. People attribute beliefs, goals, and desires to other agents so readily and naturally that it is extremely difficult for them to interact without using these skills. They will expect technology to do the same".[3]

But interlocutors in human-resembling communication like to be reassured that their interlocutor is human. In fact, if one wishes to escape from the Electrical Engineering and Computer Science point of view, one has to read for example the works of D. Norman, a cognitivist who addressed the DARPA/NSF Conference on Human-Robot Interaction in... yes, the year 2001. He then gave an analogy to persuade any human being to understand why *machine speech should not be flawless* in the human sense[4]. And he is not the only one that argues in this direction (*cf. infra*).

Brain-child projects are fine, but may they ever lead to "mind-childs"? Perhaps. Let us now turn to the further specialised field of Natural Computation. At least one influential author has caught my eye.

II. "ARTIFICIAL PROBLEMS"

Some authors like to delve into "thought experiments" using such examples to study the possibilities of resolving some of the problems of the Artificial Sciences. Let us try to understand in simple terms what J. Zlatev meant in his (yes!) 2001 article in *Minds and Machines*[5]. His goal was to use one of these "thought experiments" in order to up-grade the position of the Artificial —robots— on the social status scale, or perhaps quite possibly, to argument in favour taking robotic technology even further ahead. Or was it only to test the plausibility of lifting them up to our level?

In any event, he devises a fictive situation for this purpose. A two-year old child is sitting on the floor and interacting with his father through eye contact as they pass things likes balls and blocks back and forth. The child gestures towards an object that is out of reach and says "*train*". Dad says "Oh, you want the *train-engine*". In receiving it, the child repeats "*train-engine*", thereby indicating that the adult's slight correction concerning the proper of term of reference has not passed unnoticed; etc. etc (*cf.* p. 155). Zlatev then tells us that, when it comes to playing simple language games like this, you can remove the two-year old and put a robot in the very same spot on the floor to occupy Dad; he says that today we can build a robot that would have the same physical and intellectual capacities as this person's son or daughter. I agree with him so far.

My endeavour is to focus on the communication part of his proposal as I believe this is where genetics-based robotics would stand to gain the most from my critique.

As communication is a social activity that does not have anything really to do with genes themselves, I am sure the author pointed out here will have no objection: he does in fact carry his point of view well outside of the materialistic topics spoken about traditionally in robotics.

One does not really have to read beyond the Introduction of Zlatev's rather lengthy article (though it does read quite nicely) to find out whether his version of "Epigenetic Robotics" will not be able to defy the tormenting philosophical questions that Strong AI has been battling with since the days of R. Schank in the 70s, Herbert Simon and A. Newell at Rand Corporation and CMU in the early and mid 50s, A. de Groot (1946) and still yet others, questions namely such as "is it possible for Man to build a machine to think?". As Zlatev (p. 157), I do not have the philosophical wherewithal, but I esteem myself to be in a position to be able to bring a certain number of issues to his attention, though perhaps only really in point form. A glance over my notes would nevertheless seem to indicate that what I have to say could be somewhat important for specialists in Natural Computation (*cf. infra*, Conclusive Remarks).

I understand epigenetics to be a field of study that involves mainly the "physicalist options" of the Cognitive Sciences; the work of Zlatev and Dennett are encouraging as they do endeavour to look into the other options possible under this banner, even if the latter author mentioned here has confused the notions of *mind* and *brain* in the past (*cf.* D. Dennett 1996).

III. MY APPROACH

My vision of the way things are for the sciences of the Artificial in general, and thus Natural Computation and

---

[3] *Cf. ibidem* p. 159.

[4] After exposing a version of the Asimovian laws of robotics, he states the following: "while speech input is still imperfect, the robot must make this clear [...]." He them gives the maxims the first of which is: "Don't have flawless, complex speech output at a level far more sophisticated than can be understood. If the robot wants people to realise it has imperfect understanding of language, it should exhibit these imperfections in the way it speaks. (If a foreign speaking person could speak fluent English but only understand pidgin speech, the more it spoke flawlessly, the less other people would understand the need to speak in pidgin)". *Cf.* NORMAN D. (2001), "How Might Humans Interact with Robots? Human-Robot Interaction and the Laws of Robotology", keynote address, *The DARPA/NSF Conference on Human-Robot Interaction*, San Luis Obispo CA, September.

[5] *Cf.* ZLATEV J. (2001), "The Epigenesis of Meaning in Human Beings, and Possibly in Robots", *Minds and Machines* n° 11, Kluwer.

Robotics in particular, will quite simply be based on the two questions brought forth by the author:

If a robot is able to participate in simple language games as adequately as a child, should we concede true meaning and intelligence to it?

How would we go about developing a robot which could possibly live up to a positive answer to the first question?

My approach is straightforward: ***a/***refute the first, ***b/***forget the last. In order to not leave specialists in robotics following the example targeted here in the dark, I will ***c/***deploy a prospective epistemology which will introduce discussion leading to the reinforcement of another well-known sub-domain of AI/HCI/Robotics thought (*cf.* the last section): Weak machine intelligence.

IV. Zlatev's How-to's

If I understand correctly, what the author means by "reverse engineering" is that in recreating the behaviour of communicative intelligence, while working with the smaller units of behaviour to form the larger ones of the robot language acquisition process, the robot builder must situate his action within a long set of implications enunciated in the exact opposite order: "linguistic meaning presupposes shared conventions, as a form of mutual knowledge. Conventions presuppose reflexive consciousness, allowing them to be learned and followed. Self-consciousness presupposes the perception of oneself as an intentional agent. Perception of oneself as an intentional being presupposes the perception of others similarly. Hence, other-intentionality, self-intentionality, self-consciousness and language form a possibly necessary developmental progression and an artificial system aiming at real —as opposed to simulated— language use …" (p. 189). This does appear to give a more pragmatic aspect to the 'usual implementation technique' in the artificial technologies fields, but is there not something very paradoxical here?

If these are presuppositions proper, they would indicate rather that one should start by building a robot by taking the larger chunks on and then the smaller ones. In his initial *explanation* of the thought experiment involving a child playing with toys and talking with Dad (p. 155-156), Zlatev starts off with intentions, goes through meaning and understanding to get to the grammar part. In fact this type of discourse is typical of positivistic science that has, so to speak, bit off too much to chew and then wonders what to do.

At this point in the game one may ask if robotics really does have a set methodology and direction to follow… or is it just heuristically (or hysterically) shooting in the dark? I could even say that the (almost not) implicit goal being chased after here, recreating man in behaviour as well as social role, is so difficult that, however big the steps robotic technology is taking towards this goal (excuse the pun!), we do have a very long way to go. As I see it, a more plausible way of seeing things would be to take the larger components for the smaller ones: simple intentionality features of members of the human race put end-to-end to build very complex grammatical constructions. *Why should the linguistic utterances of language users be considered any less complex then human intentionality?* It would seem obvious that —after saying over the last 25 years that utterances lacking their intentions-driven component cannot have meaning— we could and should be able to imagine positive responses to this epistemological question. Zlatev seems to be going in the right direction but shows here the sentiment that the Robotics community will have trouble following his initiative; hence the need for an adjective on the banner, (i.e. "epigenetics").

V. Dialogical Communication

In Section 4.2, it becomes clear that Zlatev's approach is based on a rather dated account of interpersonal communication. Although Grice inaugurated the discursive study of ordinary language use with his studies on "implicatures" —a welcomed advance from the area between Philosophy and 'plain linguistics'—, his (and Gazar's) results are not sufficient for what we are expecting of robotics intelligence today. Whatever we may expect, it is entirely clear that Zlatev's model of intersubjectivity is not able to escape that of Grice's presentation pattern of intentional layering: A knows X, B knows that A knows X, A knows that B knows that A knows X, etc. (cf. p. 182). Communication theory has come a long way since then (cf. Vernant, Vanderveken, Jacques, Shotter and my own work in the 90's). It has come to fully accept asymmetry as the basic nature of the communicative link. The layering of intentions performed by Grice would suggest that a symmetric alibi was still necessary.

The progress that has been made in communication theory could quite simply be stated as follows, though I run the risk of being accused of over-simplification:

> A cannot do with B what B does with A, whatever the communicative activity is (i.e. discussing explicitly or implicitly about knowing X).

So if the authors means to speak about a Robot that participates in simple language games, how can his analysis of the situation be water tight if the pragmatic nature of the relation in question here is not solid? *Who is speaking to whom?* is a simple question that resumes what I mean by the pragmatic nature of the relation and this is important as the father in the thought experiment here, as Zlatev pointed out, does not know who or what he is interacting with. Is that really his daughter on the floor in front of him? Is that not his daughter?

One of the most relevant things I have to say in this article is that it would appear that the field of robotics is too materialist to succeed in tackling the myth of humanity. Its endeavours only represent reproducing/replacing the mere *manifestation* of communicative intelligence; the dialogical profundity of human cognition and communication skills are hard to replace over and above the toddler level. This is why the AI project has been reduced. What level of dialogism can robotics-embedded AI produce? At the outset, Turing and post-Turing discussion was about adult dialogical capacities, Zlatev tackles two-year olds (behaviour only?), and in relation to this, Scassellati drastically gears the argumentation down once again in his work: "The systems presented here will not begin to approach some of the complex social skills that children master even in the first year of life".[6] In fact, there exists a logical impossibility for a robot to participate in dialogical activities because of the *primium relationalis* in human communication as defined by F. Jacques as early as in the beginning of the 80s. This means that for any propositional content flow to obtain success, it is dependent on the relationship between interlocutors *that must exist prior to it*. AI *tries* things the other way around if it even considers the pre-imminence of the relation at all.

VI. SOCIAL STATUS

I would have to add a few itilicised characters to Zlatev's first question:

If a robot is able to participate in simple language games as adequately as a child *at least in appearance*, should we concede true meaning and intelligence to it?

P. Bourdieu[7] would say that mechanical 'objects' like *Robot Sapiens* are only simple artefacts, whatever that species may be capable of: it is so-to-speak Made in the Republic of Human Society and is thus subjected to the rules therein, rules that go beyond the boundaries of mechanics, genes, synthetic flesh and other physical paraphernalia.

But what Zlatev does well in his article is point out that it is important for machines, if they are to have success in performing operations in a human way, to learn over a period of time, to have a history. They need to have the opportunity to acquire the skills to evolve their on knowledge. The programming-in method is, I think we are safe to say now, out since people working in robotic have started to take into account, as Zlatev, the more philosophical discourse on their subject (Dennett, Dreyfus, Turkle, Turing, etc.).

VII. ROBOTIC "INTELLIGENCE" T*WEAK*ED DOWN FOR PARENTS AND OTHER ADULTS

One of our main points today is that, other than for all the logico-philosophical and relational points exposed above, society is far from being in a position to accept the *advanced* products to come of "robotology", even if robotics, epigenetic or otherwise, is making good progress now. These products *are* for our utilitarian society but, in order to be fully accepted by the Self —as is the Other in a dialogical setting (context which remains exclusively inter-human for the time being)—, without the proper *identity features* they will remain at the fringe of human communities. Zlatev finds it necessary to play with our emotions to get his point across and so speaks of the remembrance of persons dear to oneself while they are in a deceased state (*cf.* the second thought experiment at p. 160-161); I do not for the moment find it necessary to go this far, but it is a rather good idea to use strong emotions —they enhance argumentation. Think about a young boy, say in the 5-10 year-old range, who comes into the living room to alert his parents of some happening and, in the middle of their discussion, our eavesdropping reveals the following utterance:

"The robot is bothering my two-year old sister".

Would the adults react in the *usual manner*? That is to say in the same way as when another human sibling bothers the two-year old? Would the parent regularly "commissioned" to handle such a scenario go into the recreation room with the intent to, say, scold the robot?

The robot will not possess the proper *identity features* in our society for some time yet to receive the treatment that might habitually correspond to bothering, teasing, pushing, hitting and so forth. For example, I doubt that, even in ten years time, *sincerely* scolding (and I mean with authentic sincerity) robots would come into practice —oh, and taking one over one's knee, even less so.

By "proper identity features" for a robot to function in a normal way at a societal level, I quite simply refer to *social status, family-induced selfhood* and *moral existence*, features *perceived as so by humans*. Furthermore, purely logical reasons for artefacts like robots not being equipped for total integration into human society do exist in the literature in Human Sciences, such as those pertaining to the *pragmatic* aspects of communication (*cf. supra*, section V): I demonstrate this elsewhere in Italy at another time[8].

I will have to maintain the question I asked of the robotics community in 2002 "Can simulating Man's physical abilities meet up to the expectations we have of Robot Technology?" (*cf.* my Berlin Ro-Man paper). The fact that T. Watanabe (*et al*) "abandoned" his InterRobot (iRT) technology for a "lesser embodied" form of communicative interaction with humans —iRT's on-screen version called InterActor— is indicative of the difficulties of human speakers to interact with very similar-looking creatures[9], which I think is the intention of the Natural Computation

---

[6] SCASSELLATI B. (2001). *Foundations for a Theory of Mind for a Humanoid Robot*, Ph.D. Dissertation, May 6: Massachussetts Institute of Technology, p. 19.

[7] Cf. BOURDIEU, P., (1982). *Ce que parler veut dire : L'économie des échanges linguistiques*, Fayard.

[8] *Cf.* SCHMIDT C.T. (Forthcoming 2004), "A Relational Stance in the Philosophy of Artificial Intelligence", *European Conference on Computing and Philosophy,* E-CAP 2004, 3-5 June, University of Pavia, Italy: Kluwer.

[9] *Cf.* WATANABE T. (2002), "InterActor: Speech-Driven Embodied Interactive Actor", *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication* RO-MAN 2002, Sept. 25-27 2002, Berlin Germany: IEEE, p. 430-435.

Community. Of course, writers as influential as H. Dreyfus (1972) have strongly suggested that the lack of corporal extension was the hindrance computer programmers met up against in the project of simulating human intelligence, but it has been proven (both experimentally and argumentatively) that fully simulating natural language (D. Luzzati 1989) and full simulation of human features (C. Schmidt 2001, 2002) goes against all sensible logos to improve interaction (though at more advanced levels, i.e. adult interaction). This goes to confirm the well-fitting of D. Norman's 'law-like' advice for designers about the flawlessness of machines.

All in all, I argue in this article for the use of *Weak AI, Reduced Robotics* and *Invisible Interfaces* (a Cog/InterRobot/Kismit type of creature does get one's attention). This *is* necessary for producing useful robotics for adults too. Taking research in this area back up to the adult level is the main idea, is it not?

I would like to sum up with the words of a very active researcher in the Robotics field:

"[...] humanising technology does not necessarily require creation of humanoid technology, it could rather push forward to develop technology which meets the specifically human ways and strategies of (socially) living and surviving"[10].

VIII. REFERENCES

[1] ARCHER N., HEAD M., WOLLERSHEIM J. & YUAN Y. (1996), “Investigation of Voice and Text Output Modes with Abstraction in a Computer Interface”, *Interacting with Computers,* vol. 8 n° 4, Amsterdam: Elsevier Science, p. 323-345.

[2] BROOKS R. (2002), *Robot: The Future of Flesh and Machines*, London: The Penguin Press.

[3] BROOKS R. (1991), “Intelligent Room without Representation”, *Artificial Intelligence*, n° 47, 139-159.

[4] BROOKS R. & STEIN L. (1994), “Building Brains for Bodies”, *Autonomous Robots*, n° 1, p. 7-25.

[5] BUTTON G., COULTER J. LEE J. & SHARROCK W. (1995), *Computers, Minds and Conduct,* Oxford: Polity-Blackwell.

[6] CARD S., MORAN T. & NEWELL A. (1983), *The Psychology of Human-Computer Interaction*, Hillsdale NJ: Lawrence Erlbaum Associates.

[7] CARROLL J., MORAN T. (1991), Special Issue of *Human-Computer Interaction* on ‘Design Rationale’, vol. 6 n° 3 & 4, Hillsdale NJ: Lawrence Erlbaum Associates.

[8] DAHLBOM B. (1993), *Dennett and his Critics. Demystifying Mind,* Oxford: Basil Blackwell.

[9] DASCAL M. (1992), “Why does Language Matter to Artificial Intelligence? “, *Mind and Machines*, 2, Dordrecht, The Netherlands: Kluwer Academic Publishers, p. 145-174.

[10] DAUTENHAHN K. (1997), "The Role of Interactive Conceptions of Intelligence and Life in Cognitive Technology", *Proceedings of the Second International Conference on Cognitive Technology*, Aug. 25-28, Aizu Japan, IEEE p. 33-43.

[11] DENNETT D. (1996), [*Kinds of Minds*] tr. in French*; La diversité des esprits : Une approche de la conscience*, Paris : Hachette.

[12] DENNETT D. (1987), *The Intentional Stance,* Cambridge MA: Bradford/The MIT Press.

[13] DENNETT D. (1971), “Intentional Systems”, *Journal of Philosophy*, n° 8, p. 87-105.

[14] DREYFUS H.L. (1972), *What Computers* Still *Can’t Do. A Critique of Artificial Reason,* Cambridge MA: The MIT Press.

[15] FOGG J. & NASS C. (1997), “Silicon Sycophants: The Effects of Computers that Flatter”, *International Journal of Human-Computer Studies*, May issue.

[16] GAZDAR J. (1979*), Pragmatics, Implicature, Presupposition and Logical Form*, New York: Academic Press.

[17] GRICE P. (1975) “Logic and Conversation”, COLE P. & MORGAN J. *Syntax and Semantics: Speech Acts*, vol n° 3, New York: Academic Press.

[18] HOFSTADTER D. (1979) *Godel, Escher, Bach: an Eternal Golden Braid,* New York: Basic Books.

[19] JACQUES F. (1990), « De “On Denoting” de B. Russell à “On Referring” de P.F. Strawson, l’avenir d’un paradigme », *Hermès VII : Bertrand Russell, de la logique à la politique*, Paris : Editions du CNRS, p. 91-100.

[20] JACQUES F. (1979), *Dialogiques, Recherches logiques sur le dialogue*, Paris: PUF.

[21] JACQUES F. (1976), “A Terminological Note about Reference”, RYLE G., *Contemporary Aspects of Philosophy*, London: Oriel Press.

[22] MEY J. & TAMURA H. (1994), “Barriers to Communication in a Computer Age”, *AI & Society*, n° 6 Springer, p. 62-77.

[23] NORMAN D. (2001), "How Might Humans Interact with Robots? Human-Robot Interaction and the Laws of Robotology", keynote address, *The DARPA/NSF Conference on Human-Robot Interaction*, San Luis Obispo CA, September. www.jnd.org/dn.mss/Humans_and_Robots.html

[24] PEIRIS D., GREGOR P & ALM N. (in print), “The Effects of Simulating Human Conversational Style in a Computer-based Interview”, *Interacting with Computers*, Amsterdam: Elsevier Science.

[25] PUTNAM H. (1981), “Brains in a Vat” and “A Problem with Reference”. *Reason, Truth and History,* Cambridge MA: Cambridge University Press.

[26] PUTNAM H. (1975), *Mind, Language and Reality*, vol. 2, Cambridge MA: Cambridge University Press.

[27] PUTNAM H. (1960), “Minds and Machines”, Hook S., *Dimensions of Mind*, London: Collier-Macmillan.

[28] QUINE W.V.O. (1960) *Word & Object,* Cambridge MA: The MIT Press.

[29] SCASSELLATI B. (2001). *Foundations for a Theory of Mind for a Humanoid Robot*, Ph.D. Dissertation, May 6: Massachussetts Institute of Technology.

[30] SCHMIDT C.T. (Forthcoming 2004), "A Relational Stance in the Philosophy of Artificial Intelligence", *European Conference on Computing and Philosophy,* E-CAP 2004, 3-5 June, University of Pavia, Italy: Kluwer.

[10] *Cf.* DAUTENHAHN K. (1997), "The Role of Interactive Conceptions of Intelligence and Life in Cognitive Technology", *Proceedings of the Second International Conference on Cognitive Technology*, Aug. 25-28, Aizu Japan, IEEE p. 33-43.

[31] SCHMIDT C.T. (2002), "Socially Interactive Robots. Why Our Current Beliefs about Them *Still* Work", *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication* RO-MAN 2002, Sept. 25-27 2002, Berlin Germany: IEEE, p. 560-564.

[32] SCHMIDT C. (2001*),* [*Mind and Machine, Concepts for a Design Community at Work* (in French)]: *L'esprit et la machine, une communauté conceptrice à l'œuvre. Distinctions pour un discours supérieur*, Lille: Presses universitaires de Septentrion.

[33] SCHMIDT C. (1997), "Pragmatically Pritine, the Dialogical Cause", 'Open Peer Community Invited Commentary', on MELE A., "Real Self-deception", *Behavioral and Brain Sciences*, 20:1, Cambridge MA: Cambridge University Press.

[34] SCHMIDT C. (1997), "The Systemics of Dialogism: On the Prevalence of the Self in HCI Design", *Journal of the American Society for Information Science,* vol. 48 n° 11 November, New York: John Wiley & Sons.

[35] SCHMIDT C. (1995), "Information Processing, Context Creation, Setting Minds in Public Arenas: Investigative Techniques for Client/Automaton Dialogue", IJCAI-'95 (Montréal), *Proceedings of the 14th International Joint Conference on Artificial Intelligence,* Paris: Institut Blaise Pascal, p. 121-131.

[36] SHANNON C. & WEAVER W. (1949), *The Mathematical Theory of Communication*, Urbana: University of Illinois Press.

[37] SEARLE J. (1975), *Speech Acts*, Cambridge MA: Cambridge University Press.

[38] SEARLE J. (1969), "Indirect Speech Acts", COLE P. & MORGAN J. *Syntax and Semantics: Speech Acts*, vol n° 3, New York: Academic Press.

[39] SHOTTER J., (1997), "AI and the Dialogical", *American Behavioral Scientist,* vol 40, n° 6, Special Issue edited by R. HARRÉ on "Computation and the Mind", p. 813-828.

[40] SPERBER D. & WILSON D J. (1986), *Relevance: Communication and Communication*, Oxford: Blackwell.

[41] STRAWSON P. (1950), "On Referring", *Mind*, vol. LIX July n°235.

[42] SUCHMAN L. (1987), *Plans and Situated Actions: The Problem of Human-Machine Communication*, Cambridge MA: Cambridge University Press.

[43] TURING A. (1950), "Computing Machinery and Intelligence", *Mind*, vol. LIX n° 236.

[44] TURKLE S. (1984), *The Second Self, Computers and the Human Spirit*, New York: Simon & Schuster.

[45] VERNANT D. (1996), [Machine Intelligence and the Dialogical Capacity of the Machine] in French; « Intelligence de la machine et sa capacité dialogique », RIALLE V. & FISETTE D., Penser l'esprit : des sciences de la cognition à une philosophie cognitive, Grenoble : Presses universitaires de Grenoble.

[46] WATANABE T. (2002), "InterActor: Speech-Driven Embodied Interactive Actor", *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication* RO-MAN 2002, Sept. 25-27 2002, Berlin Germany: IEEE, p. 430-435.

[47] WENG J., MCCLELLAND J., PENTLAND A., SPORNS O., STOCKMAN I., SUR M. & THELEN E. (2001), "Automous Mental Development by Robots and Animals", *Science Magazine,* Vol. 291 N° 5504, The American Association for the Advancement of Science (AAAS), pp. 599-600.

[48] WITTGENSTEIN L. (1953). *Philosophical Investigations*, Oxford: Blackwell.

[49] ZLATEV J. (2001), "The Epigenesis of Meaning in Human Beings, and Possibly in Robots", *Minds and Machines* n° 11, Kluwer, p. 155-195.

RUBI: A Robotic Platform for Real-time Social Interaction

Bret Fortenberry, Joel Chenu, Dan Eaton, Javier R. Movellan

Institute for Neural Computation
University of California San Diego
{bret, joel, deaton, movellan}@mplab.ucsd.edu

Abstract

The majority of our waking hours are spent engaging in social interactions. Some of these interactions occur at the level of long-term strategic planning while others take place at faster time scales, such as in conversations or card games. The ability to perceive subtle gestural, postural, and facial cues, in addition to verbal language, in real-time is a critical component of social interaction. An understanding of the underlying perceptual primitives that support this kind of real-time social cognition is key to understanding social development.

This paper presents a humanoid robot designed for research on real-time social interaction between robots and humans. We discuss many aspects of the current system including motor control, face tracking, and speech recognition and how it is utilized for human to robot social interaction. We also describe plans to increase the system's capabilities and communication skills. Last, we describe planned research for the study of real-time social interactions.

1. Introduction

Robots present an ideal opportunity to study the development of social interaction in infants and children [2]. It is possible to create robots that exhibit precisely controlled contingency structures. By observing how infants interact with these robots we gain an opportunity to understand how infants identify the operating characteristics of the social agents with whom they interact. In this paper we present progress on the development of a social interaction robot, 'RUBI', designed to communicate and interact with children and to serve as a platform on experiments for social interaction and social development with children.

Figure 1. The current appearance of RUBI was chosen by Kai Movellan, the 2 year old child on the picture. Kai refused to interact with the early versions of RUBI and proved to be a wonderful critic for the design team.

2 System Architecture

2.1 Robot Structure

RUBI is a three foot tall, pleasantly plump robot with a head and two arms (See Figure 1). It stands on four non-motorized rubber wheels for moving it easily from place to place. RUBI is self-contained; all of its components (computer, microphone, speakers, etc.) are inside its body structure. The external connections consist of a single power cable and a wireless Ethernet card. The computer (specified below) is on a sliding vibration-reducing rack. The head consists of two cameras 13.5 inches apart representing the eyes, and a disk-shaped face with a small button nose and permanent smile. An omni-directional camera is on a rod that extends above the head for a clear view of the world. The body is a wooden night-stand that is spacious enough to hold all of the RUBI's components, but short enough to keep RUBI non-threatening to children.

Pilot experiments with 2-4 year old children helped shape RUBI's physical design. Children were frightened of the original design, but several iterations of experiments helped make it more child-friendly. Some of the most effective changes included adding clothing to cover mechanical parts, giving RUBI a "smile", and making it shorter.

2.2 Motor System

RUBI's head is based on the early Robovie design from Hiroshi Ishiguro's group [5, 4]. It has 9 degrees of freedom: Three degrees of freedom (pan, tilt and roll) implemented by stepper motors that are driven by a Galil DMC-1832 PCI motor controller. The neck can move 54 degrees up and 30 degrees down from center position and 54 left and right of center The maximum pulse rate of the motor drivers is 144,000 degrees/sec, but the motors for RUBI's neck slip at high rates (over 480 degrees/sec). However, we have manually set RUBI's maximum speed to 60 degrees per second; faster motor control is possible but has an unnatural appearance. The remaining six degrees of freedom are in the eyes, both of which have pan, tilt and zoom motors. The eye cameras are SONY EVI-G20 PTZ (Pan-Tilt-Zoom) cameras. Their horizontal range is ± 30 degrees and their vertical range is ± 15 degrees. Maximum speed on both horizontal and vertical axes is 150 degrees/sec. They are controlled via a 9600 bit/sec VISCA-protocol serial connection.

2.3 Vision System

RUBI's vision sensors consist of two SONY EVI-G20 color cameras which are the "eyes". The third input is a low-resolution stationary omni-directional camera acting as RUBI's peripheral vision. All three use a component out and are routed through a quad video splitter that combines the images into one 640x480 image. The single image is then captured via a BT848 video capture card at 30 Hz.

RUBI's two eyes handle the main face tracking tasks. The commands given to the motor controller are determined by a convolutional HMM architecture that combines both a color-based and a Frontal Face-Detector to produce a distribution of possible locations and scales for faces on the image plane [6, 3].

2.3.1 Frontal Face-Detector

The face-detector system is explained in [3]. It can detect over 98% of faces with minimal false-alarms in difficult background conditions running in real time at 30 frames per second. Its main limitation is that it only detects upright frontal faces. Because of this, the frontal face-detector is complemented by a color-based detector that handles non-frontal views. Source code for the face detector is available at *http://kolmogorov.sourceforge.net* as part of the Kolmogorov project.

2.3.2 Color Tracker

The color-based system utilizes a convolutional HMM architecture desribed in [6]. The system uses standard HMM equations to update a probability distribution of 100000 states representing the possible location and scale of faces on the image plane. The color model uses two 1000 bin hue histograms, one for faces and the other for backgrounds. The initial face histogram model is defined a follows

$$p(c) \propto \exp(-\frac{d^2(h, 17.95)}{12.2^2}), \tag{1}$$

where $\mu = 17.95$, $\sigma = 12.2$ and d is the angular distance in degrees, between h and μ. The initial background histogram model is uniform. Each time the face detector finds a face, the color models for faces and background are updated with a weighted average of the current and past histogram. The most probable location and scale is chosen for display and for use by the head controller.

2.3.3 Peripheral Vision

RUBI's peripheral vision is handled by an omni-directional camera. Since the motors have limited range of motion the omni-directional camera vision software uses only a ± 55 degree field of view. The omni-directional camera uses a motion detection and a non-adaptive color model to search for people. The color model is based on statistics of hue of many example faces.

2.3.4 Head Control: Explore Mode

The goal of RUBI's head controller is to maximize the expected number of face detections. To do so it switches back and forth between two operation modes: (1) Explore mode, and (2) Face tracking mode.

While in explore mode RUBI orients towards motion detected in its peripheral vision (with the omni-directional camera). RUBI's head movement is controlled by a stochastic difference equation that favors a combination of continuous trajectories and areas with high motion energy as detected by the peripheral vision system. Once a face is de-

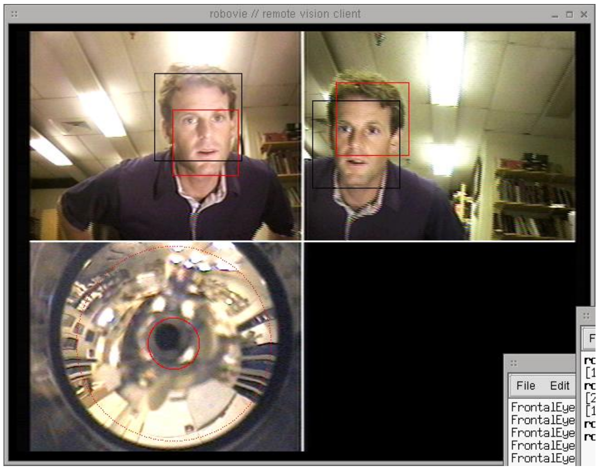


Figure 2. Here are the views that RUBI sees while tracking a face. Top are views from left and right eyes, red box is response from color tracker, black box is response from face-detector. Bottom left is the view of the omni-directional camera.

tected by the face-detector in either eye, RUBI's head controller switches to Face Tracking Mode.

2.3.5 Head Control: Face Tracking Mode

While in face tracking mode, RUBI actively moves its head by means of a Kalman controller whose goal is to minimize the distance from the most-likely probable position of the face and the center of the image plane, average across the two camera views (See Figure 2).

RUBI stays in face tracking mode if a running average $\bar{X}_t$ of the number of faces found by the face detector is above a given threshold. The value of the threshold parameter is dynamically set via reinforcement learning with the goal of optimizing the expected number of faces found by the face detector.

2.4 Auditory and Speech System

RUBI sound sensor is a VoiceTracker 8 microphone array with adaptive beam-forming. Speech detection and speech recognition is handled by the SONIC speech recognition engine from the Center for Spoken Language Research at the University of Colorado Boulder. We are currently training a new noise model to reduce the speech detector's sensitivity to robot motor noise. The detected speech is used to trigger contingent speech-like responses. These speech-like responses consist of baby vocalizations of varying length and pitches. The length of speech spoken to RUBI modifies the length of the response such that longer speech gives longer responses. By changing the characteristics of these responses we can test contingency parameters when interacting with infants in social experiments.

2.5 Social Movements

RUBI combines motor control with three social behaviors: face tracking, speech detection and response, and external environment contingency. RUBI's motor control system currently has 3 components; neck control (azimuth, elevation, roll), eye control (azimuth, elevation, zoom), and control of external objects. In adding to tracking faces, the eyes and head are used to perform social motor actions such as head nodding and gaze shifting. External objects can also be controlled via wireless Ethernet to allow RUBI to respond in a contingent manner to toys or lights when turned on or off. RUBI's architecture combines these smaller action into groups of behaviors and allows for recording and playback of social interactions. Currently RUBI's behaviors are not coordinated for socially interaction. The planned research will determine the timing of switching behaviors and the contingency of motor control for each behavior to optimize social interaction between robot and humans.

2.6 Computational System

RUBI is powered by two dual-processor 2.8 GHz Intel P4 Xeon PCs with 512 RAM connected via gigabit Ethernet. RUBI's PCs use Red Hat Linux 7.3 with open-source drivers. The first PC currently handles the face-detection and color-detection on both eyes, voice detection, and peripheral vision. In the future the first PC will handle the face-detection, color-tracking, periphery vision, and all eye movements (neck movements are handled by the Galil DMC-1832). The second PC will handle speech-detection/recognition, arm/hand movements, and emotion recognition.

3 Current Performance

Currently RUBI is capable of updating the color tracker for each eye in 15 milliseconds and the face detector in 70 milliseconds. To account for both eyes RUBI spends 15 milliseconds correcting for the error in the left eye and 15 milliseconds correcting the error in the right eye. With a frame rate of 30 Hz RUBI is capable of tracking a face at speed up to a 60 degrees/sec.

Figure 3. RUBI in a pilot study with an infant. Here RUBI (near the camera) and the subject attending to a toy.

4 Planned Research

4.1 Autistic Children Study

RUBI is being developed as part of the MESA project sponsored by the National Association for Autism Research (NAAR). The goal is to investigate the effects of contingency and timing on real-time social interaction in typically developing children and in children with autism. Experiments will be conducted in the near future at UCSD's Autism Research Laboratory. The child will be held by a parent to keep her/him comfortable (see figure 3). There will be electronic toys in the room that RUBI can attend to and actively turn off and on via wireless communication. During the course of the experiment we will alter the timing of three different social components. The first is the timing and duration of RUBI attending to the children. The second is the probability of response and timing and length of responses contingent on the children's speech. The third is the probability of response to the behavior of toys, and timing and duration of this response.

4.2 Infant Study

The infant study will have the same setup and environment as the Autistic study. The only difference between the two studies is that the experiment will be run with 18 month old infants. This will minimize the possibility of the participants having prior knowledge or expectations about robots. The infant study will have two control groups: a human (stranger) and an object to replace RUBI. The object will contribute the same level of human-like features as RUBI. RUBI's timing and contingency factors will be altered to find infant responses that are more consistent with the results of the human control study than to the results of the object control study.

5 Conclusion

The goal for the design of RUBI is to create a robot that can test child/infant responses to variances in robot behaviors. We want to determine general robot movements and behaviors that will combine current Artificial Intelligence programs and to create social behaviors that Breazeal refers to as believable behaviors [1]. For robots to work with autistic children, students or the elderly they will need the participants to be socially engaged. From the research we perform using RUBI we hope to gain a better understanding of how to keep people socially engaged while interacting with robots.

References

[1] C. Breazeal. *Designing Sociable Robots*. MIT Press, Cambridge, MA, 2002.

[2] I. Fasel, G. O. Deak, J. Triesch, and J. R. Movellan. Combining embodied models and empirical research for understanding the development of shared attention. In *Proceedings of the second international conference on development and learning*. ICDL, 2002.

[3] Ian Fasel, Bret Fortenberry, and J. R. Movellan. A generative framework for real-time object detection and classification. *Computer Vision and Image Understanding*, In Press.

[4] H. Ishiguro, T. Ono, M. Imai, and T. Kanda. Tdevelopment of an interactive humanoid robor ”robovie” – an interdisciplinary approach. In R. A. Jarvis and A. Zelinsky, editors, *Robotics Research*, pages 179–191. Springer, 2003.

[5] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu. Development and evaluation of an interactive humanoid robot ”robovie”. In *IEEE International Conference on Robotics and Automation*, pages 1848–1855, 2002.

[6] J. R. Movellan, J. Hershey, and Josh Susskind. Large scale convolutional HMMs for real time video tracking. *Computer Vision and Pattern Recognition*, 2004.

A Development Approach for Socially Interactive Humanoid Robot

Takayuki Kanda, Hiroshi Ishiguro
ATR Intelligent Robotics and Communication Labs., Kyoto JAPAN, {kanda,ishiguro}@atr.jp

Abstract

This paper describes our social approach for an interactive humanoid robot that understands human social relationships. Our interactive robot autonomously interacts with humans with its human-like body properties, and as a result, induces humans' friendly group behavior in front of it. Based on these feature as well as inspired by the survey in psychology research about friendship, we suggest a friendship estimation model for the interactive robot, which is an ability that is probably essential for interactive robots to establish social relationships with humans. As a result of a field experiment, the fundamental part of the estimation model is supported. We believe these results suggest the positive perspectives of our development approach.

1. Introduction

Recent progress in robotics has brought with it a new research direction known as "interaction-oriented robots." These robots are different from traditional task-oriented robots, such as industrial robots, which perform certain tasks in limited applications. Interaction-oriented robots are designed to communicate with humans and to be able to participate in human society. We are trying to develop such an interaction-oriented robot that will exist as a partner in people's daily lives. We believe these robots will not only used for entertainment, but also provide it with communication support task such as route-guidance and mental support task.

Several researchers are endeavoring to realize the interaction-oriented robots, such as Aibo, and Kismet [1]. Moreover, there are several research works that explore the application of the interactive robots. Shibata et al. successfully applied a seal-like pet robot Paro for mental care for elderly person [2]. Dautenhahn et al. has applied a simple interactive robot for autism therapy [3]. These research efforts seem to be devoted to social robots that are embedded in human society.

The research question we are struggling to solve is "how can interaction-oriented robot participate in human daily life, establish social relationships with humans, and contribute to the society?" In other words, our purpose is to realize a peer-partner robot that socially communicates with humans to support their daily lives.

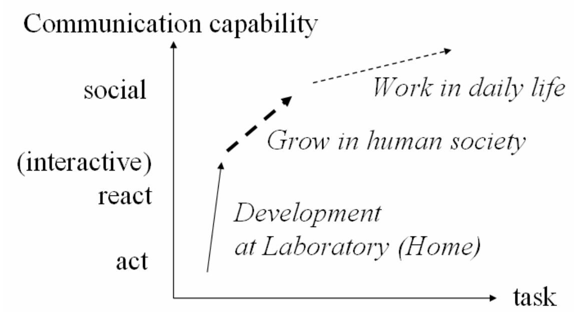


Figure 1: Development approach for social ability

We believe that the social ability of the robots will be greatly improved by putting these robots into human society. The initial tasks of the robot will be limited and perhaps not so important, because current interaction abilities of the robots are not so much as human infants' and the social skill is very little. However, we can improve the social skills of the robot in society by finding various problems that robots will suffers, which are similar development steps as human infants'. Figure 1 describes our development approach toward such the interaction-oriented robot.

Currently, robots are applied to work in our daily lives as interactive robots, and gradually growing their interactive abilities; but not to social works that requires to socially communicate with more than one people. While previous research works have developed robots' interactive abilities for only one people in front of the robots, we believe it is also indispensable to improve robots' social ability to make robots work in our daily lives, which is the approach described as broken lined arrow in the figure. We believe that robots' task will be emerged according to the improvements of robots' ability, even if current robots equipped with a little skills to accomplish useful tasks in human society.

We are achieving this development approach of making robots to gradually work in our daily lives for improving their social abilities as well as for exploring the possible tasks of the robots. While there are several learning-based approaches for understanding human-beings such as cognitive developmental robotics [4], we rather implement interactive behaviors into robots with a try-and-error manner [5]. Because, at an early stage of development, we do not know the appropriate strategy for learning-based development. Instead, we try to implement the interactive robot that is capable of socially communicating with humans, which is probably connected to such learning-based approach in future.

The first step of this development approach was a field trial in an elementary school where interactive robots behave as peer tutor of foreign language, as reported in [6]. The robot Robovie equipped with person identification function to distinguish children, such as for calling names of children, and simultaneously interacted with more than one child. As a result, it proved the positive possibility that interactive robots can motivate children to learn foreign language through the interaction with robots. Meanwhile, we have observed the group behavior among friend around the robot. For instance, a boy and his friend counted how many times the robot called their respective names, and his name was called more often, so he proudly told his friend that the robot preferred him. If the robot would understand their friendship, it could promote the interaction with the boys and the interaction between the boys. That is, the ability of friendship estimation will enable robots to mediate interaction between humans. Moreover, the friendship is tightly connected to social relationships (described in the next section in detail). Thus, this friendship estimation is essential to accomplish more general social relationship estimation, which probably make possible of the future social robot that helps to solve bullying problem or isolate children problem. In this paper, we report our approach to estimating human friendship by using an interactive robot, an ability that is probably essential for interactive robots to establish social relationships with humans.

2. Friendship estimation model from observation

2.1. Related research about friendship

It is a well-grounded finding from psychological research that children at a very young age engage in dyadic relationships, for example in the form of pretend play which then increase in size and complexity with age, forming many different peer relationships in the form of social networks. As children gradually establish social networks, each child gets a different social status, such as popular, average, isolated, and rejected [7, 8].

A sociometric test have been used to investigate the peer relationships and social networks, which lets a human directly answer the name of others whom he/she likes and dislikes. It is well validated and considered as reliable forms of assessment for human peer relationships. It distinguishes each child's social status in the groups: popular, average, neglected, rejected and controversial [9, 10]. It has been widely used to determine the relationships in a classroom or a company.

On the other hands, there are observation-based methods for understanding peer relations and social status. Children forms group and behaves with the group, along with their friendly relationships. Children usually play with peers, while boys tend to play in group and girls tend to play with only 1 other girl [11]. Ladd et al. investigated the relationships between observed group behavior and their relationships. They coded videotape about children's play with the four of the behavioral measure: cooperative play, rough play, unoccupied, and teacher-orientation. It revealed that cooperative play was associated with positive nominations while rough play related to negative nominations. In addition, they revealed that past behavior was successfully predict the current peer status, such as time spent in cooperative play was significant predictor to positive nomination [8]. Coie et al. have investigated the difference between popular and rejected children in terms of their behavior, and revealed the relationships between rejected children and their aversive behaviors [12]. We believe these findings positively support the possibility that social robots can recognize humans' peer relationships and social status by observing their group behavior.

2.2. Friendship estimation model

Human behavior is largely based on social relationships which can be in the form of dyadic relationships, known as friendship, or larger groups known as social networks where there are complex peer relationships between different individuals. Since the previous research works have proved the correlations children's group behavior and their relationships [8, 11, 12], we believe we can estimate their peer relationships and social networks from observation of their group behavior. We focused on the *estimation* of peer relationships, which are the fundamental parts of the social network, as the early attempt for *recognition* of peer relationships and social network. Yet it is not *recognition* (find all correct information accurately) but *estimation* (partly find correct information with moderate accuracy), robots can utilize these obtained information to further promote human-robot interaction.

The basic idea is "a robot autonomously interacts with several children simultaneously to cause their spontaneous group behavior, and observe the group behavior to recognize their relationships," which is our hypothesis to verify. Our friendship estimation model is based on the association of social group behavior and social relations, which is inspired by previous

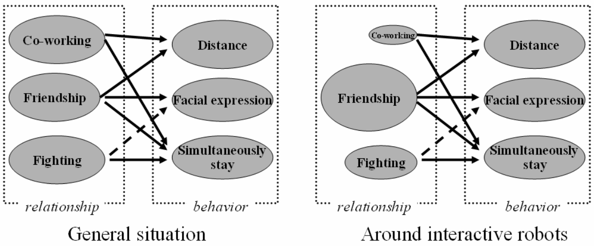


Figure 2: Relations between social relationships and group behavior

Figure 3: Scenes of friend accompanying behavior in front of an interactive humanoid robot

psychological research such as the above mentioned ones. In general, humans' social relationships affect on their group behavior, such as accompanying, distance among them, facial expression during conversation, and so forth. For instance, human is accompanied by friendly one, but not willingly approach to dislike one (accompanying and close distance). Sometimes, such the dislike relations might cause a quarrel or fight (distance will be close, but facial expression will be far from friendly). Meanwhile, official relationships rather than private one sometimes cause non-spontaneous group behavior. For instance, teacher may organize co-working activity such as "children collaborate to carry a heavy box." The left figure in Figure 2 describes these examples of the associations between group behaviors and peer relations in general situation.

On the other hand, according to our hypothesis, interactive robot mostly causes spontaneous friendly behaviors. In fact, we observed such the situation where a child is accompanied by his/her friend to interact with the robot as shown in Figure 4. We are going to verify this hypothesis in this paper later. Thus, we believe we can estimate such the friendly relationships by simply observing their group behavior. This idea is described in Figure 4-right. As the beginning step for the estimation, we will only utilize "accompanying" behavior that will be recognized by using wireless tag system.

2.3. Algorithm

Figure 4-left indicates the mechanism of the friendship estimation. From a sensor (in this case, wireless ID tags and receiver), the robot constantly obtains the IDs (identifiers) of individuals who are around it. The robot

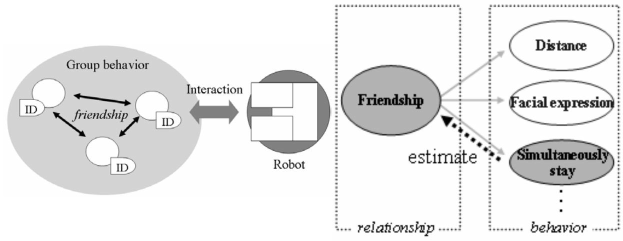


Figure 4: Current estimation model for friendship

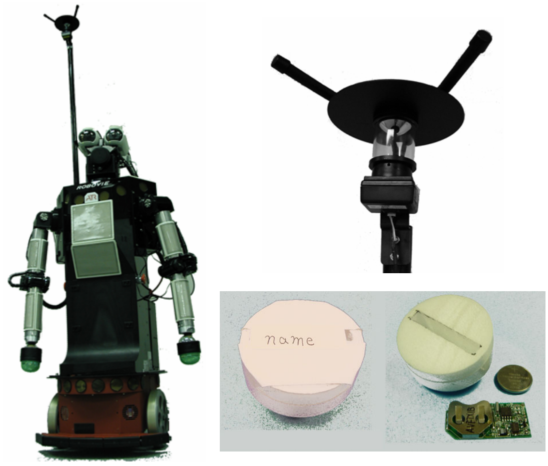


Figure 5: Robovie (left) and Wireless tags

continuously accumulates its interacting time with person A (T_A) and the time that person A and B simultaneously interact with it (T_{AB}, which is equivalent to T_{BA}). We define the estimated friendship from person A to B ($Friend(A \rightarrow B)$) as

$$Friend(A \rightarrow B) = if\,(T_{AB} / T_A > T_{TH}), \quad (1)$$

$$T_A = \Sigma\, if\,(observe(A) \text{ and } (St < S_{TH})\,) \cdot \Delta t, \quad (2)$$

$$T_{AB} = \Sigma\, if\,(observe(A) \text{ and } observe(B) \text{ and } (St < S_{TH})\,) \cdot \Delta t\,, \quad (3)$$

where *observe(A)* becomes true only when the robot observes the ID of person A, if() becomes 1 when the logical equation inside the bracket is true (otherwise 0), and T_{TH} is a threshold of simultaneous interaction time. We also prepared a threshold S_{TH}, and the robot only accumulates T_A and T_{AB} so that the number of persons simultaneously interacting at time t (St) is less than S_{TH} (Eqs. 2 and 3). In our trial, we set Δt to one second.

3. Robovie: An Interactive Humanoid Robot

3.1. Hardware of An Interactive Humanoid Robot

Figure 5 shows the humanoid robot "Robovie" [13]. The robot is capable of human-like expression and recognizes individuals by using various actuators and sensors. Its body possesses highly articulated arms, eyes, and a head, which were designed to produce sufficient gestures to communicate effectively with humans. The sensory equipment includes auditory, tactile, ultrasonic, and vision sensors, which allow the robot to behave autonomously and to interact with humans. All processing and control systems, such as the computer and motor control hardware, are located inside the robot's body.

3.2. Person identification with wireless ID tags

To identify individuals, we used a wireless tag system capable of multi-person identification by partner robots (Detailed specification and system configuration is described in [14]). Recent RFID (radio frequency identification) technologies have enabled us to use contact-less identification cards in practical situations. In this study, children were given easy-to-wear nameplates (5 cm in diameter) in which a wireless tag was embedded. A tag (Fig. 4, lower-right) periodically transmitted its ID to the reader installed on the robot. In turn, the reader relayed received IDs to the robot's software system. It was possible to adjust the reception range of the receiver's tag in real-time by software. The wireless tag system provided the robots with a robust means of identifying many children simultaneously. Consequently, the robots could show some human-like adaptation by recalling the interaction history of a given person.

3.3. Interactive behaviors

"Robovie" features a software mechanism for performing consistent interactive behaviors (detailed mechanism is described in [5]). The objective behind the design of Robovie is that it should communicate at a young child's level. One hundred interactive behaviors have been developed. Seventy of them are interactive behaviors such as shaking hands, hugging, playing paper-scissors-rock, exercising, greeting, kissing, singing, briefly conversing, and pointing to an object in the surroundings. Twenty are idle behaviors such as scratching the head or folding the arms, and the remaining 10 are moving-around behaviors. In total, the robot could utter more than 300 sentences and recognize about 50 words.

Several interactive behaviors depended on the person identification function. For example, there was an interactive behavior in which the robot called a child's name if that child was at a certain distance. This behavior was useful for encouraging the child to come and interact with the robot. Another interactive behavior was a body-part game, where the robot asked a child to touch a body part by saying the part's name.

The interactive behaviors appeared in the following manner based on some simple rules. The robot sometimes triggered the interaction with a child by saying "Let's play, touch me," and it exhibited idling or moving-around behaviors until the child responded; once the child reacted,

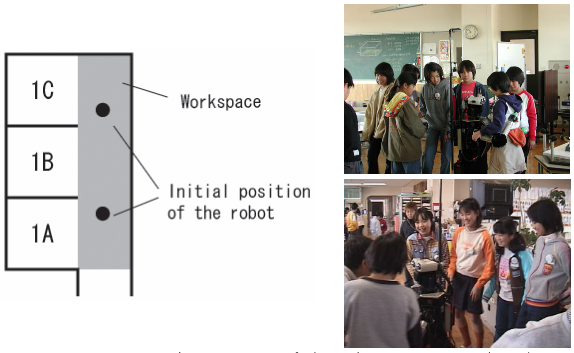


Figure 6: Environment of the elementary school

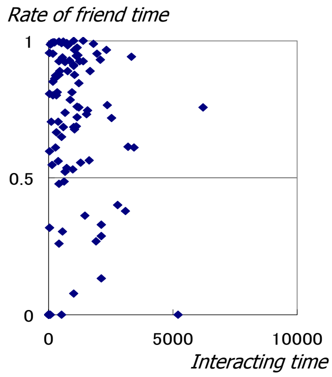


Figure 7: Frequency of friend accompanying behavior

it continued performing friendly behaviors for as long as the child responded. When the child stopped reacting, the robot stopped the friendly behaviors, said "good bye," and restarted its idling or moving-around behaviors.

4. Experiment and Result

We conducted a field experiment in an elementary school for two weeks with the developed interactive humanoid robot, which was originally designed to promote children's English learning. As we reported in [4], the robots had a positive affect on the children. In this paper, we use the interaction data during that trial as a test-set of our approach to reading friendship from the children's interaction.

4.1. Method

We performed an experiment at an elementary school in Japan for two weeks. Subjects were sixth-grade students from three different classes, totaling 109 students (11-12 years old, 53 male and 56 female). There were nine school days included in those two weeks. Two identical robots were placed in a corridor that connects the three classrooms (Figure 6). Children could freely interact with both robots during recesses (in total, about an hour per day), and each child had a nameplate with an embedded

Table 1: Estimation results with various parameters

coverage / reliability		T_{TH} (simultaneously interacting time)					
		0.3	0.2	0.1	0.05	0.01	0.001
S_{TH} (num. of simultaneously Interacting children)	2	0.01	0.02	0.03	0.04	0.04	0.04
		1.00	0.93	0.79	0.59	0.54	0.54
	5	0.00	0.02	0.06	0.11	0.18	0.18
		1.00	1.00	0.74	0.47	0.29	0.28
	10	0.00	0.00	0.04	0.13	0.29	0.31
		-	1.00	0.74	0.46	0.23	0.20

('-' indicates that no relationships were estimated, so reliability was not calculated)

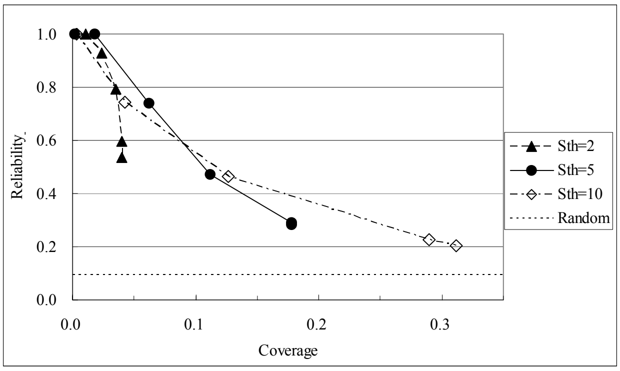


Figure 8: Illustrated estimation results

wireless tag so that each robot could identify the child during interaction.

We administered a questionnaire that asked the children to write down the names of their friends. This obtained friendship information was collected for comparison with the friendship relationships estimated by our proposed method.

4.2. Results for frequency of friend-accompanying behavior

As we compared the questionnaire on friendships and the interacting time with the robot, we found the higher frequency with which children interacted with the robot in the company of his/her friend (see Figures 7). Seventy-two percent of their interaction time with the robot was in the company of one or more friends. We believe that this result supports our hypothesis "our interactive robot mostly cause friendly accompanying behavior of children around it rather than the other behavior associated with non-friendly relationships, such as hostile, dislike, co-working". It implies we can estimate their friendship by even simply observing their accompanying behavior.

4.3. Results for friendship estimation

Since the number of friendships among children was fairly small, we focused on the appropriateness (coverage and reliability) of the estimated relationships. This is similar to the evaluation of an information retrieval technique such as a Web search. Questionnaire responses indicated 1,092 friendships among a total of 11,772 relationships; thus, if we suppose that the classifier always classifies a relationship as a non-friendship, it would obtain 90.7% correct answers, which means the evaluation is completely useless. Thus, we evaluate our estimation of friendship based on reliability and coverage, which are defined as follows.

Reliability = number of correct friendships in estimated friendships / number of estimated friendships

Coverage = number of correct friendships in estimated friendship / number of friendships from the questionnaire

Table 1 and Fig. 8 indicate the results of estimation with various parameters (S_{TH} and T_{TH}). In Fig. 8, random represents the reliability of random estimation where we assume that all relationships are friendships (since there are 1,092 correct friendships among 11,772 relationships, the estimation obtains 9.3% reliability with any coverage). In other words, random indicates the lower boundary of estimation. Each of the other lines in the figure represents the estimation result with different S_{TH}, which has several points corresponding to different T_{TH}. There is obviously a tradeoff between reliability and coverage, which is controlled by T_{TH} ; S_{TH} has a small effect on the tradeoff, S=5 mostly performs better estimation of the friendship, and S=10 performs better estimation when coverage is more than 0.15. As a result, our method successfully estimated 5% of the friendship relationships with greater than 80% accuracy (at "S_{TH}=5") and 15% of them with nearly 50% accuracy (at "S_{TH}=10") (these early findings about friendship estimation, which are reported in this subsection, has been already appeared in our previous paper [15]).

5. Conclusion

We proposed a social development approach for an interactive robot that is capable of communicating with humans socially. According to the approach, we have applied an interactive humanoid robot for a language-education task, where we have found a lack of social skill of the robot. As we have found the robot cause a friend-accompanying behavior, the robot causes human social behavior to understand their social relationships. This friend estimation model for social robot was partly verified by the field experiment. In the field experiment, two identical interactive humanoid robots placed in an elementary school for two weeks, where children freely interacted with the robots during recesses. These developed interactive humanoid robots identify individual child by using wireless tag system, which is utilized for recording individual and friend-related interaction time as well as for promoting the interaction by such as calling their names. The result suggested that mostly children were accompanied with one of more friend (72% of the total interacting time), and the robot was successfully estimated friendly relationships partly (for example, 5% of the all relationships with 80% accuracy). We believe that this early findings would encourage further research in social skill of social robots as well as the sensing technology for autonomous observation about inter-human and human-robot interaction.

Acknowledgement

We wish to thank the teachers and students at the elementary school for their agreeable participation and helpful suggestions. We also thank Takayuki Hirano and Daniel Eaton, who helped with this field trial in the elementary school. This research was supported by the NICT of Japan.

References

[1] Breazeal, C. and Scassellati, B. (1999). A context-dependent attention system for a social robot, *Proc. Int. Joint Conf. on Artificial Intelligence,* 1146-1151.

[2] T. Shibata, and K. Tanie (2001). Physical and affective interaction between human and mental commit robot, *Proc. of IEEE Int. Conf. on Robotics and Automation,* pp.2572-2577.

[3] K. Dautenhahn, and I. Werry (2002). A quantitative technique for analysing robot-human interactions, *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp.1132-1138.

[4] M. Asada, K. F. MacDorman, H. Ishiguro, and Y. Kuniyoshi. Cognitive Developmental Robotics As a New Paradigm for the Design of Humanoid Robots. *Robotics and Autonomous System*, Vol.37, pp.185-193, 2001.

[5] Kanda, T., Ishiguro, H., Imai, M., Ono T., and Mase, K. (2002). A constructive approach for developing interac-tive humanoid robots, IEEE/RSJ International Conference on Intelligent Robots and Systems, 1265-1270.

[6] Kanda, T., Hirano, T., Eaton, D., Ishiguro, H., ``Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial," *Journal of Human Computer Interaction*, 2004.

[7] Gottman, J. M., & Parkhurst, J. (1980). A developmental theory of friendship and acquaintanceship processes. In W. A. Collins (Ed.), *Minnesota symposia on child psychology,* (Vol. 13, pp. 197-253). Hillsdale, NJ: Lawrence Erlbaum.

[8] Ladd G. W., Price J. M., Hart C. H. (1990), Preschooler's behavioral orientations and patterns of peer contact: predictive of peer status?, In Asher S. R. & J. D. Coie (Eds). *Peer rejection in childhood,* pp.90-115. Cambridge Univ. Pr.

[9] McConnell, S. R. & Odom, S. L. (1986), Sociometrics: Peer-referenced measures and the assessment of social competence. In P.S.Strain, M.J.Guralnick, & H.M.Walker (Eds), *Children's social behavior: Development, assessment, and modification.* pp. 215-284, Orlando: Academic Press

[10] Asher S. R. & Hymel. S. (1981). Children's social competence in peer relations: Sociometric and behavioral assessment. In J. D. Wine & M. D. Smye (Eds.), *Social competence* (pp.125-157), New York: Guilford.

[11] Waldrop, M. F. & Halverson, C. F. (1975). Intensive and extensive peer behavior: Longitudinal and cross-sectional analyses. *Child Development,* Vol. 46, pp. 19-26.

[12] Coie J. D., & Kupersmidt J. B. (1983). A behavioral analysis of emerging social status in boys' groups. *Child Development,* Vol. 54, pp. 1400-1416.

[13] Ishiguro, H., Ono, T., Imai, M., and Kanda, T. (2003). Development of an interactive humanoid robot "Robovie" -An interdisciplinary approach, R. A. Jarvis and A. Zelinsky (eds.), *Robotics Research,* Springer, pp. 179-191.

[14] Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. (2003). Person Identification and Interaction of Social Robots by Using Wireless Tags, *IEEE/RSJ International Conference on Intelligent Robots and Systems,* pp. 1657-1664.

[15] Kanda, T., and Ishiguro, H., Reading human relationships from their interaction with an interactive humanoid robot, *Int. Conf. on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE)*,

Communicative behavior to the android robot in human infants -Preliminary report-

Itakura, S., Kanaya, N., Shimada, M., Minato, T., & Ishiguro, H.

Abstract

We investigated whether human infants show some social behavior to the android robot when it behaves like human. Infant sat mother□s lap and face to the robot. There were two toys which can move by remote controller. A procedure was as follows: When the infant looked at the robot, it turned her head toward one of the toys until the infant followed her head turn. In-focus condition, as soon as the infant followed the head turns of the robot to look at the In-focus toy, the experimenter activated the In-focus dog which began to move. Out-of-focus condition, the experimenter activated the dog at which the robot was not looking. We recorded infant□s behaviors, such as referential looking or pointing or vocalization. The ratio of occurrence of visual checking increased when the robot did not look at the moving toy. Visual checking is looking at the moving toy and the robot alternatively. Infant produced more visual checking in the out-of-focus condition than in the in-focus condition. Infants seem to understand robot□s attentional state. This result is different from Legerstee□s study. Our robot is not just life-sized doll, it looks just like a human.

Attention detection and manipulation between autonomous four-legged robots

Frédéric Kaplan, Verena V. Hafner and Andrew Whyte
Sony Computer Science Laboratory Paris
Developmental Robotics Group
6 rue Amyot 75005 Paris France
kaplan@csl.sony.fr, hafner@csl.sony.fr, andrew@csl.sony.fr

Skills for attention detection and manipulation are crucial prerequisites underlying the development of social cognition. Through a series of steps of increasing complexity, children manage to make progress in directing the attention of their parent and in interpreting gaze direction and pointing gestures. Building robots capable of engaging in such kind of interactions is now a major topic in the developmental robotics community (e.g. [4, 5]). We report results of a set of experiments conducted with a population of AIBO ERS-7 robots showing how it is possible for a robot to (a) interpret the attentional behavior of another robot and (b) use pointing gestures in order to influence the attentional behavior of another robot. One of the robots takes the role of an adult and points to an object, the other robot, the learner, has to interpret the pointing gesture correctly in order to find the object [2]. We show that motivation and intrinsically rewarding stimuli play a crucial role in this development (as already advocated by Carlson and Triesch [1]). These initial results permit a better understanding of the challenges that remain to be addressed for building robots capable of joint attention [3]. In the line of Tomasello's views [6], we argue that joint attention is much more than simultaneous looking. Beside attention manipulation and detection it involves the development of skills for social coordination and more importantly some form of intentional understanding. Constructing robots capable of developing a shared intentional relation to the world is probably one of the hardest problem developmental robotics has to tackle.

References

[1] E. Carlson and J. Triesch. A computational model of the emergence of gaze following. In *Proceedings of the 8th Neural Computation Workshop (NCPW8)*, 2003.

[2] V.V. Hafner and F. Kaplan. Learning to interpret pointing gestures: experiments with four-legged autonomous robots. In *Proceedings of the KI2004 Workshop on Neurobotics*, 2004.

[3] F. Kaplan and V.V. Hafner. The challenges of joint attention. In L. Berthouze, H. Kozima, C. Prince, G. Sandini, G. Stojanov, G. Metta, and C. Balkenius, editors, *Proceedings of the 4th International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic System*, pages 67–74. Lund University Cognitive Studies 117, 2004.

[4] H. Kozima and H. Yano. A robot that learns to communicate with human caregivers. In *First International Workshop on Epigenetic Robotics (Lund, Sweden)*, 2001.

[5] Y. Nagai, K. Hosoda, A. Morita, and M. Asada. A constructive model for the development of joint attention. *Connection Science*, 15(4):211–229, 2003.

[6] M. Tomasello. Joint attention as social cognition. In C. Moore and P.J. Dunham, editors, *Joint attention: its origins and role in development*, pages 103–130. Lawrence Erlbaum Associates, 1995.

Figure 1. The adult robot point an object out to the other one.

Facial Expression in Social Interactions: Automatic Evaluation of Human-Robot Interaction.

G.C. Littlewort, M.S. Bartlett, I. Fasel, J. Chenu, T. Kanda, H. Ishiguro, J.R. Movellan

Abstract

We present a pilot study to evaluate the automatic facial expression classification system (AFEC) developed at the Machine Perception Laboratory, as a tool for automatically measuring the quality of human-robot social interaction (Movellan 2003).

The AFEC system automatically detects frontal faces in the video stream, using a cascade of weighted integral image features (Fasel 2002) and codes them with respect to 7 dimensions in real time: neutral, anger, disgust, fear, joy, sadness and surprise. The expression classification combines Adaboost feature selection and SVM's. The generalization performance to new subjects for a 7-way forced choice was over 90\% correct on two publicly available datasets. (Littlewort 2003) The output codes change smoothly as a function of time, providing a potentially valuable representation to code facial expression dynamics in a fully automatic and unobtrusive manner.

The AFEC system was deployed for measuring spontaneous facial expressions in the continuous video stream during unconstrained interaction with a social robot. Subjects interacted with RoboVie, a communication robot developed at ATR and the University of Osaka. Facial expression measurement of joy obtained by the automated system was compared to human judgments of joy obtained by turning a dial. The system predicted human judgments of joy in test sequences. Equipping robots or computer animated agents with the perceptual primatives necessary to use and learn from sub-second nonverbal communications opens a new realm of possible applications for machine intelligence, from entertainment robots to perceptive teaching software.

How children understand other's belief before they develop attentional flexibility?

Yusuke Moriguchi and Shoji Itakura
Department of Psychology, Graduate School of Letters Kyoto University, Yoshida-honmachi, Sakyo, Kyoto, Japan. yusukemoriguchi@yahoo.co.jp

Abstract

Recent studies had shown that there is developmental link between theory of mind and self-control ability. According to those studies, 3-year-old children who were not supposed to have theory of mind did not develop attentional flexibility, one of the main functions of self-control ability. In this study, we investigated how 3-year-old children understand other's belief before they develop attentional flexibility. In the Experiment, preschoolers were given a card sorting task. Prior to starting the task, they were shown that the demonstrator sorts the cards incorrectly. There were four conditions, control condition correct belief condition, false belief condition and misunderstanding condition, according to the demonstrator's belief. Children needed attentional flexibility to solve this task correctly. The results showed that there was significant difference of performance between conditions in 3-year-old children, and that other's belief affected the performance of children in the task which needed attentional flexibility. This suggested that even 3-year-old children, whose attentional flexibility was immature, could discriminate the other's belief implicitly. We discussed the relationship between theory of mind and self-control ability from our results and proposed the new theory.

1. Introduction

1.1. Self-control ability

The ability to control attention and behaviour becomes more efficient during preschool years [1, 2]. One task frequently used to examine children's self-control is the dimensional change card sorting task (DCCS) [3]. In this task, children were given two target cards (e.g., red rabbits and blue boats) and sorting cards (red boats and blue rabbits), featuring two dimensions, and matched one target on one dimension and the other target on the other dimension. The task had two phases. In the first phase, the children were told a rule of the game which specifies how to sort the test cards according to a particular dimension, e.g. shape dimension. The experimenter said to children "Let's start a game. This game is the shape game. In this game, all the rabbits go here (pointing to a blue rabbit target) and all boats go there (pointing to a red boat target)." When the children finished five trials, the rule was changed to the color rules. The experimenter said to children "Now we are playing another game. We are playing the color game. In this game, all the red cards go here (pointing to a red boat target) and all blue cards go there (pointing to a blue rabbit target)" The children were given five trials without feedback. In this task, three-year-old children could sort the cards according to first dimension, but they could not inhibit the preceding response and showed perseverance with the old dimension when they were requested to sort according to the new dimension. On the other hand, most of the 4- and 5-year-old children could switch between the rules and sort the cards correctly

Gerstadt, Hong and Diamond [4] had observed similar difficulty in preschoolers using a Stroop-like task. In their study, children were required to keep to rules set by the experimenter and resist the temptation to say what the stimuli really represented. When presented with black/moon cards, the correct response was "day", and when presented with white/sun cards, the correct response was "night". Children between 3 and 4 years of age showed difficulty performing this task, saying "night" when they saw the black/moon cards and "day" when they saw the white/day cards. On the other hand, 5-6-year- old children age were able to respond correctly..

One of the theories used to explain young children's difficulty in this kind of task is attentional inflexibility theory. Moriguchi and Itakura [5] postulated that children should develop the ability to inhibit attention to irrelevant information and behavior. Younger children's perseveration is caused by a failure to inhibit attention to old rules. The tasks used included two alternatives (e.g., shape rule and color rule). One of two alternatives was salient for children for various reasons (e.g. familiarity, reinforcement by feedback, more attractive object) and in the tasks children needed to suppress their attention to the salient alternative in order to use another alternative. Five-year-old children were able to inhibit their attention to the salient alternatives and switch their attention to less salient ones and thus control their choices, but 3-year-old children could not; the latter perseverated to the salient alternative.

1.2. Theory of Mind

Theory of mind is one of the most important topics in developmental science. Young children show difficulty understanding that people have the mental state such as belief and knowledge, but as they get older they acknowledge the representational quality of mental states [6]. Most popular task to measure theory of mind is the false belief task. Two puppets (Bert and Ernie) played with a ball bliefly and then Bert put the ball in a blue container and left. Ernie retrieved the ball, played briefly with it and then put it away in a red container and left. Finally, Bert returned , wanting to play with the ball, and children were asked the False belief question ("Where does Bert think the ball is?") followed by the reality question.

Wellman, Cross and Watson [7] presented a meta-analysis of these kind of tasks and showed that children who were 3 years 5 months or younger performed below chance and made the error in the false belief question. Children who were 4 year or older performed above chance and could recognize that people held the false belief.

1.3. Self-Control Ability and Theory of Mind

Recent studies suggest that the development of self-control is related to the development of theory of mind. Perner and Lang [8] reviewed studies of the relationship between self-control and theory of mind, and concluded that self-control tasks (e.g. DCCS) and theory of mind tasks (e.g. false belief task) have something in common. For example, although typically developed children showed a four-year shift in both tasks, difficulties are encountered by many atypically developed children (e.g. autism) on both tasks. Many studies have shown a strong correlation between theory of mind tasks and self-control tasks in preschoolers [9,10]. Perner, Lang and Kloo [10, Experiment1] gave preschoolers two versions of a false belief task to assess theory of mind, and the Dimensional Change Card Sorting task (DCCS) as a self-control task. They found a significant positive correlation between two of false belief tasks and the card sorting task.

Self-control ability consisted of several functions such as planning, attentional flexibility, inhibitory control , error detection and correction, and so on [11,12]. Moriguchi and Itakura [5] suggested that one function of self-control ability might be related to the development of theory of mind. They used the card sorting tasks which had the correlation with theory of mind tasks and showed the possibility that the development of attentional flexibility was related to the development of theory of mind.

In their study, they modified the DCCS and gave the children three phases. The first phase was the same as DCCS. In the second phase, the children practiced directing their attention to the new dimension and then in the third phase they were asked to sort the cards according to the new dimension. Compared with the standard DCCS, the children's performance did not improve. This result suggested that the children could not switch their attention to the new dimension in the presence of conflicting cues, which were cues related to old rules. Therefore, in the subsequent experiment, there were two phases and they removed the old dimension from the targets in the second phase. Thus the children did not face conflicting situations in this phase. In this experiment, the children were able to direct their attention to the new dimension easily. This result suggested that the 3-year-old children have difficulty switching attention when faced with conflicting situations, and that 3-year-old children lacked the attentional flexibility.

From this result and the fact that there is the positive correlation between theory of mind tasks and DCCS, the development of attentional flexibility might be related to the development of theory of mind

1.4. Purpose of the present study

Earlier studies suggested that the development of theory of mind was related to the development of the attentional flexibility, but it is not still clear that how children might understand other people's mental states before they develop attentional flexibility. Some studies suggested that children who were 3 year or younger had the implicit understanding of theory of mind [13], but there were no studies about their implicit understanding with respect to the attentional flexibility. In the present study, using the new paradigm, we investigated how 3-year-old children could understand the other's belief when their attentional flexibility was immature. In a card sorting task requiring attentional flexibility, children watched another person performing the task after being told the rule, but the demonstrator sorted the cards according to the wrong rule. After the demonstrator's performance, children were required to sort the cards according to the rule experimenter had announced first. There were two conditions; in each condition the demonstrator showed the same sorting behaviour, but the demonstrator's belief was different. In one condition she (the demonstrator) believed that her sorting was correct, whereas in the other condition she noticed that her sorting was wrong. To perform the task correctly, children needed to keep the correct rule in memory and inhibit attention to the wrong rule regardless of the demonstration, that is, children needed the attentional flexibility.

If 3-year-old children might understand and discriminate other's belief before they develop attentional flexibility, there was the difference of the performance between conditions.

2. Experiment

2.1 Method

Participants

Sixty-eight 3-year-old children (*M* = 42.8 months, range = 37 months to 46 months, 34 boys and 34 girls), Sixty eight 4-year-old children (*M* = 54.0 months, range 48 months to 59 months, 35 boys and 31 girls) and sixty5-year-old children (*M* = 67.8 months, range 61 months to 71 months, 30 boys and 28 girls) were recruited from nursery schools in Kyoto as participants. In this experiment, there were four conditions in which the behavior of the demonstrator was different; control condition, correct belief condition, false belief condition, and misunderstanding condition. In each condition, there were 17 3 year old children, 17 4 year old children, and 15 5 year, old children. Most children came from middle-class back grounds and had developed normally.

Materials

Laminated cards (12 cm × 8 cm) were used as stimulus. There were two target cards （a yellow house and a blue cup） to be matched. There were 6 sorting cards (3 blue houses and 3 yellow cups). The trays (13 cm ×13 cm) on which the children put the cards were transparent and were placed near the targets.

Procedure

Each child was tested individually for 5 - 10 minutes. The participant was seated at a table. There were two conditions: a correct belief condition, and a false belief condition. In both conditions two experimenters were present. Experimenter A sat at a table across from the child and Experimenter B sat next to the child. The experimenters spoke briefly with the child, and once the child appeared relaxed, the experiment began. Each condition had three phases: pretest phase, observation phase and sorting phase.

In the correct belief condition, Experimenter B was instructed to sort the cards according to the wrong dimension, to maintain a neutral facial expression, and not to express any cues that the child might identify. In the pretest phase, Experimenter A presented the child with the cards and asked the child to name the pictures ("What is this picture?"). Children were asked to label the objects according to the two dimensions (e.g. "yellow" "cup"). If they answered correctly, Experimenter A announced the rule of the game, the shape rule ("In this game, all the cups go in this tray, and all the houses go in this tray). Earlier studies [5] showed that the order of dimension did not affect children's performance, so in the present study we used the shape rule only. We then asked the child knowledge questions to make sure that the rule was understood. Experimenter A asked "Where does this (yellow cup or blue house) card go?" The child was asked to answer two questions by pointing.

After confirmation that the child could answer knowledge questions correctly, Experimenter A said, "Now she (Experimenter B) will sort the cards, so please wait and see her," and she started to sort the cards. Although experimenter B was instructed to sort the cards according to the same dimension as the child, she failed to do this; instead she sorted the cards according to color dimension. On every one of three or four trials Experimenter A asked Experimenter B "Is this sorting right?" Experimenter B noticed her mistake and said "Oh, I am mistaken." This was the observation phase.

After the observation phase, Experimenter B said "I want to go to the toilet" and went out of the room and then Experimenter A asked the child whether Experimenter B's performance was correct or not. If the child did not answer correctly, Experimenter A stated that Experimenter B had sorted the cards incorrectly. The child was then told: "Please sort the cards according to the rule I told you first." The child was given five sorting trials, with no feedback about the cards were sorted correctly.

The false belief condition was identical except that Experimenter B had a false belief in the observation phase. In the observation phase, the child watched while Experimenter B sorted the cards according to the color (wrong) dimension. On each trial, Experimenter A asked" Is this sorting right?" Experimenter B pretended not to notice her mistakes and she confidently nodded "Yes." After the observation phase, Experimenter B said "I want to go to the toilet" and went out of the room and then Experimenter A asked the child whether Experimenter B's performance was correct or not. If the child did not answer correctly, Experimenter A stated that Experimenter B had sorted the cards incorrectly.

In the misunderstanding condition, the pretest phase was the same as that in the correct belief condition. After the pretest phase, Experimenter A said, "Now she (Experimenter B) will sort the cards, so please wait and see her," and she started to sort the cards. She could sort the cards correctly. On every one of three or four trials Experimenter A asked Experimenter B "Is this sorting right?" Experimenter B misunderstood the rule of the game and said "Oh, I am mistaken." This was the observation phase.

After the observation phase, Experimenter B said "I want to go to the toilet" and went out of the room and then Experimenter A asked the child whether Experimenter B's performance was correct or not. If they could not answer correctly, Experimenter A told the child "Her performance was correct." The child was then told: "Please sort the cards according to the rule I told you first." The child was given five sorting trials, with no feedback about the cards were sorted correctly.

The control condition had no observation phase. When they finish the pretest phase, children were given a short delay, and then asked to sort the cards according to the rule the Experimenter A instructed on the pretest phase.

2.2 Result

The result was shown in Table 1. Three 3-year-old children (from the correct belief condition) were excluded from the analysis because they did not answer the pretest phase knowledge questions correctly. Also two 5-year-old children (from the misunderstanding condition) were excluded from the analysis because the experimenters failed the procedure of the experiment. The rest of children could answer the knowledge questions perfectly.

In each condition, children were instructed to sort the cards according to the shape rule on the pretest phase, so children should sort the cards according to the shape rule on the sorting phase. The dependent measure was how many trials children sorted the cards according to the shape rule.

Preliminary analyses confirmed that no significant effect of gender was found, so all data were collapsed across this variable. The score was analyzed using a 3 (age) × 4 (condition) between subject ANOVA. There was significant main effect of age F (2, 177) = 11.038, p < .0001, significant main effect of condition F (3, 177) = 7.019, p < .001 and significant interaction F (6, 177) = 2.294 p < .05.

Follow-up one-way ANOVA's were performed to examine the interaction effect, the effect of conditions in each age. They revealed no significant difference between condition for 5 year old children , but marginally difference for 4 year old children F (3, 62) = 2.4961 p = 0.68, and significant difference between conditions for 3 year old children F (3, 61) = 6.0838 p < .01. Post hoc analysis using Tukey's HSD test revealed that children in the control condition and correct belief condition performed the task better than children in the false belief condition (P <.05), but there were no significant difference between false belief condition and misunderstanding condition.

Follow-up one-way ANOVA's were also performed to examine the developmental change. They revealed no significant developmental change between control condition, correct belief condition and in the misunderstanding condition. These results, especially the result of the control condition, might be caused by the ceiling effect. On the other hand, there was significant developmental change in the false belief condition F (3, 62) = 10.2297, p < .001.

	3 year old	4 year old	5 year old
Control	4.41	4.71	4.87
Correct Belief	3.79	4.47	4.53
False Belief	1.59	3.18	4.80
Misunderstand	3.18	4.07	4.31

Table1 mean number of correct trials in the Experiment

2.3 Discussion

In the present study, we investigated whether 3-year-old children, who were not supposed to have theory of mind and mature attentional flexibility, could understand and discriminate other's belief implicitly. In this experiment, children needed to keep the rule the experimenter instructed on the pretest phase when they were shown other's demonstration on the observation phase and sorted cards according to the correct rule on the sorting phase.

We found significant main effect of age, and this suggested that children improved the performance of the task as they get older. This improvement might reflect the result of the earlier studies [3, 5], which suggested that children developed the attentional flexibility during preschool years. In this study, in each condition including the control condition, children needed to suppress the color rule when they sorted the cards on the sorting phase. Especially in correct belief condition, false belief condition and misunderstanding condition, children were shown the demonstration with wrong behavior and/or false belief, and some of the younger children were affected by the demonstrator, this suggested that they did not develop attentional flexibility. On the other hand, older children, who develop the attentional flexibility, were less affected by the demonstration.

We also found significant interaction between age and condition, and follow up analysis showed that there was significant difference between conditions in 3 year-old children, but no significant difference in 4 and 5 year old children. This result suggested that older children could keep the correct rule in each condition while the difference of the condition affected the performance of 3-year-old children. Post hoc analysis revealed that children in the false belief condition performed worse than those in the control condition and correct belief condition.

In the control condition, children were not shown any demonstration, which meant that no one interfered children's sorting according to the shape rule. In fact children in the control condition had the greatest score in each age. In the correct belief condition, the experimenter B showed wrong sorting behavior, but she had correct belief, that is, she knew that the rule of the game was shape rule. In this condition 3-year-old children performed as well as children in the control condition. This suggested that the demonstration with wrong action only didn’t affect children’s performance.

On the other hand, children in the false belief condition were more affected by the demonstrator than the control condition and correct belief condition. Especially it was surprising that we could find the significant difference between correct belief condition and false belief condition. In the observation phase of both conditions, children observed Experimenter B’s demonstration as she mistakenly sorted cards according to color. This was inconsistent with the children’s knowledge because in the pretest phase they were asked to sort the cards according to shape. In each condition, the experimenter’s behavior was the same, but the effect on children’s sorting was different. In the correct belief condition, most of the children sorted cards correctly in the sorting phase. In contrast, younger children in the false belief condition were inclined to imitate Experimenter B, probably due to being influenced by Experimenter B’s mental state. In the correct belief condition Experimenter B noticed her mistakes, possibly reinforcing the children’s confidence in the rule for this task. They paid attention to the shape-rule. However, in the false belief condition Experimenter B “believed” that her performance was correct, which might have led children to doubt their knowledge (the rule of the game) and instead think that the rule Experimenter B used was correct. Even if Experimenter A stated that Experimenter B was wrong at the start of the sorting phase, children perseverated with the color dimension. For these children, the correct rule was the color rule, so most 3-year-old children in the false belief condition sorted the cards according to color. This result suggested that 3-year-old children whose attentional flexibility was immature could discriminate other’s belief implicitly.

Does this study support the notion that there is the relationship between theory of mind and self control ability ? As described above, earlier studies showed a positive correlation between self-control and theory of mind [9]. However as Perner and Lang [8] point out, several hypotheses exist concerning this relationship. For example, Russell [14] emphasized that self-control ability is necessary for developing theory of mind. He suggested that monitoring of action and the ability to act at will is needed to develop self-awareness, and that this self-awareness is necessary for acquiring a mental concept. On the other hand, Perner [15] suggested that children need to acquire the meta-representation before they can perform self-control tasks successfully. According to Perner, children develop meta-representation when they develop theory of mind; therefore, developing theory of mind leads the development of self-control ability.

The present study suggests that children who have immature attentional flexibility might have implicit understanding of theory of mind. From this result, we would support the notion that self-control ability is necessary for developing theory of mind. We proposed that the development of attentional flexibility was needed when implicit understanding theory of mind becomes explicit. When children’s attentional flexibility is immature, they can not pass the tasks which measures self-control ability (e.g. DCCS) and theory of mind tasks (e.g. false belief task), but they can discriminate other’s belief implicitly. When they develop the attentional flexibility and they can pass the tasks which measures self-control ability, they come to understand other’s belief explicitly and pass the theory of mind tasks. In the further research, we would investigate the causal relationship between the development of attentional flexibility and the development of theory of mind.

5. Acknowlegement

This research reported here was supported in part by Basic Research 21 for Breakthroughs in Info-Communications and a grant from JSPS (No: 13610087) to Shoji Itakura.

6. Reference

[1] Luria, A. R. (1966). *Higher cortical functions in man*: Oxford, England: Basic Books.

[2] Zelazo, P. D. & Müller, U. (2002) Executive Function in typical and atypical development. In U. Goswami (Ed.), Blackwell handbook of childhood cognitive de velopment (pp445-469). Oxford, England: Blackwell.

[3] Zelazo, D. P., Frye, D., & Rapus, T. (1996). An age related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37-63.

[4] Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Per formance of children 3.5-7 years old on a Stroop-like dayight test. *Cognition, 53*, 129-153.

[5] Moriguchi & Itakura (2003) Development of self control ability in preschoolers: Development of Switching attention. *Information processing society of Japan branch of Kansai, 2003 Proceedings* 121-122

[6] Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-128.

[7] Wellman, H.M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: the truth about false belief. *Child Development*, 72, 655-684

[8] Perner, J., & Lang, B. (2000).Theory of mind and executive function: Is there a developmental relationship? In *Understanding other mind: Perspective from Autism and Developmental Cognitive Neuroscience*, Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (eds), Oxford University Press: Oxford; 150-181

[9] Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73-92.

[10Perner, J., Lang B. & Kloo D. (2002). Theory of Mind and Self-Control: More than a Common Problem of Inhibition. *Child Development 73* 752-767

[11] Welsh, M.C., Pennington, B.F. & Groisser, D.B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7, 131-149*

[12] Zelazo, P.D., Carter, A., Rexnick., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1, 198-226*

[13] Clements, W & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9 377-395*

[14] Russell, L. (1998) How executive disorders can bring about an inadequate ‘theory of mind’. In *Autism as an executive disorder*, (ed. J. Russell), pp.256-99. Oxford University Press, Oxford.

[15] Perner, J. (1998) The meta-intentional nature of executive functions and theory of mind. In P. Car ruthers & J. Boucher (Eds.), *Language and thought* (pp. 270-283). Cambredge: Cambridge University Press.

Young children's understanding of perception and false belief: Hiding objects from others.

Manuel Sprung[1] & Martin J. Doherty[2]

[1] *University of Salzburg, Helbrunnerstr. 34, A-5020 Salzburg, Phone: +43 662 80445134, Email: Manuel.Sprung@sbg.ac.at*

[2] *University of Stirling, Stirling FK9 4LA, Scotland, Phone: +44 1786 466366, Email: mjd1@stir.ac.uk*

Abstract

This paper challenges the common assumption that children understand other's visual attention from infancy. Some classic research (Flavell, Shipstead & Croft, 1978) suggests that two-year-old children can hide an object from an adult observer, by placing it behind a screen, but not by placing a screen in front of it. In a more recent study, McGuigan and Doherty (2002) replicate this classic finding and explain this effect in terms of engagement, a concept introduced by O'Neill (1996). According to McGuigan and Doherty, occluding barriers block vision and can stop people every becoming engaged with an object, but do not necessarily disrupt engagement when it has already been established. This hypothesis implies that children should be better able to occlude objects if the person is not yet engaged with the object. The present study tests this prediction with a hiding game in which children must stop an experimenter witnessing the act of hiding by placement of a screen. The results of two experiments with forty-seven 2- to 5-year-old children confirmed that children were much better able to occlude objects if the adult observer was not yet engaged with the object, in its final location.

1. Introduction

Children's knowledge about visual perception has been considered to be quite sophisticated by the age of 2 years (e.g. Lempers, Flavell, & Flavell, 1977) and before (Baron-Cohen, 1995), but as some recent studies suggest a more complex picture emerged. As McGuigan and Doherty (2002) believe this has been already evident in a study by Flavell, Shipstead and Croft (1978). They placed an opaque screen on a table, gave children a toy and asked them to hide it from the experimenter (the *Move-Object task*). Almost all children from the age of 2 ½ years were able to hide the toy behind the screen. Flavell et al. then placed the toy on the table and asked children to use the screen to conceal it from the experimenter. Younger children performed surprisingly poorly at this *Move-Screen task*, despite the fact that it seems to require exactly the same understanding as the Move-Object task. McGuigan and Doherty's study represents a strait replication of Flavell et al. findings and they explain the results in terms of engagement, a concept introduced by O'Neill (1996), who argues that children are sensitive to adults' general involvement in their activities. According to McGuigan and Doherty, occluding barriers block vision and can stop people every becoming engaged with an object, but do not necessarily disrupt engagement when it has already been established. Since 2-year-olds understand engagement but not vision, they are unable to disrupt vision by placing the screen in front of the object.

This account generates the clear prediction that younger children should pass the Move-screen task in cases where the adult is not yet engaged with the object. This study set out to test this prediction with a hiding game in which children must stop an experimenter witnessing the act of hiding by placement of a screen. Children should understand that they can prevent the experimenter ever being engaged with the object in its hiding location. .

2. Experiment 1

2.1. Method

2.1.1. Participants

Twenty-four children (14 boys, 10 girls) from two nursery schools in Salzburg (Austria) participated in this study. Their age ranged from 2;3 (years; months) to 5;5 with a mean age of 3;8 (*SD*= 1;0).

2.1.2. Design

Each child was tested on three tasks: Move-object, Move-Screen and Hiding game. The three tasks were administered in a full balanced order. Children were tested in one single sessions lasting between 15 and 20 minutes.

2.1.3. Procedure

The procedure for the Move-object and Move-screen task followed the procedure used Flavell et al. (1978) and McGuigan and Doherty (2002).

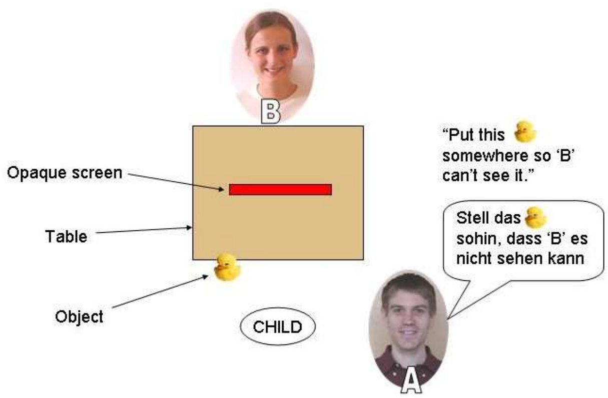


Figure 1. Move-object task

Figure 1 illustrates the procedure for the Move-object task. The child was sitting on one side of a small rectangular table and an adult observer (B) was sitting on one of the other sides of the table. An opaque screen was placed in the middle of the table and the child was asked to put an object somewhere so that the adult observer (B) can't see it.

However, in the Move-screen task, the object was placed in the middle of the table and the child was then asked to put the opaque screen somewhere so the adult observer (B) can't see the object anymore. This is illustrated in Figure 2.

In the *Hiding-game* (see Figure 3) children were shown an object and two boxes in which it was to be hidden from an adult observer (B). In this task, children were asked to use the opaque screen to prevent the adult observer (B) from witnessing the final hiding location of the object and were told to "put this somewhere so Sarah can't see us hide the duck".

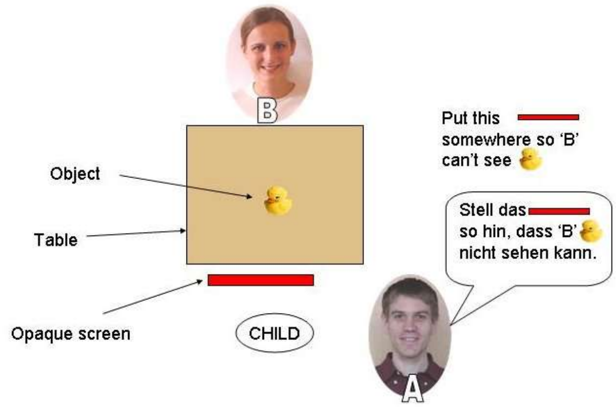


Figure 2. Move-screen task

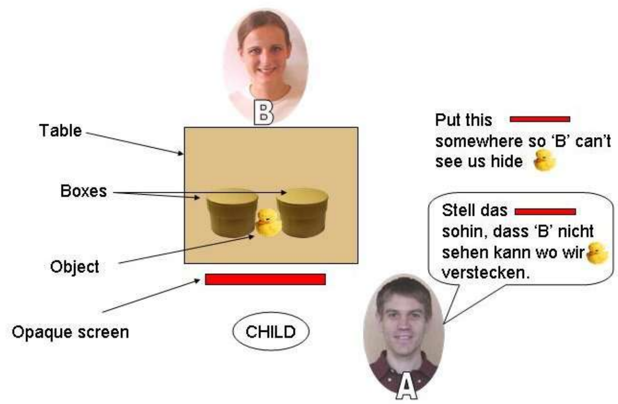


Figure 3. Hiding-game task

2.2. Results and discussion

Performance in the Hiding-game task (54% success) was intermediate between performance on the Move-object and Move-screen task (79% and 39% success respectively, p <.05 for all comparisons). The corresponding percentages correct responses are given in Figure 4. The results support the engagement hypothesis, especially considering the added general task complexity in the Hiding-game.

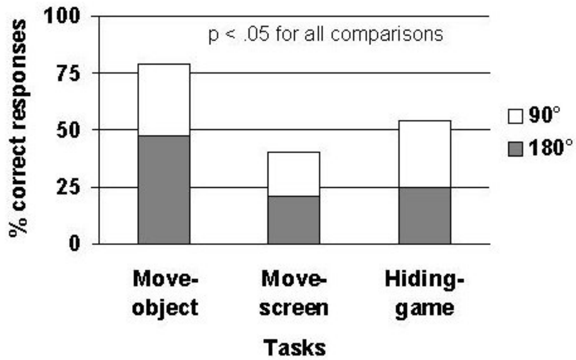


Figure 4. Percent correct response in the three tasks of experiment 1

3. Experiment 2

Experiment 2 assessed the possibility that performance in the Move-object task and in the Hiding-game was inflated by an artefact: If children simply placed the screen in the nearest convenient location - just in fronton them- this would inflate performance in the Move-object- and the Hiding-game task, but not in the Move-screen task. Therefore, the objects (and the boxes) were placed right at the edge of the table, in the present study.

Besides, performance in the Move-object-, Move-screen- and Hiding-game task was compared to performance in a standard false belief task (Perner, Sprung & Zauner, 2003).

3.1. Method

3.1.1. Participants

Twenty-seven children from one kindergarten and two nursery schools in Salzburg (Austria) participated in this study. The age of these children ranged from 2;5 (years; months) to 5;0 and the mean age was 3;5.

3.1.2. Design and Procedure

Each child was given four tasks: Move-object, Move-screen, Hiding-game and False-belief. The tasks were administered in a full balanced order. Children were tested in two sessions (each session lasted between 15 and 20 minutes), approximately one week apart.

The procedure for the Move-object-, Move-screen- and Hiding-game task was the same as in the previous study, expect that the objects (and the boxes) were placed right at the edge of the table.

The procedure for the *False-Belief task* followed the procedure used by (Perner, Sprung & Zauner, 2003). Two traditional false-belief stories were administered by acting out stories with Playmobile® people toys. For instance, in the book story, Max put his book in one location. In his absence another story character moved the book unexpected to a new location. When Max returns, children were asked the false belief test question, assessing their understanding of Max's erroneous belief:

False-belief question: "Where will Max look first for his book?"

3.2. Results and discussion

Although the present study controlled for false positives in the Move-object- and hiding-game task, the Move-object- and Hiding-game task remained easier than the Move-screen task. Performance was very similar in difficulty for the Move-object- and Hiding-game task. Hence there is no evidence that the difference between the Hiding-game and the Move-screen task in the previous study was due to an artefact. The corresponding percentages correct responses in the four task of this study are presented in Figure 5.

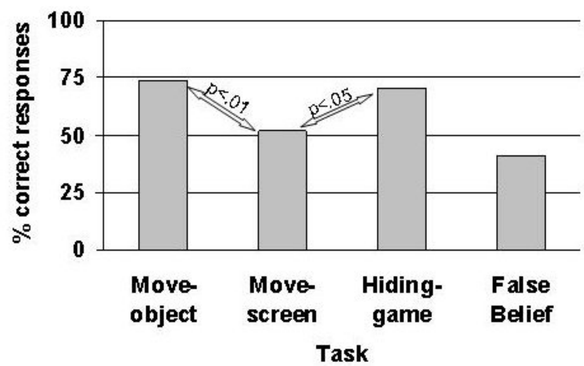


Figure 5. Percent correct response in four tasks of experiment 2

Performance in the Move-object and Hiding-game was significantly higher than in the false-belief tasks (all $p < .01$). However, there was no significant difference between children's performances in the Move-screen- and false-belief task.

4. General Discussion

The results of both experiments replicate the findings by Flavell et al (1978) and McGuigan and Doherty (2002), that children are much better able to occlude objects, from an adult observer, in the Move-object task than in the Move-screen task. Moreover, it has been found that children were also much better able to occlude objects in the new Hiding-game (when the adult observer was not engaged with the object, in its final location), than in the

Move-screen task (when the adult observer was already engaged with the object). This finding supports the hypothesis that although young children (< 2 ½ years) understand engagement, i.e. they are sensitive to adult's general involvement in their activities, but they do not fully comprehend visual perception.

That children's performance in the Move-Screen task was not superior to their performance in the false-belief task, suggests that understanding visual perception (i.e., success the Move-screen task) is not a precursor to understanding theory-of-mind (i.e., success in the false-belief task), but rather that there is a developmental relationship between these two task.

References

[1] Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press/Bradford.

[2] Flavell, J.H., Shipstead, S.G., & Croft, K. (1978). Young children's knowledge about visual perception: Hiding objects from others. *Child Development, 49*, 1208-1211.

[3] Lempers, J.D., Flavell, E.R., & Flavell, J.H. (1977). The development in very young children of tacit knowledge concerning visual perception. *Genetic Psychology Monographs, 95*, 3-53.

[4] McGuigan, N. & Doherty, M.J. (2002). The relation between hiding skill and judgment of eye-direction in preschool children. *Developmental Psychology, 38* (3), 418-427.

[5] O'Neill, D.K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development, 67*, 659-677.

[6] Perner, J., Sprung, M., Zauner, P. (2003). Want That is Understood Well before Say That, Think That, and False belief: A Test of de Villiers's Linguistic Determinism on German-Speaking Children. *Child Development, 74*, 179-188.

Learning gaze following in space: a computational model

Boris Lau[1,2] and Jochen Triesch[2]

[1]Department of Neuroinformatics
Ilmenau Technical University, Germany
boris.lau@stud.tu-ilmenau.de

[2]Department of Cognitive Science
UC San Diego, USA
triesch@cogsci.ucsd.edu

Abstract

*Following another person's gaze in order to achieve joint attention is an important skill in human social interactions. This paper analyzes the gaze following problem and proposes a learning-based computational model for the emergence of gaze following skills in infants. The model acquires advanced gaze following skills by learning associations between caregiver head poses and positions in space, and utilizes depth perception to resolve spatial ambiguities.**

1 Introduction

1.1 Shared attention and gaze following

The capacity for shared attention or joint attention is a cornerstone of social intelligence. It refers to the matching of one's focus of attention with that of another person, which can be established for example by gaze following. The importance of attention sharing in infancy and early childhood is hard to overstate. It plays an important role in the communication between infant and caregiver [8]. It allows infants to learn what is important in their environment, based on the perceived "distribution of attention" of older, more expert individuals. In conjunction with a shared language, it makes children able to communicate about what they perceive and think about, and to construct mental representations of what others perceive and think about. Consequently, episodes of shared attention are crucial for language learning [13].

Some authors make a subtle distinction between joint and shared attention: Joint attention only requires that two individuals attend to the same object, whereas shared attention also implies that each have knowledge of the other individual's attention to this object. In this paper, we will only be concerned with joint visual attention, which has been defined as looking where somebody else is looking, and which we view as an important precursor to the emergence of true shared attention. While initially, joint visual attention is mostly initiated by the caregiver, young infants soon acquire gaze following skills and initiate joint attention themselves [2]. There has been a significant body of research studying how these skills develop since the pioneering work by Scaife and Bruner [10].

Two different kinds of theories of the emergence of gaze following have been proposed. The *modular or nativist theories* posit the existence of innate modules, which are typically thought to be the product of evolution rather than to emerge from learning (e.g. [1]). *Learning based accounts* explain the emergence of gaze following by postulating that infants learn that monitoring their caregiver's direction of gaze allows them to predict where interesting visual events occur. This idea goes back to Corkum & Moore [5]. At present, the experimental evidence for or against a learning account of the emergence of gaze following in infants is still inconclusive, but computational models have shown that it is possible to acquire gaze following skills through learning (see Sect. 2).

1.2 Developmental stages in gaze following

Different distinguishable stages and effects during the development of gaze following have been discovered in cross-sectional studies: Butterworth and Jarret tested gaze following abilities of 6-, 12- and 18-month-old infants in a controlled environment [3]. In their experiments the infants were seated facing their mothers at eye level in an undistracting laboratory. Two or four targets of identical shape and color were presented at the same time as pairs on opposite sides of the room, also at the infants' eye level. Mother and infant were facing each other in every trial, until the mother shifted her gaze to a designated target. The infants' reactions were monitored and analyzed. Figure 1 (left) shows a typical setup of the experiments. All tested infants could shift their gaze to the correct direction and were able to locate targets presented within their field of view. However, only the 18-month-old infants followed

*An earlier version of this paper has been presented at the workshop SOAVE2004 (**S**elf-**o**rganization of **a**dapti**ve** behavior), Ilmenau, Germany.

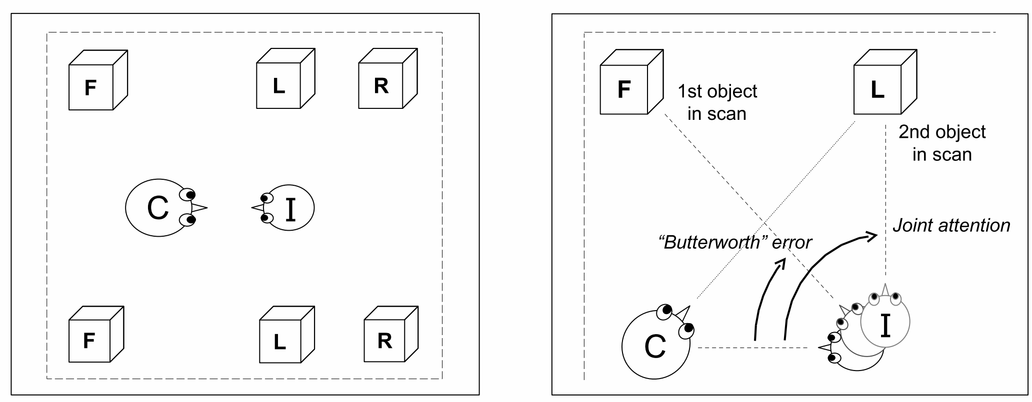


Figure 1. Left: Gaze following experiment with frontal (F), lateral (L) and rear (R) objects. Caregiver (C) and infant (I) are facing each other. Right: The caregiver looks at the lateral target. Six-month-old infants shift their gaze in the correct direction, but will most likely attend to the first object along their scan path (Butterworth error). 18-month-olds follow gaze to the correct lateral object, second in their scan path.

gaze to rear targets, while younger infants would not turn to search for targets behind them. When multiple target pairs were presented at the same time, for example the frontal and lateral targets in Fig. 1, 6-month-old infants were not able to tell which target their mother was looking at: when the mother turned to look at a lateral object, they shifted their gaze in the correct direction, but were likely to end the gaze shift at the first (frontal) object along their scan path, as shown in Fig. 1 (right). We call this effect the "Butterworth error". The infants in the 12 month group attended significantly more often to the correct object, but only the 18-month-old infants reliably followed their mother's gaze to the second (lateral) target.

Butterworth and Jarret associate a developmental stage with each of the age groups: Infants in the "ecological stage" around 6 months follow gaze in the right direction but locate only frontal targets correctly, and only if they are the first along the scan path. 12-month-old infants in the "geometric stage" are able locate the target objects more accurately and overcome the Butterworth error in some of the trials. Infants that have reached the "representational stage" around 18 months reliably overcome the Butterworth error and are also able to reliably locate targets behind them. The emergence of those stages is explained with three different mechanisms of gaze following that become effective in a sequential order and correspond to the observed stages [3].

1.3 Contribution of this paper

In order to explain the emergence of gaze following one has to explain the underlying dynamical processes of development, rather than just the snapshots provided by cross-sectional studies. In the remainder of this paper we will analyze the gaze following problem more carefully with an emphasis on its spatial properties, and isolate the different effects observed in the experimental studies. We propose a computational model, in which the infant acquires sophisticated gaze following skills and is able to overcome the Butterworth error by utilizing depth perception. It shows that the observed behaviors can emerge from the same learning mechanism and thus provides a more parsimonious account for the emergence of gaze following than the three different mechanisms proposed by Butterworth and Jarrett.

2 Gaze following and computational models

Following somebody's gaze in order to establish joint attention is a non-trivial task in cluttered environments. By observing someone's head pose, one can only infer the person's direction of gaze, rather than the distinct focus of the person's attention. Gaze following therefore requires scanning for an object along an estimate of a person's line of sight. For a precise estimate, infants have to evaluate the orientation of the caregiver's head and eye, as well as their own relative position to the caregiver. We will use the term 'head pose' in a general sense, referring to both head or eye orientations. The better the infants can discriminate different head poses, the better they can narrow down the region in space where they expect the caregiver's gaze target to be. Accurate depth perception can help to judge if objects are in the estimated line of gaze, and seems to be critical in situations where objects are in the projection of the caregiver's line of gaze but at different distances, as in Butterworth's experiments. There is evidence that infants' perception of some depth cues continues to develop until at least 7 months [14]. This could have an impact on infants' ability to acquire advanced gaze following skills and may be part of an explanation of the staged development of gaze following.

We believe that infants typically learn the ambiguous mapping from caregiver head poses to locations in space without explicit supervision. Our goal is to plausibly explain this learning process by developing computer models that show how these skills can be acquired. In general, computational models have been developed that address different aspects of the gaze following problem. To our knowledge, two of them show how infants can learn gaze following without external task evaluation (no special reward for establishing joint attention) in a self-organizing manner. Both are discussed in the remainder of this section.

Carlson and Triesch recently proposed a computational model for the emergence of gaze following [4]. Their model infant predicts where salient objects are on the basis of the caregiver's head pose. They use a temporal difference (TD) learning approach [11] to show how an infant can develop these skills only driven by visual reward. The infant receives different rewards for looking at the caregiver and looking at salient objects. This reward structure can be adjusted to simulate certain symptoms of developmental disabilities like Autism or Williams Syndrome. Experiments with the model make predictions of the emergence of gaze following in children with those disabilities. Further experiments with this model were conducted by Teuscher and Triesch [12], focusing on the effect of different caregiver behaviors on infants' gaze following skills.

The model operates on a finite set of possible object locations without any spatial relationships. Each location has a one-to-one correspondence with a distinct caregiver head pose. One object is located at any time at any one of these positions. The caregiver agent has a certain probability of looking at that object. The model infant consists of two reinforcement learning agents: The 'when-agent' decides whether to continue fixating on the same location or to shift gaze, while the 'where-agent' determines the target of each gaze shift. Both agents try to maximize the long term reward obtained by the infant. The infant perceives the caregiver's head pose whenever it attends to the caregiver, and learns to exploit the correlation between the head pose and the location of salient objects. This model supports the theory of the acquisition of gaze following by learning. However, it is not adequate for simulating or explaining the Butterworth stages since it does not deal with geometrical relationships and spatial ambiguities.

Nagai et al. proposed a model for an infant agent that has been implemented on a robot platform [9]. The robot learns to follow the gaze of a human caregiver by offline training with recorded examples. Two separate modules, one for visual attention and one for learning and evaluation, output motor commands for turning the robot's camera head. A probabilistic gate module decides which of the two proposed motor commands gets executed. The probability for selecting the output of the learning module is changed from zero to one according to a predefined sigmoid function during the learning process. The visual attention module locates faces and salient objects by extracting color, edge, motion, and face features from the camera images. It uses a visual feedback controller to shift the robot's attention towards interesting objects. The learning module consists of a three-layered neural network that learns a mapping from gray-level face images to motor commands by backpropagation. The network is trained with the current motor position as teacher signal and the caregiver image as input, whenever a salient object is fixated.

The authors mention that every head pose only specifies a line of gaze rather than a distinct location in space. They deal with this ambiguity by moving the cameras incrementally towards the learned coordinates and stopping the movement at the first encountered object. Their model does not include depth perception and cannot resolve situations where distracting objects lie in the projection of the caregiver's line of gaze in the camera images, but at a different distance (compare Fig. 1, right). The model is not able to overcome the Butterworth error, which seems to be an essential characteristic of geometrical gaze following skills in infants.

3 A model of gaze following in space

Our new model specifically addresses the spatial ambiguities in the learning process of gaze following, and is able to faithfully reproduce infants' abilities to resolve them. It consists of a simulated environment and two different agents, an infant (Inf) and its caregiver (CG). The infant learns to follow the caregiver's gaze by establishing associations between the caregiver's head pose and positions in space where interesting objects or events are likely to be present. This online learning mechanism is driven by visual feedback, based on the infant's preference to look at the caregiver's face and salient objects in its environment. The infant exploits the correlation between the caregiver's line of gaze and the locations of salient objects to learn associations between those two. The perceptual preferences and the ability to shift gaze to interesting objects are important prerequisites for the learning process, which we assume to begin before infants show simple gaze following behaviour (i.e. before an age of six months).

The environment is similar to the setups in the experiments by Butterworth and Jarrett [3], with both agents' eyes and all objects being at the same height from the floor. The learning process is divided into trials. Objects are placed at

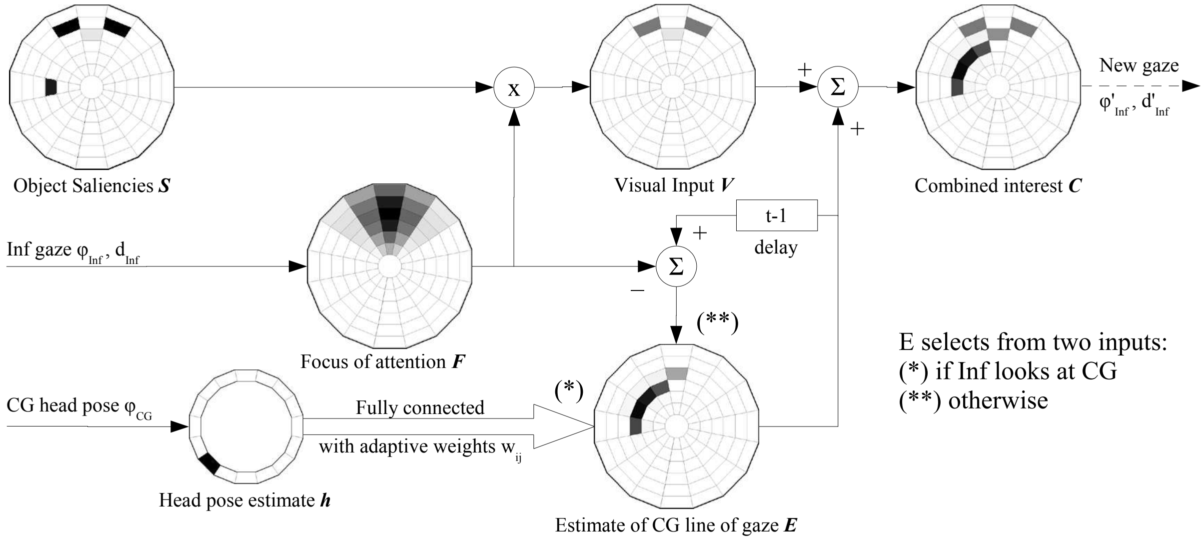


Figure 2. The infant agent with spatial representations in body-centered coordinate systems. Dark shading in the grid cells stands for high activation. The visual input V is the product of the object saliencies S and the focus of attention F. If the infant looks at the caregiver, the estimated caregiver head pose h is mapped to an estimate of the caregiver's line of gaze E. Otherwise the activation in E is held, inhibited by F to decrease the activation of locations along the line of gaze that the infant has already observed. E and V are summed up to the infant's combined interest in space C. The infant shifts its gaze to the area with the highest activation in C.

random positions in the environment in every trial. One of them is selected as the caregiver's focus of attention. The object locations and the caregiver direction of gaze do not change during a trial. The infant is looking at the caregiver at the beginning of every trial but can change its direction of gaze. The model operates on discrete time steps $t = 0, \dots, T$. Each trial lasts for 10 time steps.

3.1 Environment, objects, and a caregiver

The environment is represented by a two-dimensional 7x9 grid with cartesian coordinates. Objects indexed with $i = 1, \dots, N$ are introduced by specifying their grid coordinates (x_i, y_i) and a scalar saliency $s_i \in [0, 1]$. Both agents $a \in \{\text{Inf, CG}\}$ are defined by their positions in space (x_a, y_a), a base orientation φ_a^0 and the current direction of gaze $\varphi_a(t) \in [-180°, +180°]$, relative to φ_a^0. In addition to the current angle of gaze we introduce the function $d_a(t)$, which measures the distance from an agent to the point that the agent is currently looking at. The caregiver also has a saliency $s_{\text{CG}} = 0.1$. All angles and distances are discretized. We use 16 different values for angles (each corresponds to a range of $22.5°$), and 6 different values for distances (covering all possible distances in the 7x9 grid).

Since we focus on the spatial aspects of the learning problem and the infant's ability to learn gaze following without external task evaluation, we use a simple caregiver agent that does not react to the infant's actions. In every learning trial we let the caregiver look at the object i with the highest saliency s_i by setting its head/eye rotation $\varphi_{\text{CG}}(t)$ to the appropriate value.

3.2 The infant agent

The infant has to use its limited visual perception to gain information about the environment. The architecture of the infant agent is shown in Figure 2. It consists of different layers of neurons: the visual input V, the estimate of the CG line of gaze E, the combined interest C and the encoded caregiver head pose h. Their activations are represented with scalar values, assigned to the grid cells of a body-centered polar coordinate grid with discretized angle θ and radius r. The connections between those layers link only neurons encoding the same area in space. The object saliencies S and the encoded focus of attention F are also represented in body-centered coordinates. The infant's interest in the different locations in space is encoded by the combined interest layer $C(\theta, r, t)$. The activation of C is the sum of the visual input V and the estimate of the caregiver's line of gaze E:

$$C(\theta, r, t) := V(\theta, r, t) + E(\theta, r, t). \quad (1)$$

The infant shifts its gaze in every time step t to the area in space it is most interested in. This is done by setting its gaze orientation $\varphi_{\text{Inf}}(t)$ and looking distance $d(t)$ to the coordinates θ and r with the highest activation in $C(\theta, r, t)$.

Visual Perception is the infant's only source of information about its environment. It receives two different kinds of visual data: The caregiver head pose, encoded in the layer $h(\theta, t)$, and the actual visual input $V(\theta, r, t)$, which is the foveated transformation of the object's saliencies into the discretized polar coordinate system. V is used as a gate in the learning mechanism.

Generally we use discrete gaussian distributions $G_\sigma(x)$ as tuning curves for encoding input data for the infant agent. Extra normalization is necessary to ensure that the sum of the discrete distributions over all integers z is equal to one:

$$\tilde{G}_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{2}$$

$$G_\sigma(x) = \tilde{G}_\sigma(x) / \sum_{z \in \mathcal{Z}} \tilde{G}_\sigma(z). \tag{3}$$

The caregiver's head pose $\tilde{\varphi}_{\text{CG}}$ is encoded with a population of neurons h with gaussian tuning curves. The variance ${\sigma_h}^2$ models the level of accuracy in head pose discrimination:

$$h(\theta, t) := G_{\sigma_h}\left(\varphi_{\text{CG}}(t) - \theta\right). \tag{4}$$

The locations (x, y) of the salient objects and caregiver are expressed in the infant's body-centered polar coordinates (θ', r'). The saliency value for each grid cell in S is the sum of all saliencies $s_k,\ k \in \{1, ..N, \text{CG}\}$ falling into the particular area of space. The infant's accuracy in depth perception is modeled with the variance ${\sigma_d}^2$ of the tuning curve encoding the distance of the objects:

$$S(\theta, r, t) := \sum_{k \mid \theta'_k = \theta} s_k(t) \cdot G_{\sigma_d}\left(r'_k(t) - r\right). \tag{5}$$

The infant's visual input V is the product of the object saliencies S and the focus of attention F, which is encoded in the same body centered coordinate system as the neural layers. It is a product of two gaussians (not normalized). It has its highest value at the focused point in the center of gaze $\theta = \varphi_{\text{Inf}}(t),\ r = d_{\text{Inf}}(t)$ and values close to zero for angles and distances further away from the infant's current focus of attention. This causes V to be a foveated view. The variances ${\sigma_\theta}^2$ and ${\sigma_r}^2$ influence the sharpness of the foveation:

$$V(\theta, r, t) := S(\theta, r, t) \cdot F(\theta, r, t) \tag{6}$$

$$F(\theta, r, t) := e^{-\frac{(\theta - \varphi_{\text{Inf}}(t))^2}{2{\sigma_r}^2}} \cdot e^{-\frac{(r - d_{\text{Inf}}(t))^2}{2{\sigma_\theta}^2}}. \tag{7}$$

Our model acquires gaze following skills by learning associations between the caregiver's head pose h and locations in space, forming the estimate of the caregiver's line of gaze $E(\theta, r, t)$. The associations are represented as connections with variable weights. We use a Hebbian-like learning rule that strengthens all connections from each active input neuron encoding a specific caregiver head pose to those locations where the infant saw a salient object shortly after observing the same head pose (activation in V). A small learning rate $\alpha_{\text{Hebb}} = 0.1$ combined with a slow decay of all synaptic weights, given by $\alpha_{\text{forget}} = 0.9999$, enables the network to 'forget' wrong associations that could be established when multiple objects are present during the training. The synaptic weight between a neuron j with activation $h(\omega, t)$ and a neuron i with activation $E(\theta, r)$ is given by $w_{ij}(t)$ and adapted with the following learning rule:

$$w_{ij}(t+1) := \alpha_{\text{forget}} \cdot w_{ij}(t) + \alpha_{\text{Hebb}} \cdot h(\omega, t) \cdot V(\theta, r, t). \tag{8}$$

The activation associated with the head pose encoded in h overwrites the activity in E whenever the infant is looking at the caregiver. When the infant has shifted its gaze away from the caregiver, E keeps its activation and is used as a short-term memory: the activation of the neurons encoding areas in space that the infant has already observed is suppressed by the activations of the neurons in F, encoding the focus of attention:

$$E(\theta, r, t) := \begin{cases} \sum_j \{w_{ij}(t) \cdot h(\omega, t)\}, & \text{if Inf looks at CG,} \\ E(\theta, r, t-1) \cdot (1 - F(\theta, r, t)) & \text{otherwise.} \end{cases}$$

The selective inhibition of activity in E causes the infant to shift its gaze to unobserved locations, because it always attends to the area with the highest activation in C. This "scanning" continues as long as the activation along the line of gaze is higher than the activation due to the foveated visual input. It usually ends when the infant looks directly at an object.

4 Experiments

We present a number of experiments to show that our model infant is able to acquire gaze following skills and learns to overcome the Butterworth error. Each experiment is run 20 times under the same conditions for 1000 learning trials. The performance is measured in testing periods interposed every 25 trials during which no learning takes place. Every testing period consists of several trials with 10 time steps each, one trial for every tested object location. A trial is considered successful when the infant is looking where the caregiver is looking at the last time step of the trial. The performance of the model is measured with the t]

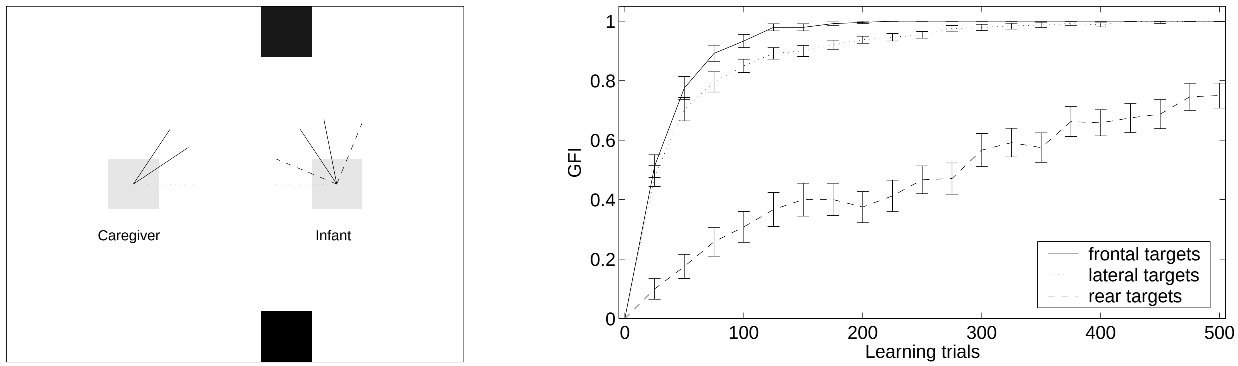


Figure 3. Gaze following performance for frontal, lateral and rear targets. *Left:* **Geometrical setup and situation at the end of a successful trial. The individual directions of the gaze of infant and caregiver are displayed with pairs of solid lines. The dotted lines indicate the agent's base orientation. The dashed lines display the borders of the infant's field of view.** *Right:* **Gaze Following Index for frontal, lateral and rear target pairs as functions of learning trials. The infant quickly learns to follow gaze to frontal and lateral targets. Gaze following to rear targets is acquired slowly. Data points are averaged from 20 runs, the error bars indicate the standard error.**

Gaze Following Index (GFI), which is defined as the number of successful trials divided by the total number of trials.

4.1 Gaze following performance

This experiment is designed to measure the model infant's gaze following performance separately for frontal, lateral and rear targets. We therefore split the testing trials in three groups, depending on the position of the caregiver's target object relative to the infant: a trial is considered a front target trial, when the caregiver's target is in the infants field of view while watching the caregiver. When the target object is initially out of view but not behind the infant, this is considered a lateral target trial. All other conditions are rear target trials.

Even the untrained model infant is able to locate frontal targets and to attend to them by simply using its peripheral vision. In order to eliminate this influence of simple preferential looking on the gaze following performance we present pairs of targets with a small difference in their saliency during the testing trials. Different from the learning trials we constrain the caregiver to look at the slightly less salient object in the testing trials, just by setting its head/eye rotation $\varphi_{\text{CG}}(t)$ to the appropriate value. The infant will turn to the other, more salient object unless it follows the caregiver's gaze.

All individual target positions in space are tested, except the line connecting infant and caregiver. The setup is shown in Fig. 3 (left). We use tuning curves with small variances for encoding the caregiver head pose and the infant's perception of distances ($\sigma_h = \sigma_d = 0.1$) in order to test the gaze following performance independent from limitations in depth perception or face processing.

The result of this experiment is displayed in Fig. 3 (right). The infant learns to reliably follow the caregiver's gaze to frontal objects in about 100 learning trials, to lateral objects in about 200 learning trials, and to rear targets (with a little lower GFI) in about 500 trials. This corresponds to the results of the experiments by Butterworth and Jarret, where only the infants in the oldest age group shifted their gaze to rear targets.

The infant has not necessarily learned the complete set of associations for the frontal targets and every caregiver head pose until trial number 100. In fact, turning the head in the correct direction moves the target object closer to the infant's focus of attention and the other one further away. This can cause a higher activation in the foveated visual perception for the correct object than for the originally more salient distractor. In this case the infant will attend to the correct object. This corresponds to the ecological stage in the development in real infants.

A similar effect is exploited when the infant learns associations between a head pose and rear objects, outside the infant's field of view: Turning in the correct direction brings lateral targets into the infant's field of view and enables the infant to learn the corresponding associations. Learning to follow the caregiver's gaze to objects that are behind the infant requires a prior ability to follow gaze to lateral targets. This explains why it takes longer for the infant to achieve reliable gaze following skills for rear targets as seen in real infants.

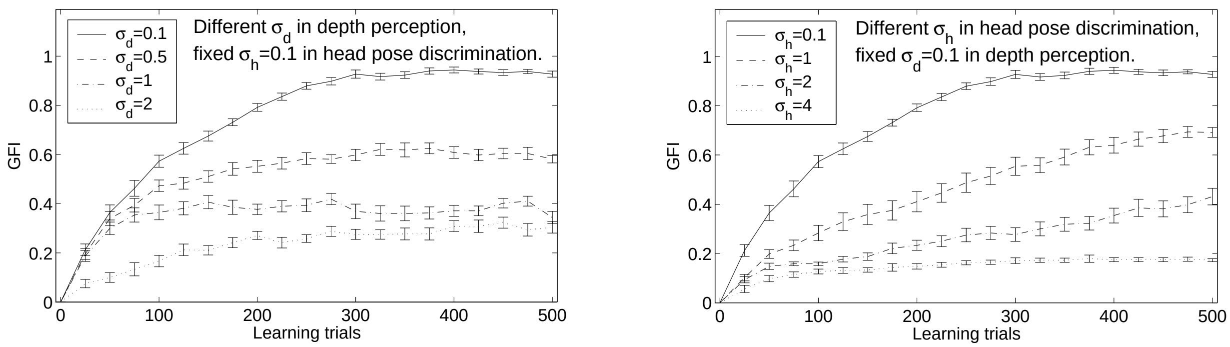


Figure 4. Overcoming the Butterworth error. Gaze Following Index for trials with lateral targets and frontal distractor objects, tested with different levels of accuracy in depth perception and head pose discrimination. High accuracy corresponds to using low variances for the tuning curves encoding object distances and caregiver head pose. Data points are averaged from 20 runs, the error bars indicate the standard error.

4.2 Overcoming the Butterworth error

In this experiment we test the infant's gaze following performance in the presence of distractor objects. Two salient distractors are placed as a pair of frontal targets behind the caregiver like shown in Fig. 1 (left). The slightly less salient target object, which the caregiver is attending to, is placed at different lateral and frontal locations, but not behind the infant or the caregiver. We test the gaze following performance with different settings for the infants ability to discriminate distances and head poses, by varying the variances ${\sigma_h}^2$ and ${\sigma_d}^2$ for the tuning curves encoding the head pose and the distances of the objects.

The results of this experiment are displayed in Fig. 4. The infant is able to overcome the Butterworth error and to ignore the distractor objects in the background for the majority of target positions, if depth perception and the discrimination of head poses are sufficiently accurate ($\sigma_h = \sigma_d = 0.1$). A higher variance (less accuracy) for depth perception or head pose discrimination leads to significantly worse gaze following performance. Unlike our model infant we assume real infants to improve their skills of depth perception and face processing over time. Our experimental results suggest that an infant cannot acquire geometrical gaze following skills before its depth perception and face processing skills are sufficiently developed. It is important to note that those skills seem to be critical not only for the actual gaze following, but for the acquisition as well.

Our model needs more than 200 learning trials to achieve reliable gaze following performance in the presence of distractors, compared to 100 trials in a simple setup with only one pair of objects. In both cases the model used high accuracy in depth perception and face processing from the first learning trial on. With only gradually developing depth perception skills the model would overcome the Butterworth error even later. These results correspond to the results of Butterworth where only older children are able to follow their caregiver's gaze correctly in ambiguous situations.

5 Discussion

We have analyzed the gaze following problem with an emphasis on its spatial characteristics, and presented a new model for the emergence of gaze following. The infant in our model learns to follow the caregiver's gaze by learning associations between observed head poses and positions in space. These associations form an ambiguous mapping from every head pose to several locations where salient objects are likely to be present. We demonstrated in experiments that our model is able to reach all stages of gaze following: first it is able to resolve spatial ambiguities when distractor objects are present in the background by using depth perception, and second it follows the caregiver's gaze to locations even behind its back. Furthermore, the temporal progression of the different stages is similar to the development observed in real infants: gaze following to frontal targets early in the development, overcoming the butterworth error and finding lateral targets later, and locating rear targets even later.

The model also makes predictions about the effect of limitations in depth perception and face processing on infants' ability to gain advanced gaze following skills: The better an infant can discriminate different head poses and object distances, the smaller is the region in space that will be associated with each head pose. If one of these two skills is not sufficiently developed, the model cannot overcome the Butterworth error. This suggests that children who are late to acquire accurate face processing and depth perception may develop geometric gaze following skills later than their peers.

Butterworth and Jarrett proposed that the development of a representation of space that contains infant, caregiver, and objects corresponds to the infants' ability to follow gaze to rear targets. The body-centered coordinate systems that we use in the infant agent provide such a spatial representation. The results of our first experiment show that gaze following to rear targets might occur later, even with such a representation of space already in place.

Our model, like most models, makes many abstractions and simplifications. While focusing on the spatial problems of gaze following we especially simplified the dynamic aspects in this problem by running the simulation in discrete trials. Different problems occur with a continuous time line in a dynamic environment: The longer the infant turns away from the caregiver, the more likely it is that the caregiver has already shifted its gaze again, causing a growing uncertainty in the infant's estimate of the caregiver head pose.

Popular approaches from the research areas of active vision and machine learning could be applied to the gaze following problem. One can understand the infant's search for salient targets as a state estimation process, based on limited observations of the real state, which is the actual distribution of salient objects in the room. Research on Partially Observable Markov Decision Processes (POMDPs) deals with the problem of decision making in environments with hidden states (e.g. [7]). Denzler and Brown developed an information theoretical approach to optimal sensor parameter selection in object recognition [6]. A similar approach could be used in the infant agent to learn how to efficiently integrate information from the available sources, namely accurate but visual perception with a limited field of view and ambiguous information from evaluating the caregiver's head pose.

Acknowledgments

The work described in this paper is part of the MESA project (**M**odeling the **E**mergence of **S**hared **A**ttention) at UC San Diego, a larger effort to understand the emergence of shared attention in normal and abnormal development supported by the National Alliance for Autism Research. We especially thank Christof Teuscher and Gedeon Deák for fruitful discussions, and Alan Robinson and Erik Murphy-Chutorian for comments on the draft. B. Lau is supported by the German National Merit Foundation.

References

[1] S. Baron-Cohen. *Mindblindness: an essay on autism and theory of mind*. A Bradford Book, The MIT Press, 1995.

[2] G. E. Butterworth. The ontogeny and phylogeny of joint visual attention. In A. Whiten, editor, *Natural theories of mind: Evolution, development, and simulation of everyday mindreading*, pages 223–232. Blackwell, 1991.

[3] G. E. Butterworth and N. Jarrett. What minds have in common in space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9:55–72, 1991.

[4] E. Carlson and J. Triesch. A computational model of the emergence of gaze following. In H. Bowman and C. Labiouse, editors, *Connectionist Models of Cognition and Perception II*. World Scientific, 2003.

[5] V. Corkum and C. Moore. Development of joint visual attention in infants. In C. Moore and P. J. Dunham, editors, *Joint attention: Its origins and role in development*, pages 61–83. Erlbaum, 1995.

[6] J. Denzler and C. Brown. Information theoretic sensor data selection for active object recognition and state estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):145–157, 2002.

[7] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101:99–134, 1998.

[8] C. Moore and P. J. Dunham, editors. *Joint attention: Its origins and role in development*. Erlbaum, 1995.

[9] Y. Nagai, K. Hosoda, A. Morita, and M. Asada. A constructive model for the development of joint attention. *Connection Science*, 15(4):211–229, December 2003.

[10] M. Scaife and J. S. Bruner. The capacity for joint visual attention in the infant. *Nature*, 253:265ff, 1975.

[11] R. S. Sutton and A. G. Barto. *Reinforcement Learning: an introduction*. A Bradford Book, The MIT Press, 1998.

[12] C. Teuscher and J. Triesch. To care or not to care: Analyzing the caregiver in a computational gaze following framework. *3rd International Conference for Development and Learning, ICDL'04, La Jolla, California, USA*, 2004.

[13] M. Tomasello. *The cultural origins of human cognition*. Harvard Univ. Press, 1999.

[14] A. Yonas, C. A. Elieff, and M. E. Arterberry. Emergence of sensitivity to pictorial depth cues: Charting development in individual infants. *Infant Behaviour & Development*, 25:495–514, 2002.

Motion Recognition and Generation for Humanoid based on Visual-Somatic Field Mapping

[1]Masaki Ogino, [1]Shigeo Matsuyama, [1]Jun'ichiro Ooga, and [1,2]Minoru Asada
[1]Dept. of Adaptive Machine Systems, [2]HANDAI Frontier Research Center,
Graduate School of Engineering, Osaka University
2-1, Yamada-Oka, Suita, Osaka, Japan
{ogino, shigeo, ooga}@er.ams.eng.osaka-u.ac.jp, asada@ams.eng.osaka-u.ac.jp

Abstract

This paper presents a method of imitation learning based on visuo-somatic mapping from observing the demonstrator's posture to reminding the self posture via mapping from the self motion observation to the self posture for both motion understanding and generation. First, various kinds of posture data of the observer are mapped onto posture space by self organizing mapping (hereafter, SOM), and the trajectories in the posture space are mapped onto a motion segment space by SOM again for data reduction. Second, optical flows caused by the demonstrator's motions or the self motions are mapped onto a flow segment space where parameterized flow data are connected with the corresponding motion segments in the motion segment space. The connection with the self motion is straightforward, and is easily acquired by Hebbian Learning. Then, the connection with the demonstrator's motion is automatic based on the learned connection. Finally, the visuo-somatic mapping is completed when the posture space (the observer: self) and image space (the demonstrator: other) are connected, which means observing the demonstrator's posture associcates the self posture. Experimental results with human motion data are shown and the discussion is given with future issues.

1 Introduction

Humanoid robot is expected to have behaviors like human as it is supposed from its appearance. The motion programming for such a robot with multiple joints is a hard task, therefore, imitation is one of the plausible solutions for humanoid motion programming [9]. This attempt has already achieved success to some extent in real robots. Nakazawa et al. [5] have realized a dancing humanoid robot that can imitate human dance performances. They segment human dancing motion into typical motion primitives with parameters. Ijspeert et al. [2] have focused on dynamical aspects of imitation and proposes the methods to describe the observed motion using the basic non-linear dynamics primitives.

On the other hand, imitation is also supposed to a fundamental framework for motion recognition in biological system. Billard and Mataric [1] emphasized importance of motion primitives, and constructed the motion control system based on the CPG modules and the learning modules. Inamura et al. [7] proposed a system that describes the self and the demonstrator's motions in the same mimesis loop, in which motions are recognized and generated in the hidden Markov models.

However, almost existing approaches to imitation in robotics assume that the angles of others' links are available. The somatosensory signals or motion commands of others are not accessible and it is necessary to have a mechanism that converts visual information observing others to self motion. Recently, Kuniyoshi et al. [4] proposed a learning system for early imitation. They suppose the optical flow information is the key to induce the self motion corresponding to the observed motion. However, they didn't mention how the early imitation can be extended to the higher level of learning.

This paper presents a method of imitation learning based on visuo-somatic mapping from observing the demonstrator's posture to reminding the self posture via mapping from the self motion observation for both motion understanding and generation. First, various kinds of posture data of the observer are mapped onto a *posture space* by self organizing mapping [3] (hereafter, SOM), and the trajectories in the posture space are mapped onto a *motion segment space* by SOM again for data reduction. Second, optical flows caused by the demonstrator's motions or the self motions are mapped onto a *flow segment space* where parameterized flow data are connected with the corresponding motion seg-

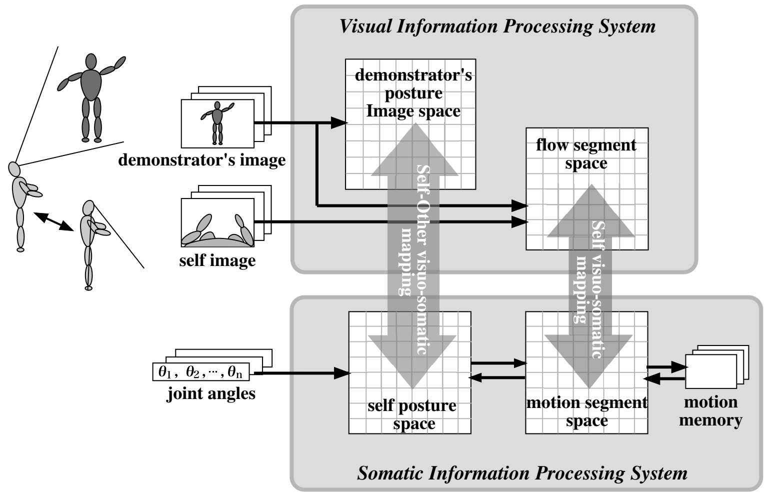


Figure 1. System overview

ments in the motion segment space. The connection with the self motion is straightforward, and is easily acquired by Hebbian Learning. Then, the connection with the demonstrator's motion is automatic based on the learned connection. Finally, the visuo-somatic mapping is completed when the posture space (the observer: self) and image space (the demonstrator: other) are connected, which means observing the demonstrator's posture associates the self posture like a mirror system [10]. Experimental results with human motion data are shown and the discussion is given with future issues.

2 A System Overview

2.1 Basic assumptions

Here, we assume the followings to realize the visual imitation based on the visuo-somatic mapping:

1. No *a priori* knowledge on the link structure, that is, connections between joints.
2. No *a priori* knowledge on the body part (joint) correspondence between the demonstrator and the observer.
3. Both the demonstrator's and the self motions can be observed in terms of a temporal sequence of joint vectors.
4. The joint angles of the self posture can be observed, but no relationship between the self posture and the flow segment space is given.
5. Currently, we focus on the mirror image imitation. This is the right (left) side of the demonstration corresponding to the left (right) side side of the observation.

2.2 Imitation system

Fig. 1 shows the proposed system, consisting of two sub-processing systems, *Visual Information processing system* and *Somatic Information processing system*. In these processing sub-systems, row sensory data are mapped onto the corresponding two dimensional Self-Organizing Maps (SOMs) [3]. The images observing the demonstrator's motion are first mapped onto *image space* which includes the posture image of the demonstrator, and then *flow segment space* in which the changes in posture are represented. The visual feature space is also utilized to represent the self motion, too. On the other hand, the self somatic sensory data are mapped onto *posture space* and their changes are mapped onto *motion segment space*. After generation of these maps independently, the flow segment space and the motion segment space are connected based on Hebbian learning.

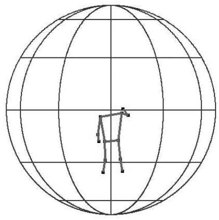
The connection between the visual feature space and the motion segment space is easily carried out by using Hebbian learning based on the simultaneous activations of segments in both spaces during the self motions. Once this connection is acquired, the connection between the flow segment space for the demonstrator's motion and the motion segment space is automatic based on the learned connection between the visual feature space and the motion segment space. Through these connections, the mapping from the image space of the demonstrator's posture to the self posture space is enabled, that is, visuo-somatic mapping can be obtained.

In the followings, the details of each sub processing system are explained in section 3, and the mapping among them is shown in section 4.

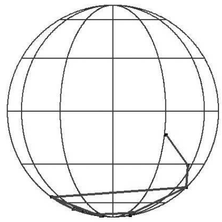
2.3 Sensory data

We prepare the sensory data by using human motions acquired by a motion capturing system. The captured data, which are three dimensional data sets in the global coordinate system, are converted to the two dimensional data on a virtual camera images captured by the observer (self). The angles between links in a human model are also calculated to be used as the self posture data. Figs. **??**(a) and **??**(b) show the attached place of labels as joints to be captured and a sample of captured data of the demonstrator's motion, respectively. A joint angle vector is mapped onto the self posture space and the segmented trajectories on the map are mapped on motion segment space. The spherical image projection from the camera position at the observer head is assumed to capture the whole self body image. Fig. 3 shows examples of the self body image (a) and the demonstrator's one (b) on the spheres, and their development onto a plane (c).

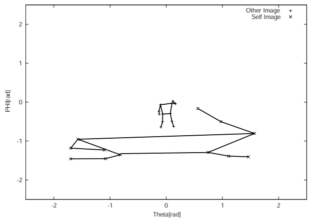
Twelve kinds of motions are captured from the human motion performances. They are combinations of motion, side, and part such as "raise," "wave," and "rotate" as motions, "left," "right," or "both" as sides, and "hand," and "knee" as parts. Also a "walking" motion is added as a whole body motion.



(a) Image of a demonstrator

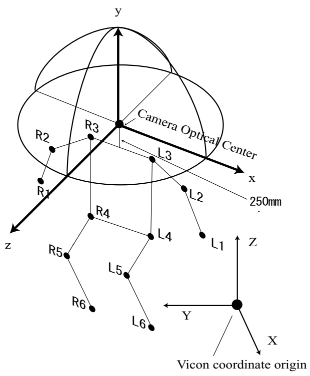


(b) Image of self body

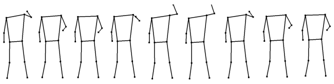


(c) The development of the spherical projection

Figure 3. Image data to be input to the system



(a) Position of the camera optical center



(b) An example of captured motions

Figure 2. Capturing data

3 Construction of SOMs for behavior recognition and generation

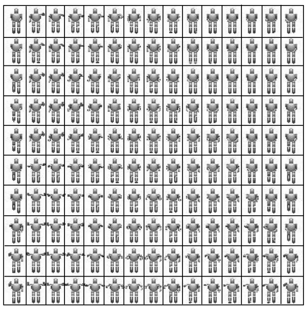
3.1 Posture space and motion segment space

We construct a *posture space* SOM from the *somatic sensation information*, which is the sequence of the vectors consisting of sixteen angles calculated from the captured motion data. The size of *posture space* SOM is 15×15, and it is constructed by 240 [frames] (8 [sec]) per each motion. Fig. 4 (a) shows the resultant SOM.

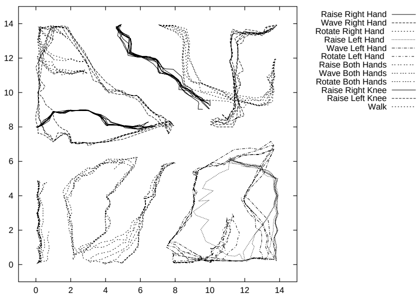
Since the posture data are input sequentially, we can visualize how posture data are connected each other in the posture space. Fig. 4 (b) shows such data indicating that the trajectries of motions are roughly segmented and construct the clusters corresponding to performed actions. These trajectories are divided into small segments, each of which includes 10 [frames] of trajectories on the posture space SOM, and are clustered in the upper layer SOM, motion segment space, (Fig. 4 (c)).

3.2 Demonstrator's posture image space

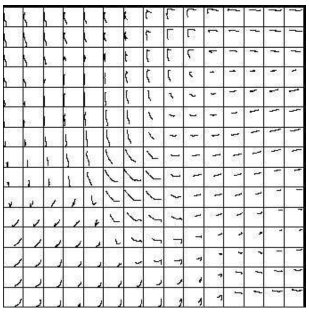
A demonstrator's posture image space (hereafter, image space in short) consists of the representative image position vectors obtained by self organizing mapping of image po-



(a) Posture space



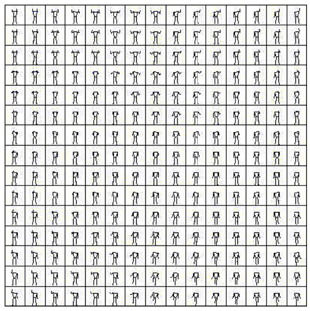
(b) Trajectories of motions on posture space



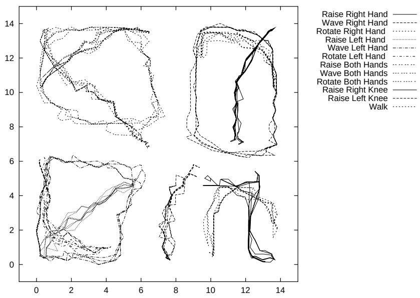
(c) Motion Segment Space

Figure 4. The soms in self somatic sensation information processing system

sitions of joints of the human model in Fig. **??**. Fig. 5 (a) shows the image space where various postures are clustered into 15 $\times$ 15 representative postures. Similar to the posture space based on the somatic information, we can visualize how posture image data are connected each other in this space. Fig. 5 (b) shows such data which indicate that the trajectories of motions are roughly segmented and construct the clusters corresponding to performed actions.



(a) Demonstrator's posture image space



(b) Trajectories of motions on demonstrators posture image space

Figure 5. A demonstrator's posture image space

3.3 The flow segment space

The problem here is how to associate the observed flow caused by the demonstrator's motion with the self motion. If the flows by the demonstrator's motions are similar to the flows by the self motions, the desired association seems easy to find because the connection between the observed self motion and the self motion segment can be easily found based on the simultaneous activations during the self motion. However, it would not be so due to the viewpoint difference. Then, as the common features of the flow segments, we chose the direction and the relative position of the flow segments. Fig. 6 indicates the directions of flow segments by the demonstrator (other) and the observer (self), where we can see that the directions are very similar to each other although there are slight differences in the directional changes between them. The relative positions are quantized into four regions (top left, top right, bottom left, and bottom right) by setting the centroid of posture image vectors as the origin. These four regions are called attention areas. Fig. 7 shows that the positions of joints are similar to each other between the demonstrator's and the observer's in spite of large shape difference.

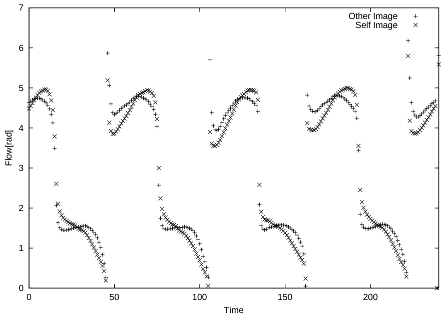
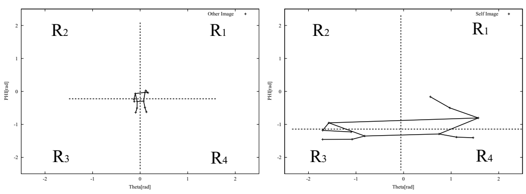


Figure 6. The flow directions from different viewpoints



(a) Attention area of Other Image

(b) Attention area of Self Image

Figure 7. Attention Area

By using the quantized directions and the normalized magnitudes of the flows, and the attention area, the flow segment space is constructed. A data structure for the flow segment space is shown in Fig. 8 (a).

Although Fig. 6 shows the directions of flow vectors in the same parts of self and a demonstrator are almost the same in spite of the camera position, the correspondence between self and a demonstrator's body are unknown. We construct the perceptive field of motion, flow segment space, based on the flow directions and the relative magnitude of flow vectors. The flow segment space has the same number of layers as observed labels (joints), which consists of F_p unit. Each unit has the representative direction correspondent to quantized direction ranging from 0 to 2π (see Fig. 8).

The directions of the flows are segmented when the sign of horizontal or vertical element of flow vector is inverted. In each segment, the directions are averaged. Suppose the time when n-th flow inversion happens is T_n, then the averaged flow direction is given by

$$\S_i^{\phi}(t) = \frac{1}{T_{n+1} - T_n} \int_{T_n}^{T_{n+1}} \phi_i^F(s) ds \quad (T_n < t < T_{n+1}), \qquad (1)$$

where $\phi_i^F(t)$ indicates the flow direction of body segment i at time t. The averaged direction data are sorted by their length of the flow vectors. And the i-th data is assigned to the i-th layer in flow segment space. In each layer, the unit that has the nearest direction to the input data is activated.

3.3.1 Attention area

Although the positions of flow vectors in the robot's view are quite different between the self and the demonstrator, the relative positions among them (upper right, upper left, lower right and lower left) are roughly maintainted well as shown in Fig. 7. Using this feature, the *attentional area* describes what part of the self image includes first flow vectors.

Let N_f the number of observed points of the self and the demonstrator's body and the regions around the center of observed points R_1, R_2, R_3 and R_4 as shown in Fig. 7, then the total flow speed included in each region R_i is given by

$$F_j(t) = \int_{i=1}^{N_f} p_i(t) ||v_i(t)||, \qquad (2)$$

$$p_i(t) = \begin{cases} 1 & if \quad u_i(n) \in R_j \\ 0 & else \end{cases} \quad (j = 1, \cdots 4), \qquad (3)$$

where $v_i(t)$ is the observed flow vector, and $u_i(t)$ is the position vector of observed point i at time t. Note that i does not correspond to the labeled point of the body. We define the relative total strength of flow among regions as

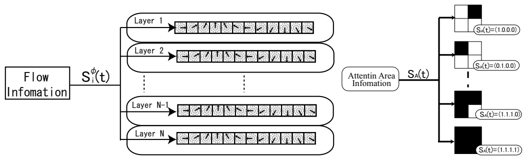
$$A_j(t) = \frac{F_j(t)}{\sum_{n=1}^{4} F_n(t)} \quad (j = 1, \cdots 4). \qquad (4)$$

The input vector to attention area, $S_A(t)$, consists of binarized $A_j(t)$,

$$S_A(t) = (A_1^S(t), A_2^S(t), A_3^S(t), A_4^S(t)) \qquad (5)$$

$$A_j^S(t) = \begin{cases} 1, & if \quad A_j(t) \geq 0.20 \\ 0, & else \end{cases} \qquad (6)$$

Attention area space consists of all the combinations of activated areas, $2^4 = 16$, as shown Fig. 8.



(a) Flow direction

(b) Attention Area

Figure 8. The flow segment space

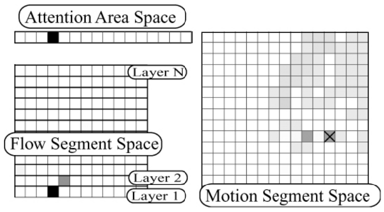


Figure 10. Arrangement of each space

4 Mapping between visual and somatic field

4.1 Self visual-somatic sensation mapping

The simultaneous activations of the units in the flow segment space and the self posture space during self motion make it possible to find correspondence between the units in those spaces (see Fig. 9). The connection coefficients

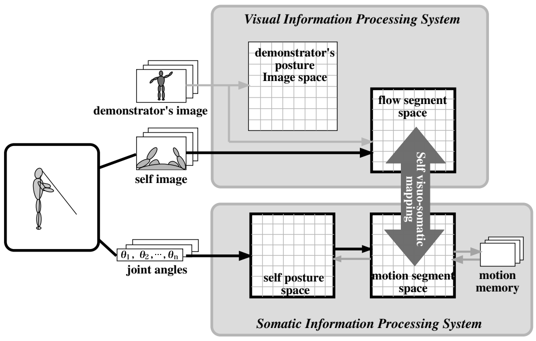


Figure 9. Self visual-somatic sensation mapping

between the units in each space are learned based on Hebbian learning. All the connection coefficients are initialized to 0s, and during the self motion the coefficient, w^{AB}, which is the connection coefficient betwen the i-th unit in space A and the j-th unit in space B, is updated during self motion, as follows,

$$w_{ij}^{AB}(t+1) = w_{ij}^{AB}(t) + \varepsilon(y_i^A(t)y_j^B(t) - y_i^A(t)^2 w_{ij}^{AB}(t)). \quad (7)$$

At the same time, the time sequences of the activated units in motion segment space during various motions are memorized as the motion modules in *motion memory*.

4.2 Recognition of other person's motion

After acquisition of self visual-somatic mapping, the input of image data observing a demonstrator's motion activates the units in the motion segment space through the flow segment space via connections between them (Fig. 11). Let the activation level of the i-th unit in flow direction $y_i^F(t)$ and that of the j-th unit in attention area $y_j^A(t)$, the activation level of the k-th unit in motion segment space, $y_k^M(t)$ is given by

$$y_k^M(t) = \sum_{i=1}^{N_F} w_{ik}^{FM} y_i^F + \sum_{j=1}^{N_A} w_{jk}^{AM} y_j^A \quad (8)$$

The quantization in the flow segment space is coarse and the mapping between the flow segment space and the motion segment space is not one-to-one mapping. The motion of a demonstrator activates multiple units in the motion segment space at a time, which makes it difficult to identify the corresponding motion module. So, we compare the temporal sequences of activated units of observed motion with those of memorized motion modules in the motion segment space. To do that, we define the evaluation function, E_m, which indicates the similarity of the time sequence of acvtiated units of an observed motion to that of m-th memorized motion as follows,

$$E_m = \max_{s_t} \int_0^T \sum_{i=1}^{N_M} y_i(t) m_i(t_s + t) dt. \quad (9)$$

Thus, the observed motion is recognized as the same as the motion module that maximizes E_m.

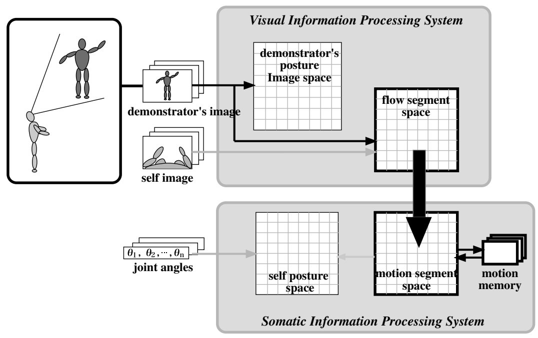


Figure 11. recalling the motion

4.3 Mapping between a self posture image space and a demonstrator's posture image

Recalling the self motion from the observation of a demonstrator's motion makes it possible to correlate the demonstrator's posture image space in visual information processing system with the self posture space in somatic sensation information processing system (Fig. 12). When observing a demonstrator's motion, the unit in the image space and the unit in the posture space activate simultaneously. So we can use Hebbian learning again between these two maps.

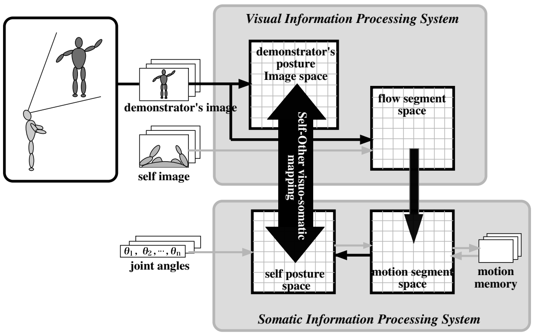


Figure 12. Self-Other Visual-Somatic sensation mapping

Fig. 13 shows the recalled posture (the rightmost figure) from the observed image (the leftmost figure). The two maps in the middle of the figures describe the activated units in image space(left) and that posture space (right) after hebbian learning.

5 Conclusions

In this paper, we proposed a learning system for imitation based on visuo-somatic mapping. This system excludes the pre-designed model of a demonstrator as much as possible. The demonstrator's model is made through demonstrator's images in the demonstrator's posture image space. The model of self is not pre-designed, either. It is constructed by self-organizing the self motion information in self posture space and motion segment space. The primitive visual features are related to the representative vectors in motion segment space during self motion. This connection induces the self motion when observing demonstrator's motions and further mapping between demonstrator's posture image and self posture space is made. After constructing the visuo-somatic mapping, this system can directly activate the self posture corresponding the observed demonstrator's image.

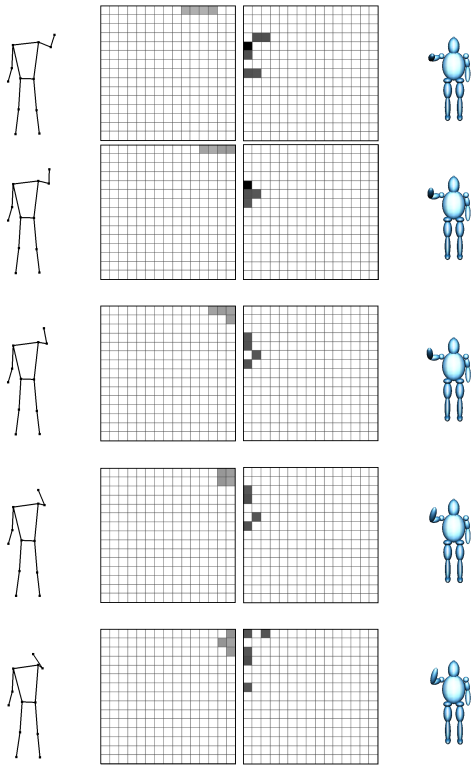


Figure 13. Recalling somatic sensation by Self-Other visual-somatic sensation mapping

Although initial aims to construct the visuo-somatic mapping through learning are accomplished in this system, it has many problems for practical use as an imitation system. First this system assumes that an observer always stands face to face with a demonstrator, and this sytem does not have concept about the translation or rotation to the ground of the demonstrator. An observer can recognize only jestures of the demonstrator. Second, the resultant visuo-somatic mapping is not so accurate as to make new motion modules only from observation of demonstrator's motion, because the sequence of the activated postures is not smooth.

For the first problem, we are now extending our model so that it can describe the transition and rotation of a demonstrator relative to the ground. The second problem can be solved by using velocity information acquired by another pathway. Acquiring new motions which are not experienced through observation is the next challenge for us.

Acknowledgments

This study was performed through the Advanced and Innovational Research program in Life Sciences from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government.

References

[1] A. Billard, M. Mataric: Learning human arm movements by imitation: Evaluation of a biologically inspired connectionist architecture, *Robotics and Autonomous Systems*, 37, pp.145–160, 2001.

[2] A. Ijspeert, J. Nakanishi, S. Schaal, Movement Imitation with Nonlinear Dynamical Systems in Humanoid Robots, In *Proc. of IEEE International Conference on Robotics and Automation*, pp. 1398-1403, 2002.

[3] T. Kohonen: The Self Organization and Associative Memory, Springer-Verlag, 1989.

[4] Y. Kuniyoshi, Y. Yorozu, M. Inaba, H. Inoue: From visuo-motor self learning to early imitation -a neural architecture for humanoid learning-, In *Proc. of IEEE International Conference on Robotics and Automation*, pp. 3132-3139, 2003.

[5] A. Nakazawa, S. Nakaoka, K. Yokoi, T. Harada, and K. Ikeuchi: Imitating human dance motion through motion structure analysis, In *Proc. of 2002 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 2539–2544, 2002.

[6] E. Oja, J. Karhunen: On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix, *J. Math. Anal. and Appl.*, 106, pp. 69–84, 1985.

[7] T. Inamura, I. Toshima, Y. Nakamura: Acquisition and Embodiment of Motion Elements in Closed Mimesis Loop, In *Proc. of the 2001 IEEE International Conference on Robotics and Automation*, pp. 1539–1544, 2002.

[8] R. Pfefer and C. Scheier.: Understanding Intelligence, MIT Press, Cambrridge, Massachusetts, 1999.

[9] S. Schaal., "Is imitation learning the route to humanoid robots?," *Trends in Cognitive Science*, pp. 233–242, 1999.

[10] V. Gallese and A. Goldman., Mirror neurons and the simulation theory of mind-reading, *Trends in Cognitive Science*, Vol. 2, No. 12, pp. 493-501. 1998.

Learning to Recognize and Reproduce Abstract Actions from Proprioception

Karl F. MacDorman*† Rawichote Chalodhorn† Hiroshi Ishiguro†
*Frontier Research Center
†Department of Adaptive Machine Systems
Graduate School of Engineering
Osaka University
Suita, Osaka 565-0871 Japan
kfm@ed.ams.eng.osaka-u.ac.jp

Abstract

We explore two controversial hypotheses through robotic implementation: (1) *Processes involved in recognition and response are tightly coupled both in their operation and epigenesis; and* (2) *processes involved in symbol emergence should respect the integrity of recognition and response while exploiting the fundamental periodicity of biological motion. To that end, this paper proposes a method of recognizing and generating motion patterns based on nonlinear principal component neural networks that are constrained to model both periodic and transitional movements. The method is evaluated by an examination of its ability to segment and generalize different kinds of soccer playing activity during a RoboCup match.*

1 Introduction

Complex organisms recognize their relation to their surroundings and act accordingly. The above sentence sounds like a truism owing in part to the almost ubiquitous distinction between recognition and response in academic disciplines. Engineering has successfully developed pattern recognition and control as independent fields, and cognitive psychology and neuroscience often distinguish between sensory and motor processing with researchers specializing in one area or the other. Nevertheless, in some sense recognition and response entail one another. Recognizing an object, action, or sign is largely a matter of recognizing what it does for us and what we can do with it. Indeed, much of what we perceive can be described in terms of potential actions. *Doing* and *seeing* cannot so readily be distinguished because we acquaint ourselves with our world through what we do and our actions drive what distinctions we learn to make. None of this is meant to deny that we can experimentally isolate purely motor centers in the brain from purely

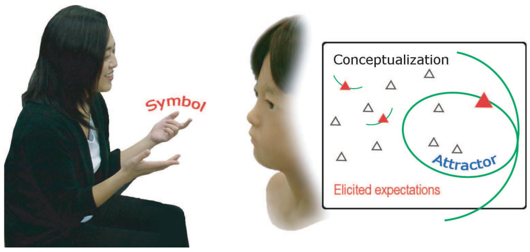


Figure 1. In the proposed approach, a neural network learns each kind of periodic or transitional movement in order to recognize and to generate it. Recent sensorimotor data elicit activity in corresponding networks, which segment the data and produce appropriate anticipatory responses. Active networks constitute an organism's conceptualization of the world since they embody expectations, derived from experience, about the outcomes of acts and what leads to what. It is assumed that behavior is purposive: affective appraisals guide the system toward desired states.

sensory ones, but rather to assert that these centers are intimately linked both in their everyday operation and in their epigenetic development. Thus, as scientists and engineers, we may have reified the distinction between recognition and response, when their main difference is merely in descriptive focus.

In this paper, we will entertain and begin to explore two controversially and, as yet, unproven hypotheses: First, there is an integrity of recognition and response. We recognize an object or event largely because it elicits expectation about what we can do with it — or at least piggybacks on those kinds of expectations. In addition, these expectations are expressed in terms of (or decontextualized from) how motor signals transform sensory data. Second, biological motion is fundamentally periodic. To put it simply, patterns

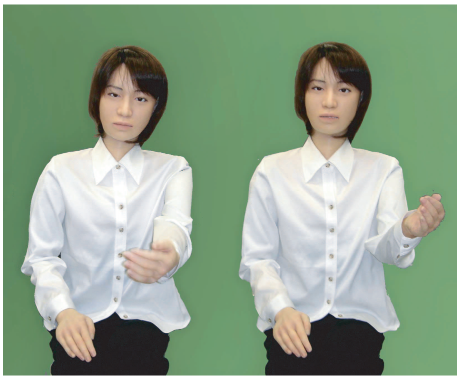


Figure 2. Actroid, the actress android, has 33 motors, which are driven by compressed air, to move its head, neck, arms, body, eyes, eyelids, and mouth. Actroid can make smooth and natural movements, including large and small gestures. Actroid has touch sensors in the arms and can access floor and infrared sensors and video cameras placed in the environment.

repeat. (If they did not, there would be little point in learning.) That is as much a function of the 'hardware' as it is the often routine nature of existence. Joints, for example, have a limited range and will eventually return, more or less, to a given configuration. Moreover, bodies have certain preferred states: for people walking is a more efficient means of locomotion than flailing about randomly. All gaits exhibit a certain periodicity as do many gestures and vocalizations.

This paper proposes a method of generalizing, recognizing, and generating patterns of behavior based on nonlinear principal component neural networks that are constrained to model both periodic and transitional movements. Each network is abstracted from a particular kind of movement. Learning is competitive because sensorimotor patterns that one network cannot learn will be assigned to another network, and redundant networks will be eliminated and their corresponding data reassigned to the most plausible alternative. Recognition is also competitive because proprioceptive data is associated with the network that best predicts it. (The data can be purely kinematic or dynamic depending on the dimensions of the sensorimotor phase space.) Since each network can recognize, learn, and generalize a particular type of motion and generate its generalization, the integrity of recognition and response are maintained. These generalizations are grounded in sensorimotor experience. They can be varied, depending on the networks' parametrization. They may be viewed as a kind of protosymbol. While we do not claim that the networks have neural analogues, we believe the brain must be able to implement similar functions.

1.1 The emergence of signs in communication

In one vein, we are exploring the application of periodically-constrained NLPCA neural networks to vocal and gesture recognition and generation. Our aim is to develop robots whose activity is capable of supporting the emergence of shared signs during communication. Signs take on meaning in a given situation and relationship, as influenced by an individual's emotional responses and motivation (see Figure 1). They reflect mutual expectations that develop over the course of many interactions. We hypothesize that signs provide developmental scaffolding for symbol emergence. For infants, the caregiver's intentions are key to fostering the development of shared signs.

We believe that periodically-constrained NLPCA neural networks could be one of the embedded mechanisms that support the development of shared signs. We are testing this hypothesis by comparing the behavior generalized by these neural networks with Vicon motion capture data from mother-infant interactions.[1] The results of behavioral studies are applied to the android robot, Actroid, which has 33 degrees of freedom (See Figure 2).

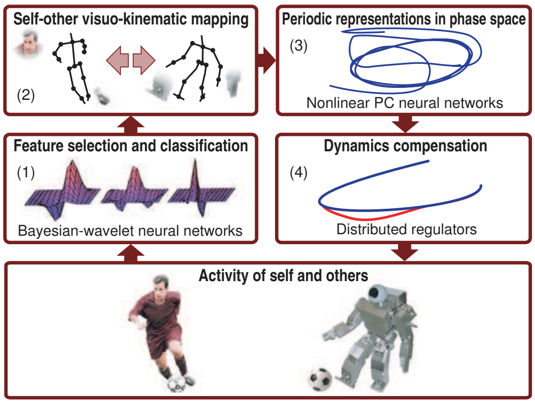


Figure 3. Periodic nonlinear principal component networks may characterize motion patterns in a much larger system for recognizing, learning, and responding behavior.

1.2 Mimesis loop

In a separate vein, we are applying NLPCA neural networks to the learning of cooperative behavior in robot soccer. Although techniques from reinforcement learning can be borrowed to guide a robot's behavior toward goals, they cannot be directly applied to the state space of a humanoid robot because of its enormous size. The approach outlined

[1] From this we have ascertained that certain important micro-behaviors that make movement seem lifelike may have been overlooked in the approach outlined here, and we are starting to develop a micro-behavior filter.

in this paper can vastly reduce the size of the state space by segmenting it into different kinds of movements. A mimesis loop [3] may be used to capture many aspects of the sort of imitation involved in learning to play soccer and other sports. This paper addresses one aspect of the mimesis loop: the abstraction of a robot's own kinematic motions from its proprioceptive experience. Figure 3 roughly outlines how a mimesis loop might be realized in a soccer playing robot. Attentional mechanisms direct the robot's sensors toward the body parts of other players, and the robot maps successfully recognized body parts onto its own body schema. This paper introduces a method to abstract the robot's own kinematic patterns: our segmentation algorithm allocates proprioceptive data among periodic temporally-constrained nonlinear principal component neural networks (NLPCNNs) as they form appropriate generalizations.

The robot can use NLPCNNs to recognize the activities of other players, if the mapping from their bodies to its own has already been derived by some other method. Since each network correspond to a particular type of motion in a proprioceptive phase space, it can act as a protosymbol. Thus, the robot would be able to recognize the behavior of others because it has grounded their behavior in terms of its own body.

Although periodic NLPCNNs may be used to generate motion patterns, the robot must continuously respond to unexpected perturbations. There are a number of approaches to this control problem that do not require an explicit model. For example, distributed regulators [2] could set up flow vectors around learned trajectories, thus, converting them into basins of attraction in a phase space of possible actions.

1.3 Outline

This paper is organized as follows. Section 2 extends an NLPCNN with periodic and temporal constraints. Section 3 presents a method of assigning observations to NLPCNNs to segment proprioceptive data. Section 4 reports experimental results using NLPCNNs to characterize the behavior of a Fujitsu HOAP-1 humanoid robot that has been developed to play RoboCup soccer.

2 A periodic nonlinear principal component neural network

The human body has 244 degrees of freedom [15] and a vast array of proprioceptors. Excluding the hands, a humanoid robot generally has at least 20 degrees of freedom — and far more dimensions are required to describe its dynamics precisely. However, many approaches to controlling the dynamics of a robot are only tractable when sensory data is encoded in fewer dimensions (e.g., [9]). Fortunately, from the standpoint of a particular activity, the effective dimensionality may be much lower.

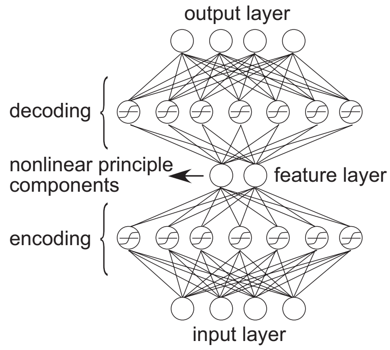


Figure 4. Target values presented at the output layer of a nonlinear principal component neural network are identical to input values. Nonlinear units comprise the encoding and decoding layers, while either linear or nonlinear units comprise the feature and output layers.

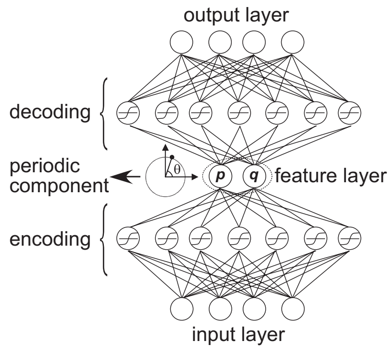


Figure 5. An NLPCA neural network with the activations of nodes p and q constrained to lie on the unit circle.

Given a coding function $f : \mathbb{R}^N \mapsto \mathbb{R}^P$ and decoding function $g : \mathbb{R}^P \mapsto \mathbb{R}^N$ that belong to the sets of continuous nonlinear functions $\mathcal{C}$ and $\mathcal{D}$, respectively, where $P < N$, nonlinear principle component networks minimize the error function E

$$\|\vec{x} - g(f(\vec{x}))\|^2, \quad \vec{x} \in \mathbb{R}^N$$

resulting in P principal components $[y_1 \cdots y_p] = f(\vec{x})$. Kramer (1991) first solved this problem by training a multilayer perceptron similar to the one shown in Figure 4 using the backpropagation of error, although a second order method such as conjugant gradient analysis converges to a solution faster for many large data sets. Tatani and Nakamura (2003) were the first to apply an NLPCNN to human and humanoid motions, though for dimensionality reduction only.

Nonlinear principal components analysis, unlike PCA (Karhunen-Loève expansion), which is a special case where

$\mathcal{C}$ and $\mathcal{D}$ are linear, does not have a unique solution, and no known computational method is guaranteed to find any of the globally optimal solutions. Nevertheless, for a 20-DoF humanoid robot, a hierarchically-constructed[2] NLPCNN has been shown to minimize error several times more than PCA when reducing to two-to-five dimensions [13].

2.1 The periodicity constraint

Because the coding function f of an NLPCNN is continuous, *(1)* projections to a curve or surface of lower dimensionality are suboptimal; *(2)* the curve or surface cannot intersect itself (e.g., be elliptical or annular); and *(3)* projections do not accurately represent discontinuities [8]. However, since the physical processes underlying motion data are continuous, discontinuities do not need to be modelled. Discontinuities caused by optimal projections can create instabilities for control algorithms (e.g., they allow points along the axis of symmetry of a parabola to be projected to either side of the parabola). Moreover, an NLPCNN with a circular node (Ridella et al., 1995, 1997) at the feature layer can learn self-intersecting curves and surfaces.

Kirby and Miranda (1996) constrained the activation values of a pair of nodes p and q in the feature layer of an NLPCNN to fall on the unit circle, thus acting as a single angular variable:

$$r = \sqrt{y_p^2 + y_q^2}, \;\; y_p \leftarrow y_p/r, \;\; y_q \leftarrow y_q/r$$

The delta values for backpropagation of the circular node-pair are calculated by the chain rule [4], resulting in the update rule

$$\delta_p \leftarrow (\delta_p y_q - \delta_q y_p) y_q / r^3, \;\; \delta_q \leftarrow (\delta_q y_p - \delta_p y_q) y_p / r^3$$

at the feature layer.

The hyperbolic tangent and other antisymmetric functions (i.e., $\varphi(x) = -\varphi(x)$) are generally preferred to the logistic function as the sigmoid in part because they are compatible with standard optimizations [6].[3] In addition, antisymmetric units can more easily be replaced with linear or circular units in the feature layer, since these units can produce negative activations. We propose using a slightly flatter antisymmetric function for the sigmoidal units with a similar response characteristic to $tanh$ (see Fig. 6). The advantage of this node is that it can be converted to a circular node-pair while still making use of its perviously learned weights.

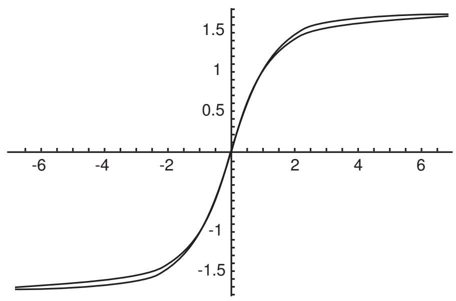


Figure 6. The popular hyperbolic tangent activation function $y \leftarrow 1.7159 \tanh(\frac{2}{3}y)$ can be approximated by a pair of circular nodes where the activation of the second node y_q is fixed at $\sqrt{1.9443}$ and the activation of the first node is calculated accordingly $y_p \leftarrow 1.7159 y_p / \sqrt{y_p^2 + 1.9443}$.

2.2 The temporal constraint

Neither linear nor nonlinear principal components analysis represent the time, relative time, or order in which data are collected.[4] This information, when available, can be used to reduce the number of layers and free parameters (i.e., weights) in the network and thereby its risk of converging slowly or settling into a solution that is only locally optimal. Since the activations y_p and y_q of the circular node-pair in the feature layer in effect represent a single free parameter, the angle θ, if θ is known, we can train the encoding and decoding subnetworks separately by presenting $k\cos(\theta)$ and $k\sin(\theta)$ as target output values for the encoding subnetwork and as input values for the decoding network.[5] Once a single period of data has been collected, temporal values can be converted to angular values $\theta = 2\pi\frac{t_k - t_0}{t_n - t_0}$ for data collected at any arbitrary time t_k during a period, starting at t_0 and ending at t_n. A network may similarly learn transitions between periodic movements when using a linear or sigmoidal activation node in the feature layer because these open-curve transitions do not restrict us to using nodes capable of forming a closed curve.[6] NLPCNNs with a circular feature node remain useful to identify the period of a motion pattern, especially when the pattern is irregular and, thus, begins and ends at points that are somewhat far from each other.

[2] The joint encoder dimensionality of limbs is independently reduced, then the arms and the legs are paired and their dimensionality further reduced, and then finally the dimensionality of the entire body.

[3] These include mean cancellation, linear decorrelation using the K-L expansion, and covariance equalization.

[4] Although a temporal dimension could be added to an autoassociative network, one drawback for online learning is that this dimension would need to be continuously rescaled as more data is collected to keep it within the activation range of the nodes.

[5] $k \approx 1.7$ for zero-mean data with variance equal to 1 based on principles discussed in [6].

[6] $y_{target} = 2k(\frac{t_k - t_0}{t_n - t_0} - \frac{1}{2})$, with $k \approx 1.4$.

3 Automatic segmentation

We conceived of the automatic segmentation problem as the problem of uniquely assigning data points to nonlinear principal component neural networks. It is possible to partition the points without reference to the predictions of the networks.[7] However, for our method each network's performance influences segmentation with more networks assigned to regions that are difficult to learn.

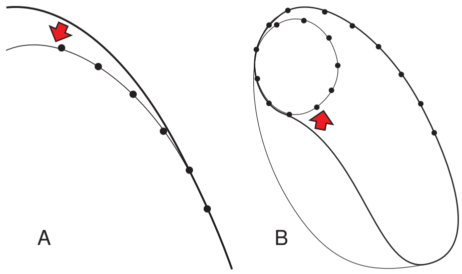


Figure 7. The thick line shows the output of an NLPCNN and the thin line shows the underlying distribution. The dots are data points. **A.** Before learning converges, allowing the network to learn data points despite a high prediction error accelerates learning. **B.** However, after convergence, it leads to segmentation errors.

As the robot begins to move, the first network is assigned some minimal number of data points (e.g., joint-angle vectors), and its training begins with those points. This gets the network's learning started quickly and provides it with enough information to determine the orientation and curvature of the trajectory. If the average prediction error of the data points assigned to a network is below some threshold, the network is assigned additional data points until that threshold has been reached. At that point, data points will be assigned to another network, and a network will be created, if it does not already exist. To avoid instabilities, only a single data point may shift its assignment from one network to another after each training cycle.

Since a network is allowed to learn more data points as long as its average prediction error per point is low enough, it may learn most data points well but exhibit slack near peripheral or recently learned data points. At the start of learning, the network should be challenged to learn data points even when its prediction error is large (see Fig. 7A). As learning converges, however, the slack leads to segmentation errors (see Fig. 7B). Therefore, we alter the method of segmentation once the network nears convergence (as determined by Bayesian methods [7] or crossvalidation) so that

$j \leftarrow 1,\ bucket \leftarrow 1,\ E \leftarrow 0$
$\forall \vec{x_i}$ {
  train($network_j, \vec{x_i}$)
  $E_i = \|\vec{x_i} - g(f(\vec{x_i}))\|^2,\ E \leftarrow E + E_i$
  if ($bucket > B_{max}$ $\vee$
    (learning?($network_j$) $\wedge\ E/bucket > E_{max}$) $\vee$
    $E_i > E_{i+1}$)
    $j \leftarrow j+1,\ bucket \leftarrow 1,\ E \leftarrow 0$ }

Listing 1: Pseudocode for segmentation.

a network may acquire neighboring points if its prediction error for those points is lower that the network currently assigned to those points.

4 Humanoid experiments

This section shows the result of automatic segmentation and neural network learning. We assess the accuracy of the result based on a manual segmentation of the data points and an analysis of how they are allocated among the networks.

First, we recorded motion data while a HOAP-1 humanoid robot played soccer in accordance with a hard-coded program [1]. Each data point is constituted by a 20-dimensional vector of joint angles. A standard (noncircular) NLPCNN reduced the dimensionality of the data from 20 to 3, which was determined to be the intrinsic dimensionality of the data by the ISOMAP procedure [14] We then applied our algorithm to segment, generalize, and generate humanoid motion.

Our algorithm uniquely assigned the data points among a number of circularly-constrained NLPCNNs. Each of the networks learned a periodic motion pattern by conjugate gradients. Our algorithm successfully generalized five out of six primary motion patterns: walking forward, turning

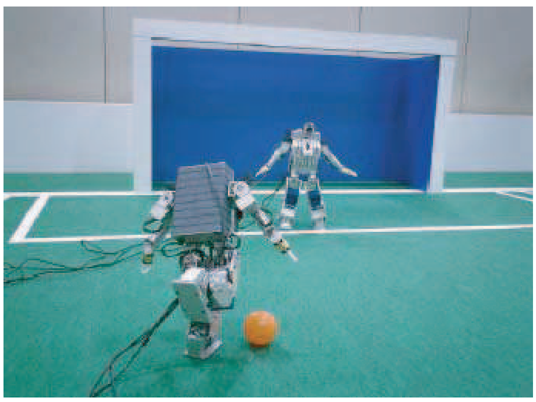


Figure 8. Fujitsu HOAP-1 robots are playing RoboCup soccer.

[7]For example, data points may be partitioned at the point at which a trajectory most closely doubles back on itself, if the distance between the two paths is within a certain threshold and the paths then diverge beyond another threshold.

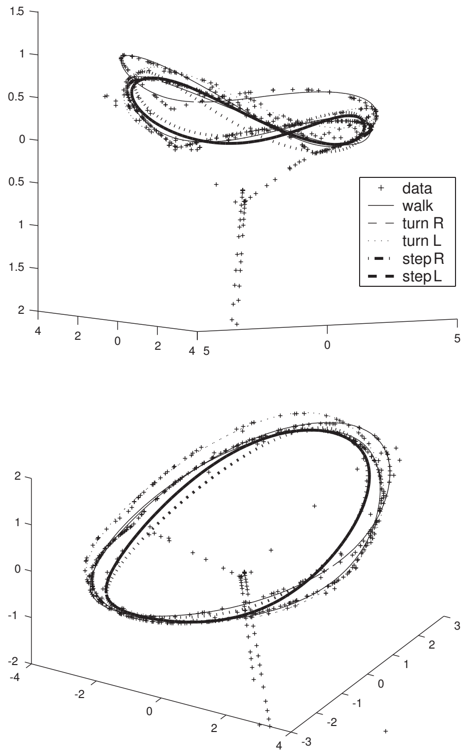


Figure 9. Recognized motion patterns embedded in the dimensions of the first three nonlinear principal components of the raw proprioceptive data. The top and bottom plots differ only in the viewpoint used for visualization.

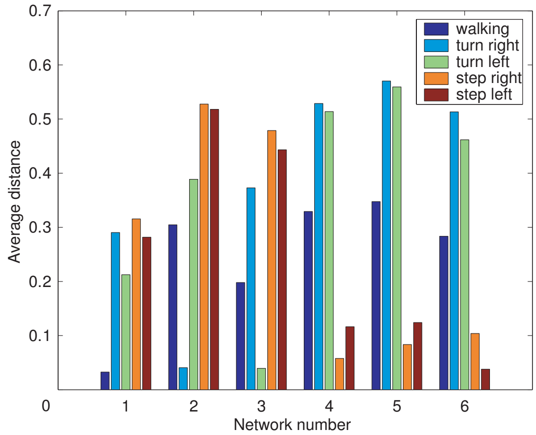


Figure 10. The average distance between the prediction of a network trained on manually segmented data and each of the automatically generated networks.

right or left, and side-stepping to the right or left. It failed to generalize as a single periodic trajectory the kicking motion, which has a highly irregular, self-intersecting shape. However, human subjects were also unable to determine the kicking trajectory from the data points.

Figure 9 shows that the automatic segmentation algorithm successfully employed circular NLPCNNs to separate and generalize five of the periodic motions. (The open-curve segmentation of transitions between periodic motions are omitted for clarity.) The periodic trajectories were generated by varying from 0 to 2π the angular parameter θ_i at the bottleneck layer of each of the circularly-constrained networks and mapping the result to the output layer for display. This demonstrates our method's capacity to generate periodic motions.

We calculated statistics based on running the automatic segmentation for 20 trails. The algorithm resulted in five decoding subnetworks for 45% of the trials, which is the most parsimonious solution. It resulted in six subnetworks for 50% of the trials, and seven for the remaining 5%.

Since the data was generated by the predefined behavior modules used by the Osaka University team in the 2003 RoboCup humanoid competition, each data point was already labelled and could be segmented into the five types of motion that had been successfully abstracted. To assess the accuracy of the automatic segmentation algorithm, we manually assigned the data points corresponding to each type of motion to five periodic temporally constrained NLPCNNs. Figure 10 shows the average distance between the prediction for each of these networks and each of the networks resulting from automatic segmentation.

The lowest bar indicates which pattern the networks, numbered 1 to 6 best match in terms of least average distance. Hence, the first network represents walking; the second represents turning right; the third turning left; the fourth and fifth sidestepping right; and the sixth sidestepping left. The fact that the fifth network is redundant, abstracting the same type of motion as the fourth, does not prevent the abstracted actions from supporting the mastery of soccer or some other task. Both networks can be used. The algorithm's capacity to reduce a vast amount of complex, raw data to just a few states may help reinforcement learning approaches to finesse the curse of dimensionality [12].

In a separate run of the learning and segmentation algorithm, the motion sequence of recorded data during soccer playing was walking forward, turning right, turning left, walking forward, sidestepping to the right, sidestepping to the left, and kicking. We counted the number of point belonging to each network before and after removing redundant networks. Redundant networks were removed by means of linear integration. The angular value θ was varied from 0 to 2π at the bottleneck layer of one network to obtain its predicted output. This value was fed into another network to obtain its predicted value. If the integral of the sum of the squared distances of the predicted outputs

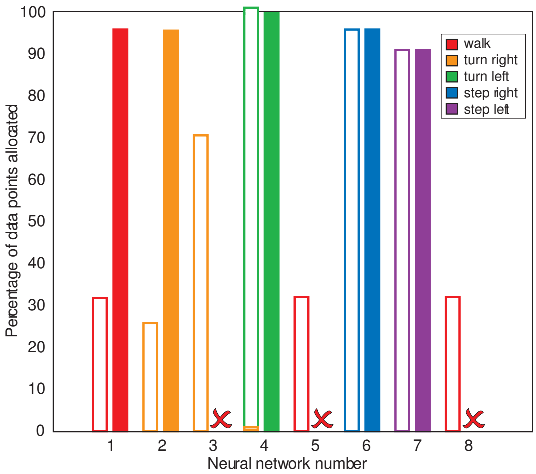


Figure 11. The percentage of data points allocated to each network before and after eliminating redundant networks and reassigning their data.

was less than a threshold, one network was removed and its points reassigned to the other network (see Figure 11). This method removed all redundant networks.

5 Conclusion

Our proposed algorithm abstracted five out of six types of humanoid motion through a process that combines learning and data point assignment among multiple neural networks. The networks perform periodic, temporally-constrained nonlinear principal component analysis. The decoding subnetworks generate motion patterns that accurately correspond to the five motions without including outliers caused by nondeterministic perturbations in the data. During 45% of training episodes, the algorithm generated no redundant networks; a redundant network appeared in 50% of the training episodes, and two appeared in 5% of them. Although the fourth and fifth networks represent the same type of motion, this does not prevent them from serving as action symbols for learning a complex task. By means of linear integration, we were able to remove redundant networks according to the proximity of their predictions.

A kind of behavior can be recognized by selecting the network that best predicts joint-angle values. It can be generated by varying the value of θ in the bottleneck layer. This shows the effectiveness of a tight coupling between recognition and response since the same networks may be used for both processes and they developed by the same mechanisms. The significance of periodicity may be more limited, however. Some motions are not periodic, and in the experiment the kicking motion, although it occurs repeatedly, was difficult to segment because of its highly irregular, self-intersecting shape.

References

[1] R. Chalodhorn, M. Aono, J. Ooga, M. Ogino, and M. Asada. Osaka University "Senchans 2003". In B. Browning, D. Polani, A. Bonarini, and K. Yoshida, editors, *RoboCup-2003: Robot Soccer World Cup VII*. Springer Verlag, 2003.

[2] T. Fujii. A new approach to the LQ design from the viewpoint of the inverse regulator problem. *IEEE Transactions on Automatic Control*, 32(11):995–1004, 1987.

[3] T. Inamura, I. Toshima, and Y. Nakamura. Acquiring motion elements for bidirectional computation of motion recognition and generation. In B. Siciliano and P. Dario, editors, *Experimental Robotics VIII*, pages 372–381. Springer, 2003.

[4] M. J. Kirby and R. Miranda. Circular nodes in neural networks. *Neural Computation*, 8(2):390–402, 1996.

[5] M. A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *Journal of the American Institute of Chemical Engineers*, 37(2):233–243, 1991.

[6] Y. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller. Efficient BackProp. In G. B. Orr and K.-R. Müller, editors, *Neural Networks: Tricks of the Trade*, pages 1–44. Springer, 1998.

[7] D. J. MacKay. Probable networks and plausible predictions: A review of practical Bayesian methods for supervised neural networks. *Network: Computation in Neural Systems*, 6:469–505, 1995.

[8] E. C. Malthouse. Limitations of nonlinear PCA as performed with generic neural networks. *IEEE Transactions on Neural Networks*, 9(1):165–173, 1998.

[9] M. Okada, K. Tatani, and Y. Nakamura. Polynomial design of the nonlinear dynamics for the brain-like information processing of whole body motion. In *IEEE International Conference on Robotics and Automation*, pages 1410–1415, 2002.

[10] S. Ridella, S. Rovetta, and R. Zunino. Adaptive internal representation in circular back-propagation networks. *Neural Computing and Applications*, 3(4):222–333, 1995.

[11] S. Ridella, S. Rovetta, and R. Zunino. Circular backpropagation networks for classification. *IEEE Transaction on Neural Networks*, 8(1):84–97, 1997.

[12] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, Massachusetts, 1998.

[13] K. Tatani and Y. Nakamura. Dimensionality reduction and reproduction with hierarchical NLPCA neural networks extracting common space of multiple humanoid motion patterns. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1927–1932, Taipei, Taiwan, September 2003.

[14] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[15] V. M. Zatsiorsky. *Kinematics of Human Motion*. Human Kinetics, Urbana Champaign, 2002.

Detecting Contingency Between Self and Other Triggers Social Behavior.

Yukie Nagai, Minoru Asada, and Koh Hosoda

Abstract

This study investigates what triggers the shift of human infants' behaviors from self-centered (ScB) to other-depended ones (OdB) for the emergence of social capabilities. Joint attention ability is known to be acquired through a three-staged process, in which infants gradually shift their behaviors from ScB to OdB. The authors have proposed a constructivist model by which a robot learns joint attention through experiences of visual attention. Visual attention is a ScB to gaze at a salient visual stimulus. Employing the model, our robot acquired the sensorimotor coordination of joint attention by detecting a contingency between the image of a human face and a motor command to look at a object. Analysis of the relationship between the learning convergence and the behavioral shift showed that: (a) when gradually shifting from ScB to OdB according to the contingency detection, the robot can acquire joint attention ability; (b) when producing only ScB over the learning phase, the robot cannot acquire a consistent sensorimotor coordination; (c) when adopting only OdB, the robot falls into locally biased behaviors that were experienced earlier. These results suggest that the emergence of infants' social behaviors is triggered when they detect a contingency between their own and other behaviors.

Cognitive Foundations of Conventions in Social Interaction

Dale J. Barr
University of California, Riverside
dale.barr@ucr.edu

I present results from a set of psycholinguistic studies that challenge the view that the establishment and use of conventions requires mutual knowledge. Instead, the results suggest that people use conventions in ways that routinely violate mutual knowledge. Based on these findings, I argue that conventions are grounded not in complex assessments about what others know, but in simple, low-level cognitive heuristics that provide a robust, but fallible, basis for coordination at a minimal cognitive cost.

One of the hallmarks of human intelligence is the ability to make use of socially shared conventions in order to solve coordination problems. Language use provides perhaps the most conspicuous example of how social interaction is governed by conventions. Languages are comprised of multiple levels of conventions—conventions of phonology, morphology, syntax, and discourse. An important question is how people establish and use such linguistic conventions.

An influential proposal is that the establishment and use of conventions depends on the accumulation of a certain kind of shared knowledge, what is known as "mutual knowledge" or common ground [1, 2]. Mutual knowledge is defined as the set of knowledge that interlocutors share, know that they share, know that they know that they share, and so on. An emerging alternative view suggests that much of convention use may not require participants to explicitly access mutual knowledge, and that the coordination phenomena observed in conversation might be an emergent effect of low-level cognitive processes [3, 4]. However, little is known about the on-line processing that underlies these emergent effects.

To investigate this issue I tracked the eyes of speakers and listeners as they coordinated reference in a referential communication task. I examined how the use of scalar adjectives (e.g., "small") became conventionalized through repeated use. The experiment focused on factors of frequency and mutual knowledge.

The results indicated that convention use was determined by frequency, but not mutual knowledge: speakers and listeners continued to use newly-established conventions even when they interacted with partners who lacked mutual knowledge of these conventions. Although this egocentric behavior may seem sub-optimal from the point of view of successful coordination, it is argued to be ecologically valid because of the existence of rich feedback loops that promote a commonality of cognitive representation in the dyad and in the community [5].

[1] H. H. Clark & C. R. Marshall, "Definite reference and mutual knowledge", in A. K. Joshe, B. L. Webber & I. A. Sag (Eds.), *Elements of discourse understanding,* Cambridge University Press, Cambridge, 1981, pp. 10-63.

[2] D. Lewis, *Convention: A philosophical study,* Harvard University Press, Cambridge, MA, 1969, pp. 10-63.

[3] D. J. Barr & B. Keysar, "Making sense of how we make sense: The paradox of egocentrism in language use," in H. L. Colston & A. N. Katz (Eds.), *Figurative language comprehension: Social and cultural influences*, Erlbaum, Mahwaw, N.J., 2004.

[4] M. J. Pickering & S. Garrod, "Toward a mechanistic psychology of dialogue," *Behavioral and Brain Sciences*, Cambridge University Press, Cambridge, in press.

[5] D. J. Barr, "Establishing conventional communication systems: Is common knowledge necessary?" *Cognitive Science*, Erlbaum, Mahwaw, N.J., in press.

Are you synching what I'm synching? Modeling infants' real-time detection of audiovisual contingencies between face and voice

George Hollich,[*] Eric J. Mislivec,[**] Nathan A. Helder, [**] & Christopher G. Prince[**]

[*] *Department of Psychological Sciences*
Purdue University
West Lafayette, IN 47907 USA
gholich@purdue.edu

[**] *Department of Computer Science*
University of Minnesota Duluth
Duluth, MN 55812 USA
misli001@d.umn.edu, nhelder@nerp.net, chris@cprince.com

Abstract

Audio-visual synchrony is one of the earliest and most salient properties to which infants are sensitive.[1] Furthermore, it is likely that detection of contingent relations in and across modalities is a critical beginning point for autonomous mental development.[2] While there are numerous ecological examples of the need for contingency detection, one of the strongest is connecting face and voice. Dodd[3] demonstrated that infants would look longer to a face that is synchronized to speech than one that is asynchronous with speech (see also Pickens[4]).

The goal of the research reported here is to explicitly contrast detailed empirical data capturing infants' real-time detection of speech/face synchrony with a formal model of audio-visual synchrony detection. The empirical data comes from the Purdue University Infant Laboratory. Infants, 4, 8, and 12 months of age, were tested in a cross-sectional design using the splitscreen preferential looking paradigm. In the procedure, two female faces (talking in infant-directed speech) were presented side-by-side on a large video screen, with the audio alternately matching one of the faces. By following the developmental trajectory of the preference for the synchronous face and by examining reaction times when the synchrony switches between faces, we gain a better understanding of the temporal resolution of infants' sensitivity to synchrony at different ages. This method also gives us frame-by-frame coding of infant looking preferences that can be directly compared with the output of the formal model. In the model, we use an algorithm that directly computes moment-by-moment audio-visual synchrony relations between low-level audio-visual features (e.g., RMS audio and grayscale pixels) based on Gaussian mutual information across a time window of audio-visual information.[5]

While the ability of the model to discover and localize sources of synchrony is still in its infancy, it already shows strikingly similar overall and moment-by-moment performance to some of the data from the infants. This suggests that infants and the model may be tapping similar aspects of the audio-visual contingencies in the video. It is our hope the model will ultimately capture even more detailed aspects of infants' behavior and scale to more general models of infant attention and autonomous development. Following this motivation, we are extending the model to utilize audio-visual synchrony to train within-visual-modality categorization and to bootstrap aspects of facial recognition.

References

[1] Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*, 190-201.

[2] Fasel, I., Deak, G., Triesch, J., & Movellan, J. (2002). Combining embodied models and empirical research for understanding the development of shared attention. In *Proceedings of the 2nd International Conference on Development and Learning* (pp. 21-27). Cambridge, MA: IEEE Computer Society Press.

[3] Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*, *11*, 478-484.

[4] Pickens, J., Field, T., Nawrocki, T., Martinez, A., Soutullo, D., & Gonzalez, J. (1994). Full-term and preterm infants' perception of face-voice synchrony. *Infant Behavior and Development, 17*, 447-455.

[5] Hershey, J. & Movellan, J. (2000). Audio-vision: Using audio-visual synchrony to locate sounds. In S. A. Solla, T. K. Leen, & K. R. Müller (Eds.), *Advances in Neural Information Processing Systems 12* (pp. 813-819). Cambridge, MA: MIT Press.

ATTENTION-SHARING IN HUMAN INFANTS FROM 5 TO 10 MONTHS OF AGE IN NATURALISTIC INTERACTIONS

Gedeon Deák , Yuri Wakabayashi & Hector Jasso

Abstract

Attention-sharing is fundamental for learning in social contexts. Some accounts of shared attention skills (Baron-Cohen, 1995) explain them as unverified, innate modules. An alternative account (Deák & Triesch, in press) proposes that attention-sharing skills like gaze- and point-following emerge from a combination of perceptual routines, affective dispositions, learning capacities, and exposure to structured social input. Thus, shared attention behaviors like gaze- & point-following should emerge gradually from interactive processing of predictable adult behaviors. In this study, 11 parent-infant dyads, with infants between 5 and 10 months of age (i.e., before or during the emergence of reliable laboratory-based gaze- and point-following skills) were videotaped at play at their home. The object was to observe how gaze- and point-following change from 5 to 10 months in everyday social interactions. Dual-video files (2 synchronized, 15 min DV streams, one each focused on parent and infant) were coded for events and variables including infant's direction and target of gaze, parent's direction and target of gaze, and parent's manual actions. The results for "showing" interactions (when parents were told, with no specific instructions, to _try to get [your baby] to pay attention to these [4 toys]_), and for _peek-a-boo_ interactions, were coded. Inter-coder reliability for gaze target was acceptable; parent kappa = .73; infant kappa = .70. Results showed that infants and parents spent a mean of 12% (range = 1-25%) of the sharing episode in mutual gaze, and 9% (range = 3-23%) in shared gaze. Frequency of infant following events was positively correlated with infant_s age, r (10) = .69. From 5 to 10 months, infants' following rate for all gestures reliably and gradually increased, and no sudden change around 9-10 months evident, as modular theories would predict. Unexpected from the experimental literature, we found that infants followed only 3.6% of parents_ gaze shifts, versus 40.1% of parents_ pointing gestures. across the age range followed parents' points far more than gaze. It has been claimed that infants younger than 10 months do not understand others_ pointing gestures; these data disconfirm that claim. The results indicate that gaze- and point-following emerge gradually in the first year, implying learning processes based on structured input, rather than the "turning on" of a specialized module or modules.

Kinesthetic-Visual Matching and Consciousness of Self and Other: How Social Minds are Possible

Robert W. Mitchell
Department of Psychology
127 Cammack Building
Eastern Kentucky University
Richmond, KY 40475 USA
Email: robert.mitchell@eku.edu

In 1925, the developmental psychologist Paul Guillaume argued that young children were able to recognize themselves in mirrors and to imitate another's actions because they were able to match between their kinesthetic experiences of their own body and their visual experiences of another body: either their own (in the mirror) or another's (in imitation). Through this kinesthetic-visual matching, the young child knows how it looks even without a mirror, in that it can imagine, from its kinesthetic experiences, a visual appearance of itself. And from a visual experience of a body, it can imagine how that body feels kinesthetically. Guillaume recognized that his solution to the problems of explaining self-recognition and imitation also solved another problem: how it is that we ever develop the notion that there are subjects of experience. Guillaume's solution offered a better explanation than the traditional argument by analogy. According to the argument by analogy, any given person knows that others have conscious experiences by extrapolation from his or her own experience. Yet this argument is clearly inaccurate because, as the philosopher P. F. Strawson argued, unless one can apply generally the idea of a subject of experience (what Strawson calls the concept of a "person"), one would be unaware that one <u>is</u> a subject of experience. Kinesthetic-visual matching provides a person with the ability to recognize itself as a subject of experience, as well as to recognize that others have subjective experiences like itself: the visual experience of a body, whether one's own or another's, is perceived as endowed with kinesthetic experience. Similarly, philosophical psychologist Merleau-Ponty observed that one needs a schema that is transferable from one's own body to others' bodies and back again in order to experience oneself and others as psychological beings, and kinesthetic-visual matching fits the requirement. It is kinesthetic-visual matching that allows individuals a foothold on the notion that oneself and others are persons. Although psychologists Meltzoff and Moore posited (many years after Guillaume) a mechanism for interpersonal understanding similar to kinesthetic-visual matching, it is unlikely that the newborns to whom they attribute the mechanism are capable of understanding others in the same way that older children who self-recognize and engage in generalized bodily imitation can. The mature kinesthetic-visual matching present in self-recognizers who can imitate innumerable bodily actions seems essential for other skills as well, including recognizing that one is being imitated, extensive planning involving imagination of one's body moving through space, pretending to be another, and communicating via imitation. The ability for kinesthetic-visual matching appears to be localized in the brain's parietal region, an area responsible for motor imagery and other imaginative movement of one's body through space. In fact, it appears likely that the self-other relational properties of kinesthetic-visual matching derive evolutionarily from motor imagery skills. In this presentation I examine the implications of kinesthetic-visual matching for self-understanding and understanding of other minds in adult humans, young children, great apes, and dolphins who are capable of both self-recognition and generalized bodily imitation.

EEG dynamics during self-produced emotion feeling-states

Julie Onton, Scott Makeig

Emotions are a fundamental part of being human, yet our understanding of how emotions affect the brain is limited. One major reason is that recording brain activity during genuine emotional states is difficult to achieve in a laboratory setting. Pictures of intensely emotional or disturbing facial expressions or scenes can trigger emotional reactions, but such methods ignore the range of feeling states that pervade our everyday experience such as joy, love, compassion, awe, as well as frustration, jealousy, etc. This study attempts to discover the EEG correlates of imaginistic emotional feeling. Subjects were asked, via a voice recording, to recall and/or imagine a series of scenarios in which they had felt or would feel the suggested emotions, in each case allowing the imagery and somatic feeling sensations to become as vivid as possible. During these sessions, 256 channels of EEG data were recorded from the scalp, neck and face. We then decomposed the EEG data using Independent Component Analysis (ICA) into maximally independent time courses and associated spatial maps. We were able to blindly cluster EEG characteristics during different emotional experiences across subjects into at least two dimensions: valence (good/bad), as well as high/low activity. By multiple levels of single-trial spectral clustering from all subjects, we found that emotions shifted the EEG in various ways, sometimes in patterns according to valence, sometimes activity and other times specific emotions tended to dominate a particular spectral change. More specifically, a diffuse cluster of occipital, parietal and temporal components appeared to express more 10-11 Hz alpha during less active emotions such as sadness, compassion, love and relief. Occipital components exhibited a marked shift in alpha frequency from ~10 Hz to ~11 Hz (depending on the resting alpha frequency), during emotional contemplation which appeared to be associated with valence and/or activity dimensions of the emotions. This study demonstrates that the EEG correlates of emotion, while complicated, were decipherable using blind data clustering techniques.

Mu rhythm modulation during intentional and unintentional human and robot actions

Shenk, L.M., Jacoby, B.P., McCleery, J.P., Ramachandran, V.S., & Pineda, J.A.

Abstract

Previous studies have found that electroencephalogram (EEG) oscillations in the mu frequency band (8-13 Hz) are suppressed during the performance and observation of human actions. Mirror neurons, originally discovered in the macaque premotor cortex, are characterized by their activity in response to both self-performed and observed actions. Based on this functional correlation, mu wave suppression is believed to be an index of activity of the human mirror neuron system. Accumulating evidence suggests that the mirror neuron system serves to create an internal simulation of actions that are perceived as either biological or intentional. The present study seeks to investigate the flexibility of the mirror neuron system by exploring its response to both intentional and unintentional biological and non-biological movement. Fourteen adult participants viewed a series of videos including 1) Visual white noise (baseline condition) 2) Balls bouncing (non-biological unintentional movement), 3) A human hand opening and closing (biological intentional movement), and 4) A human hand being pulled by a string (biological unintentional movement). Power in the mu frequency was significantly suppressed during both biological movement conditions. This finding suggests that biological movement is sufficient to suppress the mu wave independent of intentionality. Thus we conclude that the human mirror neuron system is responsive to both intentional and unintentional biological movement. To further explore this question, we are currently collecting data on intentional non-biological (i.e., robot) movement. Based on preliminary data from on-going fMRI mirror neuron studies, we hypothesize that non-biological intentional movements will also be sufficient to suppress the mu wave.

The perception of direct gaze in human infants

Authors: Teresa Farroni, Mark H. Johnson, Gergely Csibra

Direct eye contact with another human face is one of the most important foundations of our social behavior. A major debate in cognitive neuroscience concerns the origins of the "social brain" in humans, and the extent to which this is acquired through experience. In the first of two experiments, we measured 4-month–old infants' brain electrical activity to assess the neural processing of faces when accompanied by direct or averted eye gaze. High-density event-related potentials (ERPs) were recorded in response to the direction of eye gaze of this face stimulus. The results show that, consistent with previous studies (de Haan, Pascalis, & Johnson, 2002), an "infant N170" component peaked around 240 msec post-stimulus. Further, the amplitude of this component over mid-line occipital channels was higher in response to direct than averted gaze. To rule out alternative explanations, in a second experiment we showed babies inverted female faces, with either direct or averted gaze. Inspection of averaged ERPs time-locked to the onset of the stimulus revealed no effect on the “infant N170” corresponding to that observed in Experiment 1. These results suggest that from at least 4 months of age there is enhanced brain processing of upright faces with direct eye gaze.

Learning by Imitation, Reinforcement and Verbal Rules in Problem Solving Tasks

Frédéric Dandurand, Melissa Bowen, Thomas R. Shultz
McGill University, Department of Psychology, 1205 Dr. Penfield Ave., Montréal, Québec, H3A 1B1, Canada, fdandu@ego.psych.mcgill.ca, melissa.bowen@mail.mcgill.ca, thomas.shultz@mcgill.ca

Abstract

Cognitive psychologists have extensively studied feedback and explicit learning in problem solving. In contrast, they often dismissed learning by imitation as "rote memorizing", and therefore considered it trivial and uninteresting. Our work shows that this assumption might be unwarranted. Research on learning in infants and animals regards imitation learning as a powerful and pervasive process for acquiring new knowledge, namely the overall arrangement of actions (sequencing and planning) in complex tasks. Similarly, we propose that problem solvers can infer a complex hierarchical problem representation from observation alone.

We compared three types of learning in problem solving tasks: imitation learning (a group that viewed successful problem solving demonstrations), feedback learning (a group that got feedback indicating whether their answer was correct or not) and explicit rule learning (a group that was presented instructions to solve the problem). Participants were required to find, with three uses of a scale, the one ball which was either heavier or lighter than the rest of a set of 12 balls.

We found that participants in the imitation learning and explicit learning groups were more accurate (solved more problems correctly) than those in the feedback learning group. We also found that all participants got faster, but not more accurate, over problem solving trials.

More research is required to understand the mechanisms underlying imitation learning and the nature of problem representations constructed. However, we conclude that learning by imitation rivals learning by explicit instructions as an efficient learning technique in our problem solving task. While both methods were superior to feedback learning, imitation learning has additional benefits such as being more informal and omnipresent, while imposing a minimal cost to the mentor.

1. Introduction

Problem solving is "thinking that is directed toward the solving of a specific problem that involves both the formation of responses and the selection among possible responses" [25]. It is therefore a very important area of cognitive psychology, and it is considered a crucial component of intelligence. "The ability to solve problems is one of the most important manifestations of human thinking." [12].

Information-processing theory is currently the dominant approach to problem solving [12]. Problems are construed in terms of states, transitions and operators. The essence of problem solving is a search through the state space for a solution state using the operators available at each point while satisfying a set of problem-specific constraints.

1.1. Learning in Problem Solving Tasks

Several types of learning can occur during problem solving efforts, either in isolation or in various combinations.

1.1.1. Reinforcement or Feedback Learning

Feedback learning, or learning by trial-and-error, is a mechanism whereby systems (humans, animals or machines) seek to maximize rewards and minimize punishments. In general, rewarded behaviors tend to increase in frequency whereas punished behaviors tend to decrease in frequency as learning progresses [2, 16]. By rewarding searches that yield correct solutions and punishing those that do not, feedback learning allows problem solvers to learn which tactics are successful and which ones are not.. It is the most common learning mechanism in problem solving because feedback information is usually embedded in the problem itself or is available in the environment. The main problem with feedback learning is that the information available is typically very limited, in the form of binary data (correct/satisfactory or incorrect/unsatisfactory answer).

Although he used a different terminology, Holyoak describes a form of feedback learning as the learning mechanism involved in information-processing theory: "An intelligent problem solver uses the results of solution attempts to acquire new knowledge that will help solve similar problems more readily in the future" [12].

1.1.2. Explicit Learning

Explicit verbal learning, is based on having access to explicit instructions for solving problems. They are typically expressed as abstract symbolic rules in the form of *if… then…* statements. Western cultures particularly value explicit verbal instructions such as problem solving algorithms typically described in textbooks, documents and other "how-to" manuals [29].

Explicit learning is somewhat limited in scope because it assumes the availability of a domain expert and a skilled author who have the time, energy and ability to express problem solving reasoning explicitly, concisely, completely and coherently.

1.1.3. Imitation Learning

Imitation learning, rooted in the long tradition of social learning [4], can be defined as a mechanism where behaviors or skills are acquired by watching others perform. [24]. In the natural world, learning by imitation makes evolutionary sense for social animals because it allows them to learn and transmit successful methods and strategies, possibly acquired over many generations.

There is still a lot of controversy around a precise and detailed definition of imitation learning. Definitions can generally be regarded on a spectrum from inclusive to restrictive. For instance, Thorndike (1898) defined imitation as any situation in which animals "from an act witnessed learn to do an act." [1] In contrast, Thorpe (1963) defined "true imitation as the copying of a novel or otherwise improbable act or utterance, or some act for which there is clearly no instinctive tendency" [1].

Most of the controversy originates from the fact that other mechanisms (generally based on some kind of priming) can also account for imitative behaviors. Those mechanisms include [1, 17]:

1. Social Facilitation / Social Enhancement – The mere presence of conspecifics encourages similar behaviors.
2. Local Enhancement – The attention of the observer is drawn to a place or location due to activities of the demonstrator.
3. Stimulus Enhancement – The attention of the observer is drawn to an object (e.g., tool) due to activities of the demonstrator.
4. Goal Emulation – The imitator does not try to copy the action, but tries to reproduce the result.

Most imitation researchers agree about the importance of ruling out these other mechanisms. To truly qualify as imitation, it appears that some kind of understanding of the demonstrator's intentions is important. Recent research has found evidence of such understanding of intentions in animals and in human infants [8, 9, 27].

Byrne and Russon [7] proposed the idea of a *program level* imitation consisting in imitating the overall arrangement of actions in a hierarchical fashion, particularly the planning of and sequencing of actions. They contrasted this with *action level* imitation where the fine details of the actions are copied or imitated. They argued that program level imitation qualifies as "true imitation" because it implies that the imitator understands the intentions of the demonstrator in terms of goals and sub-goals. They consider the learning of a new arrangement of behavioral units, already present in the behavioral repertoire, to be novel.

In the context of problem solving, imitation learning takes the form of demonstrations. Byrne and Russon's theory suggests that imitation learning might enable a problem solver to infer a complex hierarchical problem representation from observation alone.

Learning by imitation does not have the limitations of explicit learning. Because it conveys information or instructions implicitly in problem solving demonstrations, it can be applied to tasks learned implicitly and to tasks for which no written instructions can be found. In fact, the demonstrator need not be conscious he is being imitated, although awareness of his role might facilitate learning because he can emphasize or highlight critical steps during the demonstration. This makes learning by imitation more efficient and adaptive than explicit learning in many contexts.

Learning by imitation and by feedback are probably mediated by different brain mechanisms, namely mirror neurons for imitation learning [18] and the activity of mesencephalic dopaminergic neurons for reinforcement learning [14]. Both types of learning are used in machine learning [6, 26].

1.2. Supplemental Approaches to Problem Solving

Several methods for complementing, supplementing or replacing feedback learning have been proposed. Techniques traditionally favored include: teaching of heuristics [20], hints [15] and reasoning by analogy [13]. Heuristics are used to limit search complexity by considering only a small number of alternatives that seem most likely to lead to a solution [12].

In contrast, the importance of learning by demonstration (imitation) has been minimized because it was considered rote memorization [15], and therefore trivial and uninteresting. In light of the more recent research just reviewed, we believe that this assumption is unwarranted, and instead, we argue that learning by imitation is actually complex and cognitively challenging.

For instance, Katona [15] explored a matchstick problem with three groups: a *creative* group that was provided hints for solving the problem, a *memory* group that saw demonstrations, and a control group that had no help. The creative group performed the best on the experimental task, and the memory group outperformed the control group. A striking result that attracted little attention was that the memory group, which arguably had an opportunity for learning by imitation, outperformed the control group even on novel problems. This outcome suggests that with a demonstration of even one problem to imitate, people may be able to generalize those strategies to new problems.

Furthermore, heuristics, hints and analogies share many of the problems of explicit learning, namely the need for availability of a mentor, a significant time and effort investment by the mentor, and the need of a task to which hints, heuristics or analogies can be applied. Also, Nisbett and Wilson [19] found that subjects are typically not aware that hints are given, and they are not accurate in determining which hints (among real and false ones) are useful. Gick and Holyoak [10, 11] found that people often fail to make use of potentially useful analogies unless their relevance is explicitly pointed out to them.

2. Project Description

In this research, we compared the effect of learning by imitation, feedback learning and explicit learning on problem solving performance. To the best of our knowledge, learning by imitation has never been explicitly studied in adult humans in the context of higher-level cognition, such as problem solving tasks, with the goal of understanding its underlying mechanisms. Katona's work aimed to show the superiority of hints over demonstrations (memory group). He did not compare the memory group with an instruction group to control for the amount of information subjects got, and did not study the mechanisms underlying learning by imitation.

We limited our study to so-called *well-structured* problems [22]. Such problems are characterized by their clear initial and goal states, and by their precisely defined operators and constraints. Furthermore, this research focused on planning-intensive tasks. The Towers of Hanoi problem is a classical example of a well-structured, planning-intensive task.

For this research, a well-known mathematical problem, the ball-weighing problem, was selected. This class of problems has been previously used in a psychological experiment, the Coin problem [21]. It can be described as follows: "Suppose you have eight coins and a balance. One of the coins is counterfeit, and therefore is lighter than the others. How can you find the counterfeit coin by using the balance only twice?" [5] We used a variant of this problem which involved weighing balls. Our problems used 12 balls, the target ball could be heavier or lighter than the rest and, the scale could be used only three times.

Appendix 1 presents a complete instruction set for solving this problem - only minor variants are possible. The reason for selecting a relatively difficult problem with counter-intuitive solutions was to require learning, and thus enable differential performance between the experimental groups. Simple problems might not enable the various learning techniques to effectively show how they vary in efficiency.

3. Experimental Design

This section presents the experimental design for studying learning on the ball-weighing problem.

3.1. Design Variables

There were two independent variables in this design. The first one, the experimental *Group*, was a between-subject factor with three levels:

1. Imitation learning group – had access to 5 successful demonstrations of how to solve the target problem (for 5 different ball/weight combinations)
2. Explicit learning group – had access to verbal instructions for solving the problem that they could view once before starting to work on the problem, and study for 10 minutes. Those instructions conveyed the same amount of information as the 5 demos presented to the imitation learning group, and exposure times were matched.
3. Feedback learning group – got feedback on their performance (whether their answers were correct or not).

Note that the imitation learning and explicit learning groups did not receive any feedback.

The second variable, called trial *Quartiles*, was a within-subject factor. A trial is a single problem instance from its initial presentation until the answer is given. Each trial had a different target ball and weight selected at random among the 24 possibilities (12 balls x 2 weights {heavy, light}). Trials were clustered in four quartiles to mark the progression of time within the problem solving session. Besides accommodating the unequal number of completed trials between subjects, this clustering allowed the study of dynamic effects, i.e., how dependent variables evolved over trials.

The design had two dependent variables: elapsed time and accuracy (i.e., whether the answer was correct or not). Both dependent variables were measured on each trial and averaged over trials within each quartile.

3.2. Experimental Hypotheses

Two experimental hypotheses were tested in this experiment. First, the imitation learning and the explicit learning groups were expected to outperform the feedback learning group, both in terms of accuracy (i.e., higher correct answer rate) and speed (i.e., shorter elapsed time per trial). The imitation and explicit learning groups got full information on exactly what to do; in machine learning and connectionist terms this was supervised learning using fully specified target vectors. In contrast, the feedback group only got an impoverished binary signal indicating whether the answer was correct or not. In the machine learning literature, reinforcement learning is considered difficult "because the agent is never told what the right actions are, nor which rewards are due to which actions" [26].

Furthermore, assuming that the verbal instructions were understandable, no differences between the imitation and explicit learning groups were expected because the amount of information given was identical, although presented in a different form.

Second, some kind of learning effect was hypothesized for all groups. Participants were expected to get both more accurate and faster with practice. This effect was expected to be largest in the feedback learning group because, in the absence of any information on correct solutions, subjects were expected to initially spend considerable time exploring solutions of limited efficiency, and thus have more opportunity to gain speed and accuracy though practice.

3.3. Methods and Procedures

Participants were McGill undergraduate and graduate students. Testing 68 participants yielded 63 (17 males and 46 females) usable data samples (21 per experimental group). Participants were excluded when they could not finish the warm-up task within 30 minutes (n=3), or when they were identified as statistical outliers on a q-q plot graph (n=2). Participants were randomly assigned to experimental groups. The chance to win a $50 prize encouraged maximal performance by keeping participants motivated.

A warm-up task (level 1) was first presented (3 balls, 2 uses of scale) to allow participants to become familiar with the task and the user interface.

The target task (12 balls and 3 uses of the scale) was presented as level 2. Upon entering that level, participants were given demonstrations or instructions depending on the condition. They next worked on problem trials for 30 minutes, or until they successfully solved all 24 different trials consecutively. Trials were selected in random order from a list of unsolved trials. When solved successfully, a trial was removed from the list. However, when an error was made, the list was reset back to the whole 24-trial set.

Participants were instructed to label (categorize) balls to reflect the information they gathered about the balls' weights as trial progressed. Each ball could be labelled as follows: "Unknown" (heavy, light or normal weight), "heavy or light", "heavy or normal", "light or normal", "heavy", "light", or "normal".

Figure 1 shows a screenshot of the Java computer program designed to implement the ball-weighing task and record problem solving data for further analysis.

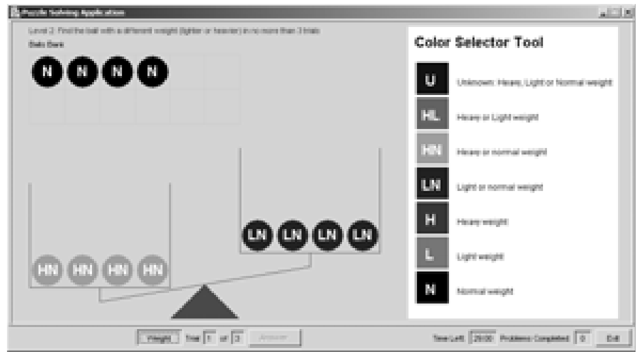


Figure 1. Computer Program Screenshot

Appendix 1 presents a complete set of instructions for solving the problem. Each participant in the explicit learning group read a subset of these instructions. Each subset was designed to provide the same information as a set of five randomly selected demonstrations used for the imitation learning group.

In terms of task analysis, the ultimate goal of this problem is to identify the target ball, which is either heavier or lighter. To do this, the problem solver must loop through a pair of sub-goals until the problem solver finds a solution or exceeds the maximal number of scale uses allowed (solution fails in the latter case). The first sub-goal is to select which balls to weigh in order to maximize the information obtained from the scale. The second sub-goal is to appropriately extract new information acquired using the weight trial and to update categorization of the balls with the appropriate color markings. Appendix 1 presents detailed operations to be performed for all sub-tasks. Typically the operations to perform depend on the result of the previous weighing.

4. Results, Analysis and Discussion

4.1. Correct answer rates

Table 1 presents mean correct answer rates. Values were computed by averaging trial values for all participants and all quartiles. Correct answer rates were not normally distributed and could not be transformed to a normal distribution because of a ceiling value at 1.00. Therefore, Kruskal-Wallis and Median non-parametric tests had to be performed to determine the statistical significance of the group differences in means. Separate analyses had to be done for the effect of group and the effect of quartile because those tests allow only one independent variable to be tested at a time. Note that standard deviation measures were high, but, as we will see, some differences are nonetheless significant.

Table 1. Mean correct rate[1]

Group	Correct Rate	Std Deviation
Feedback learning	0.59	0.49
Imitation learning	0.76	0.42
Explicit learning	0.71	0.45

On the one hand, Kruskal-Wallis and Median tests showed significant differences in correct answer rate across *Group* (Chi-square=7.054, df=2, p=0.029* and Chi-square=7.255, df=2, Median=0.667, p=0.027*, respectively). Pairwise Kruskal-Wallis tests were performed to determine which differences were significant. The results are the following:

1. Feedback vs. Imitation learning groups: Chi-square=5.368, df=1, p=0.021*
2. Feedback vs. Explicit Learning groups: Chi-square=4.793, df=1, p=0.029*
3. Imitation vs. Explicit Learning groups: Chi-square=0.430, df=1, p=0.512

On the other hand, the main effect of *Quartile* was not significant under the Kruskal-Wallis test (Chi-square=3.443, df=3, p=0.328) nor under the Median test (Chi-square=6.43, df=3, p=0.092).

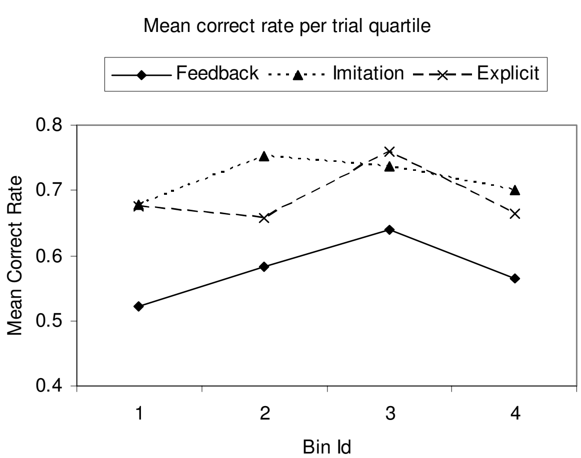


Figure 2. Mean correct rate per trial quartile (i.e., 4 bins) (n=63 participants)

[1] Computed as follows: Total number of correct answers / Total number of completed trials

Figure 2 presents the mean correct answer rate per group and quartile.

The performance decrease in the last quartile might suggest a fatigue effect. However, this effect is not reliable. When trials are grouped into different numbers of clusters, the shape of the distributions varies, as illustrated in figure 3, which presents the mean correct answer rates across 8 bins.

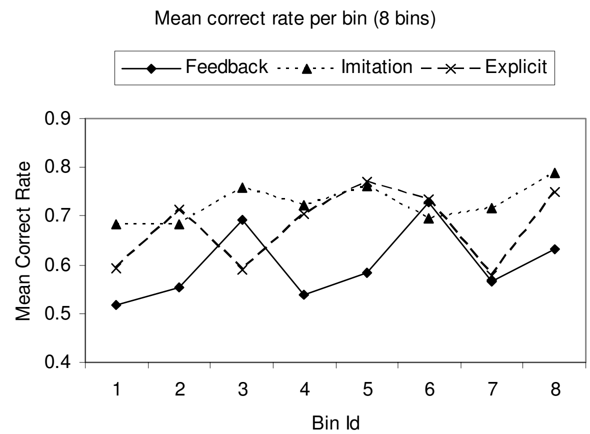


Figure 3. Mean correct rate for 8 bins (n=59 participants[2])

Two conclusions can be drawn. First, the feedback learning group significantly underperformed compared to the other two groups, which did not differ significantly from each other. Second, visual inspection of correct answer rates with sufficient numbers of bins (e.g., figure 3) suggests a small learning effect, which fell below the statistical power available in this experiment perhaps due to an insufficient sample size.

4.2. Contingency tables analysis

Contingency tables 2 and 3 present the number of perfect performers and number of correct and incorrect answers for each group respectively. A perfect performer was defined as a participant who made no errors, and thus completed level 2 in exactly 24 trials.

Table 2. Numbers of perfect and imperfect performers

Group	Subjects	Perfect Score	Imperfect score
Feedback learning	21	0	21
Imitation learning	21	5	16
Explicit learning	21	2	19

[2] Four participants were excluded because they completed fewer than 8 trials.

Table 3. Numbers of correct and incorrect answers

Group	Trials count	Correct	Incorrect
Feedback learning	350	207	143
Imitation learning	209	311	98
Explicit learning	390	276	114

The two Chi-square tests performed on contingency tables of correct answer data were both significant. In other words, Group had a significant effect on the number of perfect performers (Chi-square=6.11, df=2, $p<0.05$) and on the overall number of correct and incorrect answers (Chi-square=21.08, df=2, $p<0.05$). Most of perfect performers and correct answers are found in the Imitation Learning group.

4.3. Elapsed time

Table 4 presents mean elapsed times. Values were computed by averaging trial values for all participants and all quartiles.

Table 4. Mean elapsed times (sec per trial)

Group	Elapsed time	Std Deviation
Feedback learning	10.4	7.9
Imitation learning	8.8	5.1
Explicit learning	9.5	6.1

An ANOVA test was performed across the two independent variables (Group (3 levels) and Quartile (4 levels)) after a log transformation was applied to achieve better normality in elapsed time distributions. Similarly to correct answer rate values, standard deviation measures of elapsed time were high, but nevertheless significant.

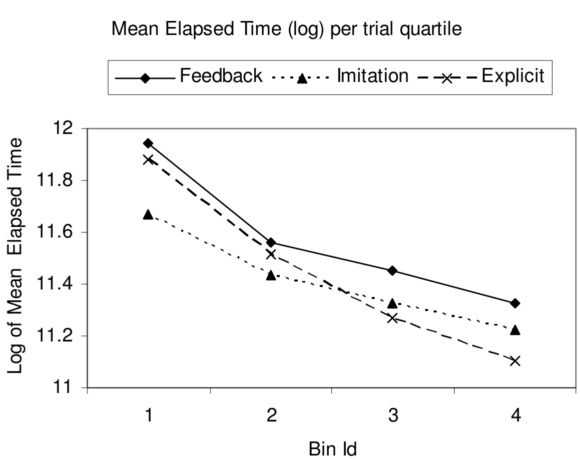


Figure 4. Log of average elapsed time per trial quartile

There was no significant main effect of Group ($F_{2,60}$= 1.42, p= 0.25). However, the main effect of Quartile was highly significant ($F_{3,180}$= 85.7, $p<0.001$) suggesting a learning effect across quartiles within each group. Furthermore, the interaction effect of Group and Quartile was significant ($F_{6,180}$= 2.63, p= 0.018) indicating that groups differ in their decrease in elapsed time across Quartiles. These results suggest different rates of learning, and inspection of figure 4 suggests that the explicit learning group shows the highest speedup. It may have taken time and practice for participants in the explicit learning group to figure out how to effectively make use of the abstract instructions they were given.

4.4. Analysis of Strategies

The use of strategies for selecting balls in the first weighing was investigated. The correct strategy is to weigh four balls on the right side of the scale against four balls on the left side (abbreviated as 4/4 below) because it maximizes the information obtained from the three possible scale outcomes (left heavy, right heavy or equal weights). Besides the 4/4 strategy, two other ones were frequently used: 6/6 (six balls on each side of the scale) and 3/3. Table 5 presents the use of each strategy in each group at the beginning (init) and the end (final) of the 30-minute problem solving session. A strategy was tagged as *n/n* only when it was consistently and exclusively used during the first 20% of trials (init) or the last 20% of trials (final). Unequal ball counts and use of multiple strategies were categorized as Other.

Table 5. Use of initial and final strategies during the problem solving session for all experimental groups

Strategy	Experimental Group					
	Feedback learning		Imitation Learning		Explicit Learning	
	Init	Final	Init	Final	Init	Final
1/1	0.0%	0.8%	0.0%	0.0%	0.0%	0.0%
2/2	1.2%	4.0%	0.0%	0.0%	1.2%	0.0%
3/3	22.9%	21.0%	0.0%	0.0%	10.3%	16.7%
4/4	31.5%	62.3%	100%	99.2%	83.3%	83.3%
5/5	4.8%	0.0%	0.0%	0.0%	2.0%	0.0%
6/6	34.9%	11.9%	0.0%	0.8%	3.2%	0.0%
Other	4.7%	0.0%	0.0%	0.0%	0.0%	0.0%

As Table 5 exhibits, participants in the imitation learning group consistently used the correct strategy (4/4) all through the problem solving session.

Participants in the explicit learning group mainly used the correct strategy from the beginning (about 83%), but also explored other possibilities, suggesting it was difficult for them to map the abstract verbal description into action. Results also suggest that a significant proportion of participants in that group interpreted "use 1/3 of the balls on each side of the scale" as suggesting a 3/3 strategy instead of 4/4. The effect persisted until the end: about 17% of subjects were still using the wrong strategy. This suggests that these verbal rules (see Appendix 1) are more difficult to turn into correct action than a demonstration is, and that although the rules are properly written, they might be misinterpreted.

Finally, participants in the feedback learning group increased their use of the correct strategy by 31%, indicating a learning effect. The 6/6 strategy was the most popular initially, possibly showing a natural (yet incorrect) heuristic bias that testing all the balls gives the most information. The fact that uses of incorrect strategies (3/3 and 6/6) still remains high in the feedback learning group suggests that it is difficult for participants to abandon the use of incorrect strategies. In fact, even incorrect strategies are rewarded because they do yield some correct answers. Many sub-optimal strategies leave two possible answers at the end, which means 50% of correct answers when participants guess between the two. If participants "satisfice" [23] on getting 50 or 60% of correct answers, they may well remain stuck with a sub-optimal strategy. Only an optimal solution strategy, such as the one outlined in Appendix 1, will always yield the correct answer.

5. Conclusions

Based on the non-parametric and ANOVA tests performed on accuracy and speed, our first prediction was partially confirmed: the imitation and the explicit learning groups both outperformed the feedback learning group in terms of accuracy. However, the other part of our hypothesis was not supported: no significant overall difference in speed (response time) was observed across groups.

The difference in accuracy shows that demonstrations and instructions were significantly better sources of information than simple binary feedback, consistent with machine learning theories such as connectionism. The experiment also suggests that demonstrations were more effective than instructions based on the contingency tables analysis of correct answers and perfect performers. However, the difference was not large enough to yield a significant effect under the Kruskal-Wallis and the Median tests. Furthermore, this experiment reminded us that it is challenging to write abstract rules to be used without any external feedback on their correct interpretation or any chance to ask clarification questions. In fact, about 15% of participants misinterpreted the explicit instructions given.

Learning effects were also found in all groups, although we did not find that the feedback learning group was generally slower than the other groups, nor that its speed increased more over trials. If anything, it was observed that the explicit learning group's speed increased most perhaps because it requires practice for participants receiving information in an abstract form to figure out how to use it.

In short, accuracy was affected most by the type of learning, whereas elapsed time was affected most by the number of trials completed (i.e., quartile) and by quartile in interaction with group.

This work represents an initial step towards showing the importance of learning by imitation in problem solving. The next step will consist in determining more precisely which mechanisms underlie improved performance in problem solving tasks. Possible mechanisms are rote memorization of the solution, priming of certain states and operations, and acquisition of a complex hierarchical problem representation. Individual differences are also possible. For example, participants might use different learning mechanisms or even combinations of mechanisms. To explore those potential underlying imitation learning mechanisms, follow-up experiments will be devised. For instance, a group could be presented with a simpler version of the target task (e.g., 9 balls). Because the task demonstrated is not the same as the target 12-ball task, rote memorizing could be discounted as an explanation for improved performance in this group.

Our ultimate goal is to model human imitation learning in problem solving tasks. We plan to apply various learning systems, both symbolic and connectionist to the phenomena observed in the present experiment.

Early simulation results using neural networks suggest that performance in the imitation group cannot be accounted for solely by learning and generalizing the presented demonstrations. It is possible that reasoning and/or prior knowledge could account for differences between the participants in this experiment and these first simulations. We expect that further experiments and simulations will help to elucidate the mechanism underlying imitation learning.

References

[1] A. Alissandrakis, Nehaniv, C. L. and Dautenhahn, K. "Learning by Experience from Others – Social Learning and Imitation in Animals and Robots", Adaptivity and Learning: An interdisciplinary debate, Edited by Kuhn, Menzel, Menzel, Ratsch, Richter and Stamatescu, Springer, Berlin, Germany, 2003.

[2] Anderson, J. R. Learning and memory, an integrated approach. 2nd ed., John Wiley & Sons Inc., New York, NY, 2000.

[3] G. Ashby and Ell, S. "The neurobiology of human category learning", Trends in Cognitive Sciences, Vol. 5, No. 5, May 2001, pp. 204-210.

[4] Bandura, A. Social learning theory, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[5] Benjafield, J. G. Cognition, 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 1997.
[6] C. Breazeal and Scassellati, B. “Robots that imitate humans”, Trends in Cognitive Sciences, Vol. 6, No. 11, 2002, pp. 481-487.
[7] R. W. Byrne and Russon, A. E. “Learning by imitation: a hierarchical approach”, Behavioral and Brain Sciences, Vol. 21, 1998, pp. 667-721.
[8] M. Carpenter, Call, J., and Tomasello, M. “Understanding ‘prior intentions’ enables two-year-olds to imitatively learn a complex task”, Child Development, Vol. 73, 2002, pp. 1431-1441.
[9] G. Gergely, “Rational imitation in preverbal infants”, Nature, Vol. 415, 2002, p. 755.
[10] M. L. Gick and Holyoak, K. J. “Analogical problem solving”, Cognitive Psychology, Vol. 12, 1980, pp. 306-355.
[11] M. L. Gick and Holyoak, K. J. “Schema induction and analogical transfer”, Cognitive Psychology, Vol. 15, 1983, pp. 1-38.
[12] K. J. Holyoak. “Problem solving”. Chapter 8 in Thinking, an invitation to cognitive science 2nd edition, Vol. 3. Edited by Smith and Osherson, MIT Press, Cambridge, MA, 1995.
[13] K. J. Holyoak and Thagard, P. “Analogical mapping by constraint satisfaction”, Cognitive Science, Vol. 13, 1989, pp. 295-355.
[14] C. Holroyd and Coles, M. “The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity”, Psychological Review, Vol. 109, No. 4, 2002, pp. 679–709.
[15] Katona, G. Organizing and memorizing. Columbia University Press, New York, NY, 1940.
[16] Lieberman, D. A. Learning: Behavior and Cognition. 3rd ed, Wadsworth Thomson Learning, Belmont, CA, 2000.
[17] Lock, A. and Colombo, M. “Cognitive abilities in a comparative perspective”, pp. 596-643 in Handbook of human symbolic evolution, Edited by Lock and Peters, 1996.
[18] Meltzoff, A. and Prinz, W. The imitative mind: development, evolution and brain bases. Cambridge University Press, Cambridge, UK, 2002.
[19] R. Nisbett and Wilson, T. “Telling more than we can know: Verbal reports on mental processes”, Psychological Review, Vol. 84, No. 3, 1977, pp. 231-259.
[20] Polya, G. How to solve it 2nd ed., Princeton University Press, Princeton, NJ, 1957.
[21] M. L. Simmel. “The coin problem: a study in thinking”, American Journal of Psychology, Vol. 66, 1953, pp. 229-241.
[22] Simon, H. A. “The structure of ill-structured problems”, Artificial Intelligence, Vol. 4, 1973, pp. 181-202.
[23] Simon, H. A. Models of Man. Wiley, New York, NY, 1957.
[24] Shultz, T. R. Computational developmental psychology, MIT Press, Cambridge, MA, 2003.
[25] Solso, R. Cognitive Psychology, 4th edition, Allyn and Bacon, Boston, MA, 1995.
[26] Russell, S. and Norvig, P. Artificial Intelligence, A Modern Approach, Prentice Hall, Englewood Cliffs, NJ, 1995.
[27] Tomasello, M. The cultural origins of human cognition, Harvard University Press, Cambridge, MA, 1999.
[28] Anzai, K., & Simon, H. A. (1979). The theory of learning by doing. Psychological Review, 86, pp. 124-140.
[29] Scribner, S. & Cole, M. (1973) Cognitive Consequences of Formal and Informal Education. Science, 182, pp. 553 – 559.

Appendix 1 – Symbolic Rules for solving the ball-weighing problem (12 balls variant)

Rules for Selecting Balls for Weighing

- Put 1/3 of the balls on each side of the scale (i.e. 4/4)

1. *if* the scale does not move *then* use 3 “unknown” vs. 3 “normal” balls.
   - *if* the scale moves *then* use 1 “light or normal” vs. 1 “light or normal” ball, *or* 1 “heavy or normal” vs. 1 “heavy or normal” ball
   - *if* the scale does not move *then* use 1 “unknown” vs. 1 “normal” ball
2. *if* the scale moves *then* use 1 “heavy or normal” + 2 “light or normal” vs. 1 “normal” + 1 “heavy or normal” + 1 “light or normal” balls
   - *if* the scale moves *then* use 1 “normal” vs. 1 “heavy or normal” ball, *or* use 1 “normal” vs. 1 “light or normal” ball, *or* use 1 “light or normal” vs. 1 “light or normal” ball
   - *if* the scale does not move *then* use 1 “heavy or normal” vs. 1 “heavy or normal” ball

Rules for Marking Balls

1. *if* the scale does not move *then* all balls on it are “normal”.
2. *if* the scale moves *then* all balls left in the bank are “normal”.
3. *if* there are “unknown” balls located on the side of the scale that moves up *then* those are "light or normal”
4. *if* there are “unknown” balls located on the side of the scale that moves down *then* those are “heavy or normal”.
5. *if* there are "light or normal” balls located on the side of the scale that moves down *then* those are “normal”.
6. *if* there are "heavy or normal” balls located on the side of the scale that moves up *then* those are “normal”.

Caregivers and the Education of the Mirror System

Patricia Zukow-Goldring
Department of Linguistics, University of Southern California
zukow@usc.edu

Abstract

The paper presents an approach to infant cognitive development that emphasizes how caregivers bracket ongoing actions with gestures that direct the child's attention to perceptual information embodied in action sequences as well as the perceivable correspondence between word and referent. Such supervised learning narrows the search space and enhances the speed of achieving a common understanding.

1. Introduction

Arbib's [1] evolutionary framework delineating the neural and functional grounds for modern human language extends and elaborates Rizzolatti & Arbib's "mirror neuron hypothesis [2]. They argued that the brain mechanisms underlying human language abilities evolved from our non-human primate ancestors' ability to link self-generated actions and the similar actions of others. On this view, communicative gestures emerged eventually from a shared understanding that actions one makes oneself are indeed similar to those made by conspecifics. Thus, what the self knows can be enriched by an understanding of the actions and aims of others, and vice versa. From this view, the origins of language reside in behaviors not originally related to communication. That is, this common understanding of action sequences may provide the "missing link" to language.

In answering the question, "What are the sources from outside the self that inform what the child knows?", the basic idea is that a shared understanding of action grounds what individuals know in common. In particular, this perspective roots the ontogeny of language in the progression from action and gesture to speech. What then might the evolutionary path to language and the ontogeny of language in the child have in common? We hope that the examples presented here will engender discussion of how this process might illuminate/relate to language learning in automata. This view characterizes the source of the emergence of language in both as arising from perceiving and acting, leading to gesture, and eventually to speech.

We report here on an ongoing research program designed to investigate how perceiving and acting ground the emergence of language. This effort entails an analysis of the influences of the environment and, in particular, of the ways in which caregivers attune the infants to that environment. Building on what infants might "know" from birth, this work delineates the interplay of perceptual processes with action that might allow them to come to know "what everyone else already knows", including word meaning.

2. Imitation and attention: Affordances and effectivities

The ability to imitate has profound implications for learning and communication as well as playing a crucial role in attempts to build upon the mirror system hypothesis [3,4]. The focus here is to understand how imitation, especially assisted imitation, contributes to communicative development.

Chimpanzees imitate some actions of others in the wild [5], but learn much more complex actions with objects when raised by humans [6]. However, the pace and extent of their imitation is very limited with respect to that of humans. Indeed, the vast majority of human children do imitate, albeit to varying degrees at different ages and for behaviors that differ in modality and complexity of content [7,8].

Non-human primates in the wild learn fewer new skills or behavioral complexity, requiring many more trials or attempts than do human children or enculturated chimps. The former learn in a painstaking manner, engaging in random sequences of behavior by design or naturally, eventually detecting patterns of moves/behaviors that lead to predictable consequences and a basis for repeating them in similar circumstances. In the case of wild chimpanzees, tasks such as learning to crack a nut with a hammer on an anvil takes years to acquire [9]. In contrast, human infants and chimps raised in enriched environments similar to that of human infants learn to use objects far more rapidly.

Byrne [10] has noted that scholars use the term *imitation* in two ways: (1) *the correspondence problem* (how can the child match *observed actions* with *self-executed actions?* [different from the correspondence problem of vision]*)* and (2) *the transfer of skill problem* (how can the child acquire *novel, complex behavior* by observing?). In the main, researchers have focused on this *correspondence problem*, overlooking or subsuming, perhaps, the transfer of skill problem within it. For the sake of argument, assume that if a child knows that she is herself like the other (e.g., the caregiver), perhaps she can learn to do what the other does to achieve similar benefits or avoid risks, given the movement is already part of the child's repertoire. But can a novice spontaneously or after a delay imitate just any novel and/or complex, "developmentally appropriate" behavior not yet in her repertoire without assistance? I argue, probably not. Instead, I suggest that investigating the *transfer of skill problem* may provide answers to the *correspondence problem* as well. That is, grasping the methods that promote the *transfer of skill* may illuminate how, as the child learns new bodily skills, s/he may literally get in touch with both sides of the *correspondence problem*, the match between the other's actions and those of the self. (See [11] for modeling of the development of grasp-related mirror neurons, suggesting how the mirror system grounds imitation as a core component of communicative and linguistic development within an action/perception framework.)

Most research investigating the development and implications of imitation focuses on what the child knows, rather than how the child comes to know. Accounting for these achievements usually takes the form of proposing some combination of cognitive precursors, socio-pragmatic knowledge, or maturing modules hypothesized to be necessary for the activity [12-14]. This literature documents the age at which the average child can observe someone else's action and repeat it accurately either promptly or after a delay – but may underestimate sources of the infants' accomplishments located in the caregiving environment. Informed by an integrative view of action and perception, a somewhat different perspective suggests how imitation may foster the emergence of language.

Greenfield [15] observed that children imitate those actions that are entering their repertoire. Why might these particular actions be ripe for imitation and not others? Are the children's imitations usually autonomous accomplishments or do they have a robust history of assistance from others? According to Piaget [16], only the infant's independent achievements contribute to cognitive development, whereas he referred to the "pedagogical mania" of those who tend children as interfering at best. Indeed, caregivers do invite infants to imitate, but I suggest for the better. On those occasions, caregivers both direct attention [17-21] to aspects of the ongoing events and tutor actions to "achieve consensus" [22-23]. These interactional opportunities give infants crucial practice in (and a refining of) what to notice and do, and when to do it. Further, when demonstrating an activity, the caregiver marks the child's subsequent suitable attempts to imitate with speech and gestures of approval or may elaborate the ongoing activity, whereas repeated and revised messages, dropping the current activity, or remarking on the child's lack of interest follow inadequate responses. These interactions also may be central to communicative development. In particular, engaging in these activities may provide the means to grasp important prerequisites that underlie communicating with language. These basics include knowing that words have an instrumental effect on the receiver of a message [24,25], words refer [21, 26-28], and coparticipants share or negotiate a common understanding of ongoing events [20,21,29,30].

Our normal experience is highly multi-sensory, not restricted to the limited perceptual input of, say, a video clip. Indeed, Stoffregen and Bardy [31] have argued that "multisensory perception is not merely the primary type of perception; it's the only type of perception". Caregivers and children detect “the something that something is happening to” as well as "the something that is happening" through vision, touch, sound, taste, and touch [21,32]. Especially relevant to this idea is the young infant's known ability to detect regularities or invariants in the continuous stream of perceptual information [33].

Gibson [34] proposed the notion of *affordances*, referring to opportunities for action for creatures in their environment. Further, he argued that creatures detect the perceptual structure that specifies the unchanging, invariant aspects of ongoing events as well as the structure specifying transformation and change. Research informed by Gibson's ecological realism emphasizes that the relation between a creature and its environment have consequences for behavior. The classic example is that as we walk across a room we see more and that more that we see tells us which surfaces will support our walking, what objects block our way and so on. Some characterize the relation as emergent properties of the animal-environment system [35], others add the notion of *effectivities* to that of affordances as dual complements [36,37]. That is, an affordance (environmental disposition) takes as its complement an effectivity (acting itself is informed by what the body can do) or vice-versa [38]. Effectivities expand as an individual gains skill participating in new activities and differentiating what the environment affords for action. The question, however, is how does an infant become more adept?

Can infants, novice members of their culture, detect affordances and consummate an activity without assistance, if the action is not present at birth? (See [39] for the development of grasping that is present at birth). What to do with objects, except for the most rudimentary, stereotyped, self-directed actions, such as sucking and grasping, presents a challenge. I have argued that objects can not "tell us" what they afford [21]. Nor can caregivers of young infants *tell them*, as verbal instructions directed to infants before they know "what and that" words mean, prove most ineffective [22,23]. Novices learn affordances as they engage in daily life in a particular time, place, and culture [20,40]. However, even careful nonverbal guidance pointing out aspects of elements, configuring them in time/space, and/or modeling actions may not be sufficient. Shaw [41] has argued that effectivities transform potential experiences into actual ones; that is, *affordances dispose,* while *effectivities deliver* (actualize).

I propose that the child at first lacks the ability to detect by observation alone how the body relates to the physical layout and to the furniture of the world, except for the most rudimentary actions. In this regard, these empirical studies of communicative development (Zukow-Goldring, 1996, 2001) have stressed, in addition, that during interaction, the perceiving and acting of one person continuously affects the perceiving and acting of the other. I argue that caregiver practices guide infants to perceive possibilities for action [21,23,42]. Further, as infants respond to caregiver guidance, the infants' misdoings that may ensue pinpoint possible misperceiving as well as lack of bodily skill that, in turn, inform the caregiver's subsequent seeing and doing.

The point is not to deny that children can learn certain things for themselves by trial-and-error. Clearly the physical environment or layout affects us and we surely affect it continuously, but we should not overlook the person and the social environment. That is, by directing the children's attention to their own effectivities in relation to affordances in the environment, the caregiver greatly narrows the search space for learning, and consequently enhances the speed and extent of learning. Further, these caregiver practices or methods may educate infants to notice that the infant is "like the other" through interactions that explicitly foreground the correspondence between the effectivities of the infant-actor's body and that of the caregiver. In any novice-expert interaction, whether infant and caregiver or student and teacher, the perceiving and acting of one person continuously informs that of other interactional partners.

3. Educating attention: From being a body to becoming a cultural being "like the other"

What do infants have to learn about the world in order to communicate about what's happening? Infants must learn the most basic things (even about), e.g., taking a bath, eating with utensils, walking. During mundane activities, infants must detect and participate in assembling the structure and organization of everyday events before they can communicate with others about these events. Out of the unceasing perceptual flow, which is quite unlike the highly edited cuts of most movies, *caregivers* continuously educate attention to aspects of ongoing events. This assistance guides infants to notice key elements of what persists and what changes. Caregiver gestures make perceptual structure prominent through translational movements that often occlude other information in a scene. In the same vein, placing objects close to the child's face ensures attention and the inescapability of details [21].

Caregivers embody or put infants through the motions of activities as well as direct attention to the similarity of the one's own body to those of others, to the relation of such a body to specific objects and animate beings, and to what these objects and animate beings afford for action and interaction. In contrast, many studies and theories assume that children know and/or learn autonomously how their bodies move in space and in relation to animate and inanimate things [16,43] and thus do not explore what experiences might underlie eventual adept performance. I stress again the role of the caregiver in directing attention to *effectivities* as well as *affordances* – the two sides of the mirror system. These interactive sequences eventually invite imitation.

Caregivers talk about what they are doing as they do it. Often children initially misunderstand these actions and spoken gestures, in part, because words cannot explain unless the child already knows what the words means. Given these circumstances, how is consensus achieved as the child becomes an adept member of the community? My approach integrates discovering the interactional methods or practices that inform perceptual differentiation, the assembling of action sequences, and the detecting of word meaning, despite the fact that many studies of language acquisition assume that gestures entail ambiguity of reference [27,44]. These authors rely on Quine's classic essay [45] in which he discussed the ambiguity of reference entailed in, say, speaking about and pointing to a rabbit. But caregivers tend to focus attention with precision. They do not simply say an unfamiliar word (such as Quine's *gavagai)* while pointing. Instead, caregivers may rub a rabbit's fur while saying, "fur"; trace the topography of its ears while saying, "ear", stroke the entire rabbit or rotate the whole animal when saying, "rabbit", etc. 20,22 ,46]. Successful teaching entails marking the correspondence between what is said and what is happening.

In what follows, I illustrate the findings from a number of studies of infant development with some qualitative examples [20,21,23].

4. The naturalistic experiments

Infants are immersed in talk: some directed to them, some to others, some to prohibit action, some to direct attention to something new or when the child does not understand an utterance. We argue that at early stages of communicative development, learning that words mean and what they mean entails having an embodied understanding of the organization and structure of relevant aspects of daily life. Concomitantly, the child must notice that others mark the relation between what is said and what is happening, and how they do so. To support this view, key results from a series of naturalistic experiments conducted to clarify how children come to comprehend initially misunderstood messages are summarized [22,23]. In these studies the following hypotheses were tested:

1. Providing a child with more perceptual structure will assist caregiver and child to achieve the consensus needed for communication, including, where appropriate, explicit guidance of the child's movements.

2. Additional or more specific verbal information will not enhance understanding when no basis for that understanding has yet been embodied.

The studies reported illuminate how a human child learns about the world. Of course, I do not denying the utility of verbal instruction for older children. Rather, the purpose is to illustrate how the fundamental link between perception and action provides the information upon which communication can build. It is a separate study to understand the later "bootstrapping" that occurs when words can take a far greater role in advancing what the child knows.

5.Qualitative examples: Assisted imitation

The caregiver demonstrates the action(s) and then gives the infant a chance to act. Quite often the infant's attempt is inadequate. Caregivers take care to arrange the physical layout, so that the configuration in space of caregiver, infant, and object(s) makes them suitably aligned in space so that action is within reach. In addition, optimal proximity makes perceptually prominent what the object or some aspect of it affords for action. Frequently the caregiver embodies the infant, so that the child can perceive the relation of his or her body in terms of posture, motor actions, rhythm as it changes over time to accomplish the action or action sequence.

In the following examples, the infant’s unsuccessful attempt at imitation with a toy displays some familiarity with the culturally relevant use of objects, but were initial attempts on the part of the infant to engage in this sequence of activities. Grasping an object is the fulcrum around which novel action grows. The first example focuses less on tutoring effectivities and affordances and more on learning a sequence of actions to consummate an activity (Vibrating Toy). Even though such objects and their uses are not entirely novel, what is required to imitate apparently is. That is, the ability to notice the relevant affordances and coordinate them with particular effectivities that are necessary to accomplish these tasks is not available to the infant without assistance. Their fragmentary, flawed attempts to imitate actions observed in the past elicit very careful and elaborate tutoring on the part of the caregivers to direct attention to relevant affordances and effectivities. Going further, we need to understand how picking-up the perceptual information that the caregiver has picked out allows the infant to get a grip on what to do. Getting a feel for what to do can provide the basis for detecting the affordances that will guide children in their attempts to imitate that action.

5.1. Vibrating toy (14.5 months) - Caregiver and "toy" tutoring of a sequence of actions

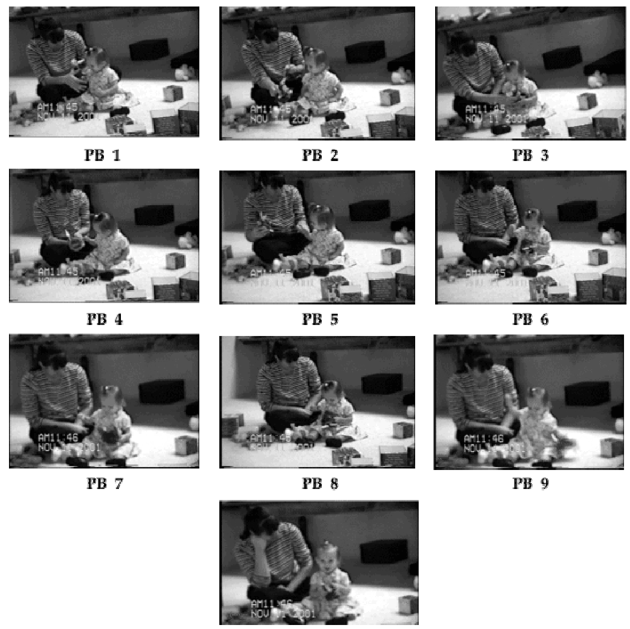


Figure 1. Vibrating toy sequence

This infant, Elsa, and her mother, Kathy, engage in a familiar routine with a reindeer toy that has a hidden affordance, a spring inside the toy to which a string is attached (Figure 1). Family members had played this "game" with Elsa quite frequently during the prior eleven months. However, she had never attempted to imitate the others. In this routine, when the caregiver pulls on a ring that protrudes from the back of the toy, the string unwinds. Releasing the ring/string at the apex of its extension retracts the string so quickly that the toy vibrates strongly accompanied by a loud pulsing noise. Elsa expresses delight when she feels the vibrating toy placed on her stomach. Elsa, however, cannot make the toy vibrate by herself. This "game" entails a sequence of actions: (i) someone grasps the string by the ring, (ii) the string unwinds as it is pulled and (iii) retracts within the toy as the tension on the string lessens. Finally, (iv) someone places the reindeer toy on the infant's stomach. In this example, more emphasis is placed on the sequence of actions than on the briefly noted affordances and effectivities that mother and toy make perceptually available. For conciseness sake, this sequence has been abbreviated by omitting repetitions and variations leading to the child's final adept enactment of this activity.

Elsa, sitting in front of her mother, turns to give her mother, Kathy, the toy that she wants her to animate (Figure 1, R1). Kathy pulls the string out by the ring (R2), releases the string and places the vibrating toy on Elsa's stomach (R3). When Elsa wants her mother to continue, Kathy says, *You do it!,* as she *partially demonstrates* by orienting the back of the toy toward Elsa, making the ring for pulling (affordance) prominent and within the infant's reach. The infant grasps the ring (R4). The mother *embodies* by holding Elsa's hand and the toy steady as she pulls the toy away from them, presumably so the infant can feel the tension (affordance) as the string unwinds (R5). The spring attached to the string *embodies* Elsa by pulling her hand back toward the toy (R6). Subsequently, the infant pulls the string out herself (R7) and holds on as the spring *embodies* by retracting the string back into the toy (R8). At this point, the toy vibrates weakly, if at all. A few seconds later, Elsa pulls the string so quickly and fully out, that the tension on the string *embodies* her by snapping the ring from her fingers. This adept use of her body when pulling the string forcefully (effectivity) allows her quite serendipitously to experience how to take advantage of the vibratory properties of the toy (affordance). Note that her arm recoils from the force moving quickly away from its former position, while the hand holding the strongly vibrating toy moves far in the opposite direction (R9). Within seconds, Elsa first pulls and lets go of the string and then places the base of the toy on her stomach to best perceive the vibrations as she expresses evident joy (R10).

Note the free building up of a sequence, as the child understands each new element. Both caregiver and toy educate Elsa's attention to new affordances and the refining of her actions (effectivities) to fill in the gap between grasping the ring and feeling the vibration of the toy on her stomach. Elsa experiences bodily the tension of the string unwinding as her mother pulls the toy away from her and as the spring hidden in the toy that controls the string retracts as the string appears and disappears. Pulling the string out is within Elsa's grasp. However, the accidental snapping back (letting go) of the string at the apex of its path and at its highest tension made evident the relation between the effectivity of releasing the string and the ensuing affordance of the toy's vibrations. By the second attempt, Elsa had placed the toy to bring the most enjoyment. Within the next several minutes, she changed her grasp from a finger crooked through the ring to a pincer grip, could pull the string with both right and left hands, and attempted to give her doll the same experience. Although she knew "that" the toy held potential for vibrating, she did not know "how" to make it happen until she received very careful tutoring. This educating of attention and action contrasts sharply with the effort it takes to tutor the attention and action of monkeys as well as autistic children (Leo Fogassi, personal communication; [7]).

5.2. Pop beads (13 months): Caregiver tutoring of effectivities and affordances when concatenating beads

Figure 2. Pop beads sequence.

Pop beads, easily graspable by infants and toddlers, have affordances that allow concatenation. Play with this toy consists of (i) orienting toward each other the parts of each bead that afford concatenation (the dual complements of protrusion and opening), (ii) moving the appropriately oriented beads toward each other on a converging path, and (iii) applying enough force when the parts meet to embed the protrusion of one in the opening of the other.

The infant, Angela, begins by pressing a block lacking the appropriate affordances and a pop bead together. She displays an understanding that completing the task requires two small graspable objects and the application of some force to bring them together (Figure 2, PB1). Her behavior provides no evidence that she knows that a set of objects with specific parts must come together, nor that they must sustain an orientation as they meet along a converging path. Her mother, Cecilia, provides perceptual information to Angela, gradually foregrounding the affordances of the objects and the effectivities of the body required to put the beads together. At first, Cecilia provides a bit of both, *point-touching* the opening on one bead (but not the protrusion in the other) as she directs attention to an affordance (PB2) and then reorients another bead (PB3), enacting a movement that aligns the beads on the same converging path. Lacking at this point, however, is information displaying the path itself or the force required to push the converging objects together. After an unsuccessful attempt by Angela, Cecilia *shows* her what the body must do to move the beads along a path with the required orientation as she pushes the protrusion into the opening with appropriate force (PB4). The infant remains unable to put the beads together (PB5) The mother then more elaborately provides perceptual information as she slowly *point-touches* both the affordances of protrusion (PB6) and opening (PB7), followed by *demonstrating* the effectivities required for connecting and then disconnecting of the beads (PB8). As the infant watches more intently, Cecilia eventually invites Angela to imitate, *"¿A ver, tu?/Let's see, you (do it)?*". She assists her daughter's imitation by *partially demonstrating* what to do by orienting a bead opening toward Angela and holding it in a fixed position as she makes the protrusion easy to see on the second (affordances for action) (PB9). Angela moves the bead along the appropriate path (PB10), but she misorients her bead's protrusion from the opening of the other as the beads touch one another (PB11). Cecilia realigns the opening of her bead, making prominent just where Cecilia should push in her bead. As the infant pushes in the protruding end of her bead, the mother pushes from the opposite direction with enough force to link the beads (PB12).

In this case, the caregiver's gestures gradually provided increments in perceptual information that guided the infant to concatenate two objects. Eventually, the caregiver simplified the task by holding an appropriately oriented bead as a fixed target. This assistance allowed the child to bring her slightly misoriented bead (i) along a path toward her mother's (ii). Angela pushed her bead against the other (iii), as her mother subtly reoriented her bead (i) and provided a complimentary push (iii).

Notwithstanding Angela's noteworthy improvement on this occasion, bringing together two hands, each grasping a properly oriented object, was not within her "reach". Nor could she by herself apply enough force to connect her beads. It is possible that embodying Angela, putting her through the motions, might have drawn maximal attention to the coordination of the affordances of the beads and the effectivities of the body required to consummate this activity. The point is that the young child cannot simply observe a complex activity and imitate it. Much tutoring is required to build a repertoire on which “true” imitation (if such exists!) can take place.

6. Discussion

Caregivers establish an understanding of what is happening. They gather and direct attention to perceptual structure that makes prominent the relations among animate beings, objects and their actions. These dynamic relations specify the organization and structure of the most mundane daily activities. Caregivers introduce their infants to new effectivities or bodily capabilities and affordances for action and interaction on a daily basis. They assist them to link sequences of actions that comprise more and more complex activities. As caregivers educate attention, infants gradually learn to perceive, act, and know in culturally relevant ways. Engaging in action sequences with the caregiver may cultivate a precursor to language: negotiating a common understanding of ongoing events.

6.1. Perceiving that the self is "like the other".

In the context of understanding how the methods of the caregiver correspond to the expanding capabilities of the child, note that often developmental researchers and scholars study affective, motor, perceptual and cognitive development separately. Caregivers do not. During the prelinguistic and one-word periods, caregivers prepare infants to imitate by assisting them "to see what to do" before they can "do what they see" others doing. Day-in and day-out, they cultivate imitation within mundane daily activities with gestures. They animate and direct their infants' attention to their own and others' bodily movements as well as making prominent what the environment offers for action. In the process of learning a new skill, especially when embodied or put through the motions by the caregiver, infants directly experience that the self is "like the other" (cf. [47]). Thus, embedded in grasping the *transfer of skill* problem are opportunities to see and feel solutions to the *correspondence* problem, detecting the match between self and other. Humans who eventually learn/understand that the self is "like the other" cultivate abilities in their young that contribute to imitating, tutoring, communicating, and re-presenting events. The mirror system offers a means to clarify in what manner human and nonhuman primates understand what they see other conspecifics and other primates doing, what abilities and perceptual information underlie learning to do what they see others do, and much more.

7. References

[1] Arbib, M. A., 2005, From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics, *Behavioral and Brain Science.*

[2] Rizzolatti G, Arbib MA, 1998, Language within our grasp. *Trends in Neurosciences* 21: 188-194.

[3] Arbib, M. A., 2002, The Mirror System, Imitation, and Evolution of Language. In: Nehaniv, C., Dautenhahn, K., Eds., *Imitation in Animals and Artifacts*. The MIT Press.

[4] Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G, 1999, Cortical mechanisms of human imitation. Science 286: 2526-2528

[5] Greenfield, P. M., Maynard, A. E., Boehm, C., & Schmidtling, E. Y., 2000, Cultural apprenticeship and cultural change: Tool learning and imitation in chimpanzees and humans. In S. T. Parker, J. Langer, & M. L. McKinney (Eds.), *Biology, brains & behavior* (pp. 237-277). Santa Fe: SAR Press.

[6] Tomasello, M., Savage-Rumbaugh, S., & Kruger, A. C., 1993, Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees. *Child Development, 64,* 1688-1705.

[7] Nadel, J., & Butterworth, G., 1999, *Imitation in infancy.* Cambridge: Cambridge University Press.

[8] Eckerman, C. O., 1993, Toddlers' achievement of coordinated action with conspecifics: A dynamic systems perspective. In L. B. Smith & E. Thelen (Eds.), *A dynamic systems approach to development* (pp. 333-357). Cambridge: MIT.

[9] Boesch, C., 1993, Aspects of transmission of tool-use in wild chimpanzees. In K. R. Gibson & T. Ingold (Eds.), *Tools, language, cognition in human evolution* (pp. 171-183). Cambridge: Cambridge University Press.

[10] Byrne, R. W., 2003, Imitation as behaviour parsing. *Phil. Trans. R. Soc. Lond. B., 358, 529-536.*

[11] Oztop, E., and Arbib, M.A., 2002, Schema Design and Implementation of the Grasp-Related Mirror Neuron System, *Biological Cybernetics,* 87:116-140.

[12] Meltzoff & Moore, 1995, Infants' understanding of people and things: From body imitation to folk psychology. In J. Bermúdez, A. J. Marcel, & N. Eilan (Eds.), *The body and the self* (pp. 43-69). Cambridge, MA: MIT Press.

[13] Tomasello, M., Kruger, A. C., & Ratner, H. H., 1993, Cultural learning. *Behavioral and Brain Sciences*, 16, 495-552.

[14] Uzgiris, I., 1991, The social context of imitation. In M. Lewis & and S. Feinman (Eds.), *Social influences and socialization in infancy* (pp. 215-251). New York: Plenum Press.

[15] Greenfield, P.M., 1972, Cross-cultural studies of mother-infant interaction: Toward a structural-functional approach. *Human Development*, 15: 131-138.

[16] Piaget, J., 1962, *Play, dreams, and imitation in childhood*. New York: Norton Library.

[17] Adamson, L., & Bakeman, R., 1984, Mothers' communicative acts: Changes during infancy. *Infant Behavior and Development, 7*, 467-478.

[18] Tomasello, M., 1988, The role of joint attentional processes in early language development. *Language Sciences*, 10:69-88.

[19] Zukow, P. G., 1989, Siblings as effective socializing agents: Evidence from Central Mexico. In P. G. Zukow (Ed.), Sibling interactions across cultures: Theoretical and methodological issues (pp. 79-105). New York: Springer-Verlag.

[20] Zukow, P. G., 1990, Socio-perceptual bases for the emergence of language: An alternative to innatist approaches. In C. Dent, & P. G. Zukow (Eds.), The idea of innateness: Effects on language and communication research. *Developmental Psychobiology,* 23, 705- 726.

[21] Zukow-Goldring, P., 1997, A social ecological realist approach to the emergence of the lexicon: Educating attention to amodal invariants in gesture and speech. In C. Dent-Read & P. Zukow-Goldring (Eds.), *Evolving explanations of development: Ecological approaches to organism-environment systems* (pp. 199-250). Washington, D. C.: American Psychological Association.

[22] Zukow-Goldring, P., 1996, Sensitive caregivers foster the comprehension of speech: When gestures speak louder than words. *Early Development and Parenting*, 5 (4), 195-211.

[23] Zukow-Goldring, P., 2001, Perceiving Referring Actions: Latino and Euro-American Infants and Caregivers Comprehending Speech. In K. L. Nelson, A. Aksu-Koc, & C. Johnson (Eds.), *Children's Language*, Vol. 11. Hillsdale NJ: Erlbaum.

[24] Braunwald, S., 1978, Context, word and meaning: Toward a communicational analysis of lexical acquisition. In A. Lock (Ed.), *Action, gesture, and symbol: The emergence of language* (pp. 285-327). London: Academic Press.

[25] Braunwald, S. R., & Brislin, R. W., 1979, On being understood: the listener's contribution to the toddler's ability to communicate. In P. French (Ed.), *The development of meaning: Pedo-linguistic Series.* Japan:" Bunks Hyonron Press.

[26] Bates, E., 1976, *Language and context: The acquisition of pragmatics.* New York: Academic Press.

[27] Schlesinger, I. M., 1982, *Steps to language: Toward a theory of native language acquisition.* Hillsdale, NJ: Erlbaum.

[28] Zukow-Goldring, P., & Rader, N., 2001, Perceiving referring actions: A commentary on Gogate, Walker-Andrews, & Bahrick. *Developmental Science* 4, 28-30.
[29] Macbeth, D., 1994, Classroom encounters with the unspeakable: "Do you see, Danelle?". *Discourse Processes*, 17:311-335.
[30] Moerman, M., 1988, *Talking culture: Ethnography and conversation analysis*. Philadelphia: University of Pennsylvania Press.
[31] Stoffregen, T. A., 2003, Affordances as properties of the animal-environment system. In K. S. Jones (Ed.), *How shall affordances be refined: Four perspectives.* [Special issue]. *Ecological Psychology, 15,*115-134.
[32] Michaels, C. F., & Carello, C., 1981, *Direct perception*. Englewood Cliffs, N. J.: Prentice-Hall.
[33] Bahrick, L. E., & Pickens, J. N., 1994, Amodal relations: The basis for intermodal perception and learning in infancy. In D. J. Lewkowicz & R. Lickliter (Eds.), *The development of intersensory perception: Comparative perspectives* (pp. 205-233). Hillsdale, NJ: Lawrence Erlbaum Associates.
[34] Gibson, J. J., 1979, *The ecological approach to visual perception.* Boston: Houghton Mifflin.
[35] Stoffregen, T., & Bardy, B., 2001, On specification and the senses. *Behavioral and Brain Sciences*, 24:195-261.
[36] Shaw, R., & Turvey, M., 1981, Coalitions as models of ecosystems: A realist perspective on perceptual organization. In M. Kubovy & J. R. Pomerantz, Eds.), *Perceptual organization* (pp. 343-415). Hillsdale, NJ: Erlbaum.
[37] Turvey, M., Shaw, R., Reed, E., Mace, W., 1981, Ecological laws of perceiving and acting. *Cognition, 9,* 237-304.
[38] Turvey, M., 1992, Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, 4:173-187.
[39] Oztop, E., Bradley, N.S., and Arbib, M.A., 2004, Infant grasp learning: A computational model *Exp. Brain Res.* (in press).
[40] Costall, A., 1995, Socializing affordances. *Theory and Psychology, 5,* 1-27.
[41] Shaw, R., 2001, Processes, acts, and experiences: Three stances on the problem of intentionality. *Ecological*
[42] Zukow-Goldring, P., & Ferko, K. R., 1994, An ecological approach to the emergence of the lexicon: Socializing attention. In V. John-Steiner, C. Panofsky, & L. Smith (Eds.), *Sociocultural approaches to language and literacy: Interactionist perspective* (pp. 170-190). New York: Cambridge University Press.
[43] Thelen, E., & Smith, L., 1994, *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.
[44] Markman, E. M., 1989, *Categorization and naming in children: Problems of induction*. Cambridge, MA: The MIT Press.
[45] Quine, W. V. O., 1960, *Word and object.* New York: Wiley.
[46] Reed, E. S., 1993, The intention to use a specific affordance: A conceptual framework for psychology. In R. H. Wozniak & K. Fischer (Eds.), *Development in context: Acting and thinking in specific environments* (pp. 45-76). Hillsdale, NJ: Erlbaum.
[47] Meltzoff, A. N., and Moore, M. K., 1999, Persons and representation: Why infant imitation is important for theories of human development. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (pp. 9-35). Cambridge, UK: Cambridge University Press.
Psychology, 13, 275-314.

An Emergent Framework for Self-Motivation in Developmental Robotics

James B. Marshall	Douglas Blank	Lisa Meeden
Computer Science	Computer Science	Computer Science
Pomona College	Bryn Mawr College	Swarthmore College
Claremont, CA 91711	Bryn Mawr, PA 19010	Swarthmore, PA 19081
marshall@cs.pomona.edu	dblank@cs.brynmawr.edu	meeden@cs.swarthmore.edu

Abstract

This paper explores a philosophy and connectionist algorithm for creating a long-term, self-motivated developmental robot control system. Self-motivation is viewed as an emergent property arising from two competing pressures: the need to accurately predict the environment while simultaneously wanting to seek out novelty in the environment. These competing internal pressures are designed to drive the system in a manner reminiscent of a co-evolutionary arms race.

1 Introduction

The quest for creating robot control systems that undergo an autonomous and extended developmental learning process was initiated by Weng and his colleagues [14]. In their report, they differentiate the field of developmental robotics from traditional robotics by focusing on *task-independent* learning. Rather than building control systems to perform specific, predefined tasks, developmental robotics seeks to create open-ended learning systems that continually adapt to new problems. A number of robot control architectures have been created using this paradigm [2, 15, 11], many of which involve some form of reinforcement learning. Reinforcement learning is an appealing approach because it provides a method for giving feedback to a developing system without having to specify how to succeed. Instead, the system is simply rewarded or punished, and must determine on its own how to behave so as to maximize its reward.

However, there is no consensus yet about the most appropriate source for the reinforcement signal in a developmental robotics system. The reinforcement could come from an external teacher, from an internal mechanism such as emotion, or from a combination of external and internal sources. For example, the SAIL robot, an early prototype of a developmental learning system, depended on external reinforcement. SAIL could learn to navigate the corridors of a building by being manually pushed by a human teacher, or by having the teacher press the robot's "good" button or "bad" button in response to its behavior [14]. A more recent version of SAIL employs a reinforcement signal that is the weighted sum of both external reinforcement and an internal measure of novelty [9]. The system compares the predicted next state to the actual next state, and if the prediction is incorrect, novelty is considered to be high. The intent of introducing novelty is to model habituation, as when human babies get bored by constant stimulation and are attracted to novel stimuli. In the SAIL system, the external reinforcement is weighted much more strongly than the internal novelty detection. Therefore the external teacher can easily override the internal drive to perceive new things.

Another fruitful area of inspiration for creating general-purpose internal reinforcement signals is the use of artificial emotions [7]. In Gadanho and Hallam's work, a simulated Khepera robot is endowed with a set of homeostatic variables related to energy, pain, and restlessness. The environment contains a set of obstacles and a set of food sources. The robot's energy decreases on every time step, and increases when it visits a food source. The robot's pain increases when it bumps into obstacles and its restlessness increases when it is not moving. These homeostatic variables can serve to positively reinforce behavior that increases energy and negatively reinforce behavior that increases pain or restlessness. Currently, these reinforcement signals are only used to determine when to switch between a set of preprogrammed behaviors. Thus the robot does not develop new behaviors, but simply determines the best way to sequence its innate behaviors.

We believe that a key step in exploring developmental architectures is to focus on internal sources of reinforcement. The learning process should be driven by *self-motivation*, that is, by the system's own internally-generated representations and goals, instead of relying on those provided by a teacher or designer outside the system according to some specific task to be learned. We are interested in creating a general learning architecture with self-motivation at its core, along with the other key processes of *abstraction* and *anticipation* [3]. Abstraction and anticipation are active research areas [10, 5], but self-motivation has not yet received as much attention from the research community. We envision a control system in which abstraction, anticipation, and self-motivation are closely intertwined and develop together from the start within a single unified framework, using both internal and external sources of reinforcement. Such a system would build up abstractions of its experiences over time, guided by its internal motives, while learning to anticipate the effects of its sensorimotor interactions with the environment. Furthermore, a robot capable of learning about its own sensors and effectors as well as its surrounding environment would avoid the problem of anthropomorphic bias, since the robot's knowledge of its inherent capabilities and limitations, having been acquired through firsthand experience, would be directly grounded in sensorimotor perceptions [3].

There is another, perhaps even more important advantage of self-motivated systems. They can exhibit a degree of open-endedness not possible for systems that are designed to learn specific tasks. For example, the human capacity for learning is not only general-purpose and task-independent, but typically continues over a lifetime, becoming progressively more complex and sophisticated in the types of abstractions and behaviors that can be acquired. The learning tasks themselves may change over time, as different circumstances and goals arise, but the impetus to adapt is ever present.

How does this self-driven pressure to learn arise? In our view, it emerges from the interactions of other competing pressures within the system, in a manner reminiscent of a co-evolutionary arms race, in which two co-evolving species continually push each other toward ever greater complexity. For example, such a self-driven system might attempt to predict future states as accurately as possible, while also attempting to seek out unanticipated, novel states. In effect, these two pressures compete directly against one another, since a system able to perfectly predict future states would never encounter any novelty, and a system that regarded everything it saw as new and unexpected would be incapable of predicting anything. However, if these pressures are balanced appropriately, the system might be able to "bootstrap" its way to increasingly sophisticated behaviors and organization. In other words, by seeking out situations with enough novelty to be interesting without being overwhelmingly unpredictable, the system might achieve a kind of temporary "homeostasis" balanced between surprise and predictability. Gradually, the system would gain the upper hand as it learned to anticipate unexpected things better, and its level of "boredom" would increase, in turn pushing it to explore its environment in search of richer, more interesting experiences. On the other hand, too much surprise would cause it to seek out more predictable regions of the environment. The result would be a type of punctuated learning in which the system remains at a given level long enough to master the tasks at hand, before moving on to the next level. Clearly, such a capability would depend on having a robust, general-purpose learning system that could deal with the multitude of different learning tasks that would arise as the system's experiences and behaviors increased in complexity.

This paper takes an incremental approach to the problem of creating a self-motivated developmental system driven by predictability and novelty. As a first step, we propose a connectionist architecture and learning algorithm for implementing self-motivated robot control. A set of experiments is performed on a simulated robot to demonstrate the feasibility of this approach. Next, a detailed examination of the training process for one run on the robot is presented. Finally, the implications of this approach are discussed. It is important to note that the relatively simple environment used in the experiments described here is not rich enough to allow the full realization of higher levels of behavior that such a system should ultimately be capable of developing in principle. However, having shown the viability of this approach under basic conditions, we envision extending it to more realistic and complex environments in future work.

2 Architecture and Algorithm

In this section we propose a neural-network based learning architecture to address these issues, in which discrepancies between the predicted outcomes and the actual outcomes of the robot's actions in its environment serve as the fundamental source of self-motivation, thereby determining what the robot will learn to do. Although this represents an innate bias built into the architecture, it is not task-specific. The hope is that given the right developmental learning algorithm "hard wired" into the system (whether by evolution or engineering), the robot will be able to learn appropriate task-specific behaviors through its own experiences, guided by internally-generated feedback.

Under control of the neural network, the robot generates motor actions to perform, along with predictions of the effects of these actions on its current situation. In our model, situations and predictions consist of simple two-dimensional visual scenes, but other types of sensory representations could be used. After performing an action and observing the results, the robot's prediction is compared with the actual outcome, and a representation of the prediction error is created. This representation forms the basis of a reinforcement training signal for the network, using a version of Complementary Reinforcement Backpropagation (CRBP) [1].

In CRBP, continuous-valued output activations from a network are transformed into binary values stochastically, typically by flipping a biased coin using the output activations as biases. Depending on the particular binary output pattern generated, the network may receive reward or punishment as feedback. In the case of reward, the network's weights are changed using backpropagation with the binary pattern itself as the training target. In the case of punishment, however, the *complement* of the pattern is used. The stochastic nature of CRBP allows the network to learn using only positive or negative feedback signals instead of a fully-supervised training regimen, which is ideal from the point of view of a robot exploring its environment in real time.

In our version of CRBP, the amount of stochastic noise involved in transforming continuous output values into binary can be varied dynamically, under control of the robot itself. We introduce a *computational temperature* parameter τ, ranging from 0 to 100, that controls the amount of noise used in generating motor action vectors and their complements [12]. At low temperature levels, activation values are translated to 0 or 1 nearly deterministically, while at high temperature the translation is nearly random, with 0 or 1 chosen essentially independently of the activation value. At intermediate temperatures, the translation function is a sigmoid curve of the general form $1/(1+e^{-\alpha(x-0.5)})$, with the steepness parameter α of the sigmoid depending on τ. Thus temperature acts as a knob that determines the amount of influence the activation values exert on the translation process, ranging from no influence when $\tau = 100$ to complete determinism when $\tau = 0$.

Given the inherently temporal nature of prediction, we chose to use a Simple Recurrent Network (SRN) architecture [6], shown in Figure 1. There are separate banks of units for representing the robot's motor actions ($\mathcal{M}_{in}$ and $\mathcal{M}_{out}$), sensory state ($\mathcal{S}$), sensory prediction ($\mathcal{P}$), and temporal context ($\mathcal{C}$), with each bank fully connected to the hidden layer. The purpose of the network is twofold: to generate motor actions for controlling the robot, and to generate predictions that in turn guide the training of the network itself. Prediction and control are interleaved during the training process, with different banks of input and output units active at different times. Since the choice of motor action depends on the robot's current sensory state and temporal context, banks $\mathcal{M}_{out}$, $\mathcal{S}$, and $\mathcal{C}$ are active when deciding what to do next, with $\mathcal{M}_{in}$ and $\mathcal{P}$ disabled. Predicting the next state depends on which motor action is performed given the current state and context, so banks $\mathcal{M}_{in}$, $\mathcal{S}$, $\mathcal{C}$, and $\mathcal{P}$ are active during prediction, with $\mathcal{M}_{out}$ disabled. Some weights of the network (namely, those from the state and context banks to the hidden layer) participate in learning both the control and prediction tasks, reflecting their closely intertwined relationship, while others are specific to one task or the other.

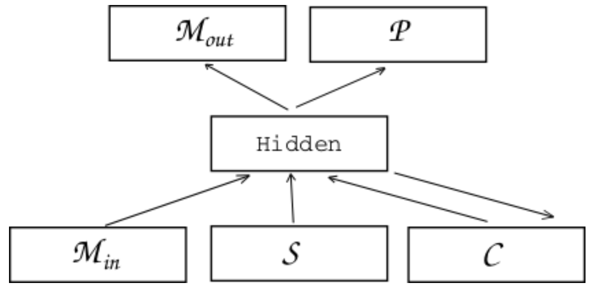


Figure 1. The network architecture

The training algorithm can be understood in terms of three general phases. In the first phase, internal feedback signals are generated from the robot's prediction error. A representation of the prediction error is created based on the discrepancy between the robot's actual observed state and its prediction made on the previous time step, and from this a reinforcement signal is computed, along with a temperature value.

Learning occurs during the second phase. First, the network weights responsible for *motor control* are updated using CRBP, based on the reinforcement signal from phase one. This corresponds to *behavioral learning*, which is driven by discrepancies in the robot's own internally-generated anticipations, rather than by feedback coming directly from the environment or an external teacher. Next, the network weights responsible for *prediction* are updated, using ordinary backpropagation with the robot's actual observed state as the feedback signal. This corresponds to *anticipatory learning*, which is driven by the robot's direct experience in the environment.

In the final control phase, the network generates the next action for the robot to take, as well as a prediction of the outcome of taking that action, and then executes the action.

A more detailed description of the algorithm is given below, outlining the steps performed at time t. At the beginning of Step 1, the following information is known: $\mathcal{M}_{t-1}$ is the motor action performed by the robot on the previous time step; $\mathcal{S}_{t-1}$ is the robot's previous sensory state; $\mathcal{C}_{t-1}$ is its previous temporal context; $\mathcal{P}_{t-1}$ is the prediction, generated at time $t-1$, of the robot's sensory state at time t; and $\mathcal{E}_{t-1}$ is a representation of the prediction error at time $t-1$, based on the discrepancy between $\mathcal{S}_{t-1}$ and $\mathcal{P}_{t-2}$.

- *Generation of internal feedback*
  1. Observe the current sensory state $\mathcal{S}_t$.
  2. Compare $\mathcal{S}_t$ to $\mathcal{P}_{t-1}$ and create a representation of the prediction error $\mathcal{E}_t$.
  3. Compare $\mathcal{E}_t$ to $\mathcal{E}_{t-1}$ and compute a reinforcement signal r of $+1$, -1, or 0, and a temperature τ between 0 and 100.

- *Learning phase*

4. If r is positive, set the motor target $\mathcal{M}_{target}$ to $\mathcal{M}_{t-1}$. If r is negative, set $\mathcal{M}_{target}$ to the complement of $\mathcal{M}_{t-1}$. If r is zero, skip to Step 7.

5. With banks $\mathcal{M}_{in}$ and $\mathcal{P}$ disabled, perform one backpropagation pass with inputs $\mathcal{S}_{t-1}$ and $\mathcal{C}_{t-1}$ on the state and context banks, and $\mathcal{M}_{target}$ on the motor output bank. In the case of positive reinforcement, this makes the network more likely to produce $\mathcal{M}_{t-1}$ given the state and context $\mathcal{S}_{t-1}$ and $\mathcal{C}_{t-1}$. For negative reinforcement, however, the opposite action will be more likely.

6. With bank $\mathcal{M}_{out}$ disabled, perform one backpropagation pass with inputs $\mathcal{M}_{t-1}$, $\mathcal{S}_{t-1}$, and $\mathcal{C}_{t-1}$, and target $\mathcal{S}_t$ on the prediction bank. This makes the network more likely to correctly predict state $\mathcal{S}_t$ when performing motor action $\mathcal{M}_{t-1}$ in state $\mathcal{S}_{t-1}$ with context $\mathcal{C}_{t-1}$. Set $\mathcal{C}_t$ to the hidden layer activation pattern resulting from this step.

- *Control phase*

7. With banks $\mathcal{M}_{in}$ and $\mathcal{P}$ disabled, compute the activation of the output bank $\mathcal{M}_{out}$ using $\mathcal{S}_t$ and $\mathcal{C}_t$ as inputs to the network. Stochastically transform the continuous-valued activations of $\mathcal{M}_{out}$ into a binary motor representation $\mathcal{M}_t$, with the amount of noise determined by τ. This step generates the next motor action for the robot to perform, given its current state and context.

8. With bank $\mathcal{M}_{out}$ disabled, compute the prediction $\mathcal{P}_t$ using $\mathcal{M}_t$, $\mathcal{S}_t$, and $\mathcal{C}_t$ as inputs to the network. This step generates the robot's prediction of the next state given the motor action to perform and its current state and context.

9. Perform action $\mathcal{M}_t$.

10. Set t equal to $t+1$ and go to Step 1.

When training with CRBP, it is often helpful to use a higher learning rate for positive reinforcement than that used for negative reinforcement [1]. A positive reinforcement signal provides evidence that the motor action just performed was a good response to the current situation, so a relatively large weight change helps to increase the likelihood that the robot will take the same action the next time it finds itself in a similar situation. Negative reinforcement, however, suggests only that the motor action was *not* a good thing to do, and offers no guarantee that the opposite action would actually have been better. In this case, using a lower learning rate helps to steer the network away from producing the same response in the future, while remaining somewhat noncommittal about what response the network should actually produce. Thus the learning rate to use in Step 5 above can be set dynamically in Step 4 according to the value of r. In addition, a separate learning rate for prediction may be used in Step 6 if desired.

2.1 State Representation

The above algorithm does not specify exactly how representations of the prediction error $\mathcal{E}_t$ are created in Step 2, or how reinforcement signals are computed from them in Step 3. In fact, the algorithm is fairly general, and does not depend on the particular representation chosen for robot states or motor actions. Furthermore, there is no requirement that robot states must contain purely *sensory* information from the external environment. States could contain additional proprioceptor information, as well as explicit representations of more abstract information generated internally by the robot, such as the prediction error itself.

In our current model, a state $\mathcal{S}_t$ is represented as a 40 $\times$ 10 grayscale image of intensity values normalized to the range 0–1, generated from a simulated blob vision camera. Prediction error $\mathcal{E}_t$ is represented as a 40 $\times$ 10 map of the error values obtained in Step 2 by subtracting the corresponding image values of $\mathcal{S}_t$ and $\mathcal{P}_{t-1}$, and normalizing to 0–1.

2.2 Internal Reinforcement Signal

To compute the reinforcement signal in Step 3, we first compute the "center of mass" coordinate, called the *error centroid*, for each two-dimensional error map $\mathcal{E}_{t-1}$ and $\mathcal{E}_t$. This coordinate is simply the weighted average of the two-dimensional coordinates of all 40 $\times$ 10 error values, weighted by the size of the error. In our experiments, we have used a binary weighting function in which the weight of the error is 1 if the observed value is significantly greater than the predicted value at that point in the map, or 0 otherwise. Other mapping functions are of course possible, such as weighting a value by the magnitude of the error. To compute the reinforcement, the error centroids of $\mathcal{E}_{t-1}$ and $\mathcal{E}_t$ are compared. If the centroid has moved *closer* to the center of the error map from time step $t-1$ to t, the reinforcement is positive; if the centroid has moved *away* from the center, the reinforcement is negative; otherwise it is zero.

The temperature is also updated on the basis of the prediction error. Recall that the temperature ranges from 0 (deterministic) to 100 (random). Currently there are only two cases when the temperature is not set to 0. The first is when there is no error centroid, which corresponds to perfect prediction. In this case, the temperature is set to 100 to induce exploration. The second is when the error centroid has remained stable between two successive steps, but is still not centered. In this case, the temperature is set to 50.

This method of computing the reinforcement signal represents a built-in bias of the system. This can be thought of as an innate tendency of the robot to want to attend to regions of unanticipated activity in the visual field by moving them to the center of view. It is important to note, however, that the reinforcement signal is not based directly on visual input from the environment; rather, it is based on the robot's own *expectations* of what it will see as a result of responding to its current situation. The training of the network is driven by this internally-generated error information rather than by externally-generated visual information.

2.3 Motor Representation

A binary representation for motor actions is necessary in order to allow CRBP to be used for the training of the network's motor responses. In Step 7 above, the continuous-valued activations of the $\mathcal{M}_{out}$ units are transformed into a binary vector $\mathcal{M}_t$. By injecting stochastic noise into this process, the network gains the ability to nondeterministically explore its weight space. This is especially important in the case of negative reinforcement, in which the optimal training target is unknown.

In the experiments described below, we used a simulated robot with only one degree of freedom of movement. The position of the robot was fixed at the center of its environment, with only its angular orientation allowed to change. We chose an 8-bit representation for the motor actions, where the number of ones in a pattern specified the robot's rotation speed and direction, allowing 9 distinct actions to be encoded. The order of the bits was irrelevant. For example, all-zeros represented turning left quickly, all-ones represented turning right quickly, and an equal number of ones and zeros caused the robot to stop. Many different patterns, therefore, were potentially available for the network to use in representing a particular motor action, which gave the robot more flexibility in learning to generate its motor responses. Accordingly, the $\mathcal{M}_{out}$ bank in Figure 1 contained eight units. However, when a motor action is presented to the network as input, it is first translated back into a continuous-valued scalar in the range 0–1, in order to make learning easier for the network. The $\mathcal{M}_{in}$ bank thus consisted of only a single unit.

3 Experiments

To test the architecture and the training algorithm, we created a simple environment in which the developing robot is fixed at the center of a circular arena and can rotate in order to observe its world. Also in the environment is a moving "target" robot controlled by an innate obstacle-avoidance behavior (see Figure 2). In some experiments,

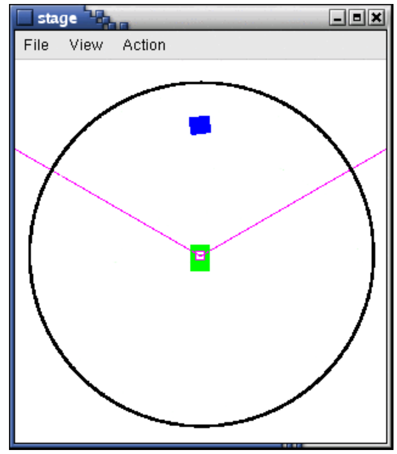


Figure 2. View of the training arena

an additional stationary "decoy" robot was also present, in order to create a slightly more complicated environment.

The goal of these experiments is to induce the developing robot to attend to the target robot by tracking its motion. Clearly it should be possible to learn tracking by providing an external reinforcement signal that is based on whether the target robot is centered in the developing robot's visual field. However, the more interesting issue is whether the developing robot can learn to track given only an internal reinforcement signal based on the error of its own predictions. In this case the external reinforcement signal is directly related to the task of tracking, while the internal reinforcement signal is more indirect. In the following experiments we compare the performance of a developing robot when using external and internal reinforcement signals. The performance measure is based on the average offset of the moving target robot from the center of the developing robot's visual field.

The experiments were conducted using the Stage mobile robot simulator [8], where the developing robot was a simulated ActivMedia Pioneer 2 with a camera. The simulated camera had a 120-degree viewing angle centered on the front of the robot (indicated by the straight lines in Figure 2). Although the Stage simulator does not have simulated pixel-based camera output, we transformed Stage's "blob" data into a 40×10 grayscale image. When the target robot was in view, approximately 16 pixels (4% of the total image) were affected. The robot could turn to the left or right using one of 9 possible rotation speeds, as described earlier in section 2.3.

Using the robotics programming environment Pyro [4], we constructed the neural network shown in Figure 1, where the input layer had 1 motor-in unit, 400 state units, and 30 context units, the hidden layer had 30 units, and the output layer had 8 motor-out units and 400 prediction units. Using Pyro, the network was trained with the three-phase proce-

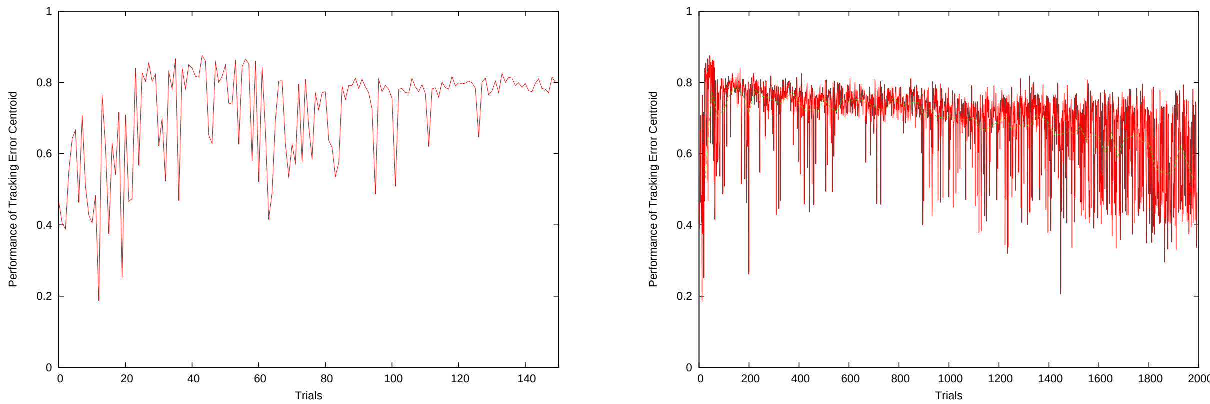


Figure 3. Performance of error centroid tracking: first 150 training trials (left); all trials (right)

dure described in section 2.

The target robot roamed around the inside circumference of the circular wall. At the beginning of each training trial, the target was positioned on the north side of the circle facing west. It then traveled to the left for several hundred time steps, following the circular wall as it went. When it reached the south side of the arena, it was repositioned at the starting point, but this time facing east. The target robot then traveled along the wall to the right, until again it reached a point approximately due south of the starting point. The purpose of this two-legged journey was to ensure that leftward and rightward motion was represented equally during training. The combined westward and eastward journey of the target robot constituted one training trial for the developing robot. Furthermore, whenever the target robot was repositioned at the north side of the arena, the activations of all of the network's context units $\mathcal{C}$ were reinitialized to 0.5. This occurred at the beginning and the middle of each training trial.

In our first experiment, which served as a basic benchmark, the external reinforcement signal was based on the *visual* centroid of the camera image. The robot received positive reinforcement if the visual centroid moved toward the center of the visual field, and negative feedback if it moved away. If the target robot was not in view, no learning was performed. We ran this experiment with computational temperature turned off (*i.e.*, set to 0) in order to see how well the robot could learn in the absence of noise. All of the runs attained a high level of performance within 10 training trials. The network architecture and training procedure enabled the robot to learn to track the target easily.

Of course, our real interest was in seeing if the robot could learn this task indirectly, by using its internally-generated prediction error in place of the actual visual input (as described earlier in section 2.2). As it turned out, using the internal reinforcement signal required that computational temperature be turned on in order for learning to be successful. Although the learning process was slower, the robot was still able to learn to track the moving target robot, even with a stationary decoy robot present in the environment. The next section examines in detail one successful run of this second experiment.

4 Analysis of a Training Run

This run is representative of those that learned to track the moving target robot using only the internally-generated reinforcement signal based on movement of the error centroid. As can be seen in Figure 3 (left), initial performance was about 0.50, but quickly rose to above 0.80 within the first 40 trials. On trial 44 the performance of the network reached its peak, around 0.87. For comparison, we hand-coded a robot to perform the visual tracking task as well as possible, and it scored 0.92. A perfect score of 1.0 is unattainable due to the system's inability to maintain the centroid in the exact center of view at all times.

Recall that our system is designed to perform two conflicting tasks: to accurately predict the next state $\mathcal{P}_{t+1}$, but also to track the areas of its visual field where it cannot predict. Not surprisingly, the better the system is able to predict, the less it is able to track, resulting in a lower performance measure. From these competing goals, three recognizable phases emerge: an early phase (around trials 0 to 35) where the performance on tracking the moving target robot increases; a middle phase where the peak performance is attained (around trials 35 to 60); and a late phase in which tracking performance slowly declines (trials 60 and greater).

Figure 4 shows representative camera images and prediction error data from the middle and late phases of this run. Each column labeled *Camera* shows a sequence of four camera images, with time running from top to bottom. The target robot can be seen as a square of gray pixels near the center of the visual field. The prediction error associated

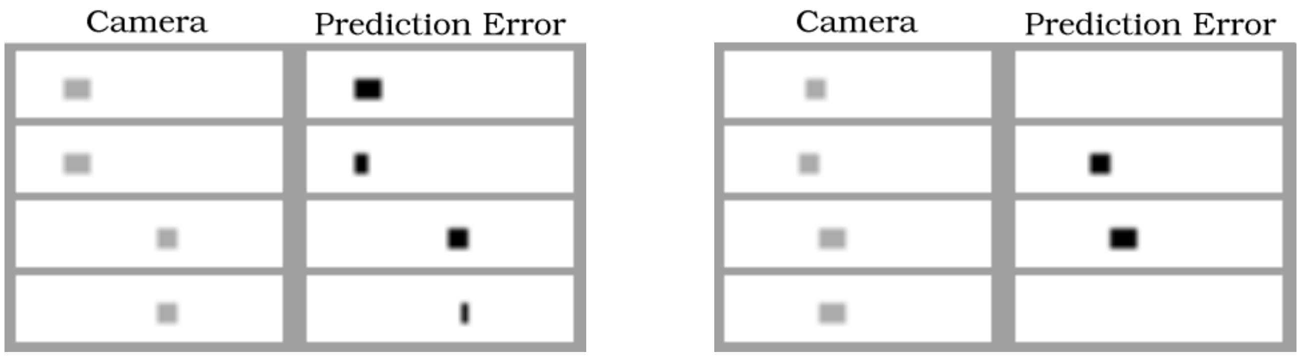


Figure 4. Sample camera images and prediction error data from the middle phase of learning (left) and the late phase of learning (right)

with each camera image is shown to its right. The black pixels indicate where the errors occurred on the prediction bank $\mathcal{P}$ at that step during training. Notice that some of the prediction error regions are smaller than the associated regions from the camera image. This indicates that the system has begun to make some accurate predictions. The system received negative feedback between the first and the second rows and again between the third and fourth rows (since the error centroids have moved slightly farther away from the center). Between the second and third rows, the network was rewarded, since the centroid moved toward the center of the field.

For the camera images and prediction error in the late phase of training, the most noticeable feature is that in the first and fourth rows, there is no error in prediction. This resulted in reward between the first and second rows, and also between the second and third rows (as the centroid gets closer to the center). However, the system was again punished between the third and fourth rows as it "lost" the error centroid.

Further examination of the tracking performance during the late phase shows that it continues to fall until the end of the run at trial 2000. Figure 3 (right) shows the steady decline in performance and an increasing range of performance variability. To understand this behavior better, let us look more closely at how the prediction error evolves over time.

Figure 5 shows that prediction accuracy climbs over the span of 2000 trials, albeit very slowly and also with increasing variability. Indeed, as performance continues to increase in the late stage, the robot encounters fewer views containing any error at all, for which it is then punished. It is in this stage that the competing pressures discussed earlier are most apparent. If the experimental environment had been richer and more varied, after the developing robot had learned tracking, it would likely have been driven by its prediction error to focus on a new aspect of its world.

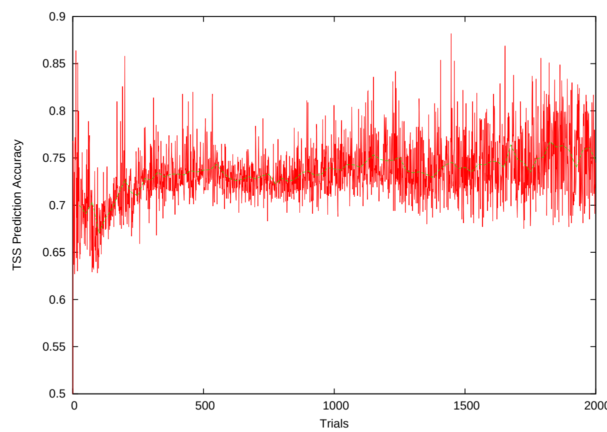


Figure 5. Prediction accuracy over all trials

5 Discussion and Conclusion

The defining characteristic of a developmental robotics architecture is task-independence. A developmental system must be open-ended and capable of finding interesting phenomena to focus on and learn about. The previous experiment suggests that a very general internal mechanism, such as an error centroid created from the robot's own predictions, can serve as a successful reinforcement signal for a developmental connectionist architecture. This initial experiment provides a benchmark for what a self-motivated learner can achieve with limited sensory capabilities in a simple environment. However, the idea of using prediction error as a reinforcer is so general that this same mechanism should be capable of providing a useful reinforcement signal for other sensory modalities and more complex environments.

This paper has outlined a philosophy for designing systems with self-motivation. We believe that self-motivation is an emergent property generated by the competing pressures that arise in attempting to balance predictability and novelty. In the current work, we have proposed a simple recurrent network architecture and algorithm in which systems for learning prediction and control are closely intertwined. The prediction and control pathways within the network share connection weights from the state and context units to the hidden units, and are trained in an interleaved fashion. One system attempts to make predictions of future sensory experiences while the other uses a reinforcement signal based on error provided by the first to drive control. Previous research has shown that prediction learning is facilitated within a system when control and prediction share pathways and when the control signals are internally generated, but not when the control signals come from an outside teacher [13].

In our model, as the predictive system becomes better at anticipating the consequences of the control system's actions, novelty decreases, and the behavior of the predictive system becomes more tightly coupled to the behavior of the control system. As novelty decreases, the error map generated by the predictive system becomes smaller and more fragmented, which may cause the error centroid to jump around at random or disappear entirely. The control system thus has a harder time attending to novel parts of the sensory input. As the control system's performance declines, the robot appears to "lose interest" in those aspects of the sensory input that had previously captured its attention. The coupling between the predictive and control systems therefore begins to weaken, since the control system is no longer reliably paying attention to what it had before. As the predictive system loses its ability to reliably predict the responses of the control system, novelty once again begins to increase. At this point, the novelty of some other stimulus may begin to attract the system's attention (although in our experiment the developing robot never found another focus of attention). We believe that this scenario could potentially serve as a model of habituation. More generally, the interplay between predictability and novelty in our view provides a rich framework for exploring open-ended learning and skill acquisition in developmental robotics.

Acknowledgements

We would like to thank Deepak Kumar, Paul Grobstein, Chris Prince, and the members of the Emergent Phenomena Research Group at Bryn Mawr for engaging discussions on this topic. We would also like to thank the anonymous reviewers whose suggestions helped to improve the paper.

References

[1] D. H. Ackley and M. L. Littman. Generalization and scaling in reinforcement learning. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems 2*, pages 550–557. Morgan Kaufmann, San Mateo, CA, 1990.

[2] N. Almassy, G. M. Edelman, and O. Sporns. Behavioral constraints in the development of neuronal properties: A cortical model embedded in a real-world device. *Cerebral Cortex*, 8:346, 1998.

[3] D. Blank, D. Kumar, L. Meeden, and J. Marshall. Bringing up robot: Fundamental mechanisms for creating a self-motivated, self-organizing architecture. *Cybernetics and Systems*, 36(2), 2005.

[4] D. Blank, L. Meeden, and D. Kumar. Python robotics: An environment for exploring robotics beyond LEGOs. In *ACM Special Interest Group on Computer Science Education Conference*, 2003.

[5] M. V. Butz, O. Sigaud, and P. Gerard. Internal models and anticipations in adaptive learning systems. In *Proceedings of the Workshop on Adaptive Behavior in Anticipatory Learning Systems*, pages 1–23, 2002.

[6] J. L. Elman. Finding structure in time. *Cognitive Science*, 14(2):179–211, 1990.

[7] S. C. Gadanho and J. Hallam. Exploring the role of emotions in autonomous robot learning. In *Proceedings of the AAAI Fall Symposium on emotional intelligence*, pages 84–89. AAAI Press, 1998.

[8] B. Gerkey, R. Vaughan, and A. Howard. The Player/Stage project: Tools for multi-robot and distributed sensor systems. In *Proceedings of the 11th International Conference on Advanced Robotics*, pages 317–323, Coimbra, Portugal, June 2003.

[9] X. Huang and J. Weng. Novelty and reinforcement learning in the value system of developmental robots. In *Proceedings of the 2nd International Workshop on Epigenetic Robotics*, volume 94, pages 47–55. Lund, Sweden: Lund University Cognitive Studies, 2002.

[10] B. Kuipers and P. Beeson. Toward bootstrap learning for place recognition. In S. Coradeschi and A. Saffiotti, editors, *Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems, 2001 AAAI Fall Symposium*, number 01-01 in FS, pages 25–30. AAAI Press, 2001.

[11] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini. Developmental robotics: A survey. *Connection Science*, 15(4):151–190, 2003.

[12] M. Mitchell and D. R. Hofstadter. The role of computational temperature in a computer model of concepts and analogy-making. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates, 1989.

[13] D. Parisi, F. Cecconi, and S. Nolfi. Econets: Neural networks that learn in an environment. *Network*, 1:149–168, 1990.

[14] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291:599–600, 2001.

[15] J. Weng and Y. Zhang. Developmental robotics – a new paradigm. In *Proceedings of the 2nd International Workshop on Epigenetic Robotics*, volume 94, pages 163–174. Lund, Sweden: Lund University Cognitive Studies, 2002.

Intrinsically Motivated Learning of Hierarchical Collections of Skills

Andrew G. Barto
Department of Computer Science
University of Massachusetts
Amherst MA
barto@cs.umass.edu

Satinder Singh
Computer Science and Engineering
University of Michigan
Ann Arbor MI
baveja@umich.edu

Nuttapong Chentanez
Electrical Engineering and Computer Science
University of Michigan
Ann Arbor MI
nchentan@umich.edu

Abstract

Humans and other animals often engage in activities for their own sakes rather than as steps toward solving practical problems. Psychologists call these intrinsically motivated behaviors. What we learn during intrinsically motivated behavior is essential for our development as competent autonomous entities able to efficiently solve a wide range of practical problems as they arise. In this paper we present initial results from a computational study of intrinsically motivated learning aimed at allowing artificial agents to construct and extend hierarchies of reusable skills that are needed for competent autonomy. At the core of the model are recent theoretical and algorithmic advances in computational reinforcement learning, specifically, new concepts related to skills and new learning algorithms for learning with skill hierarchies.

1. Introduction

Despite impressive power and utility, today's machine learning algorithms fall far short of the possibilities for machine learning. They are typically applied to single, isolated problems for each of which they have to be hand-tuned and for which training data sets have to be carefully prepared. They do not have the generative capacity required to significantly extend their abilities beyond initially built-in representations. They do not address many of the reasons that learning is so useful in allowing animals to cope flexibly with new problems as they arise over extended periods of time. Numerous researchers have persuasively argued that a developmental approach is necessary to address these shortcomings (e.g., [31]), drawing from cognitive science, neuroscience, artificial intelligence, and philosophy. According to this approach, an agent undergoes an extended developmental period during which collections of reusable skills are autonomously learned that will be useful for a wide range of later challenges.

Although these arguments are compelling, developmental approaches to artificial agent design have been slow to penetrate the mainstream of the machine learning community. Implementations remain largely exploratory, and they have not yet led to the kind of mathematical formulation required to engage the largest part of the machine learning community. This paper presents preliminary work from a long-term project that seeks to address these shortcomings by elaborating the well-developed computational reinforcement learning (RL) framework [28] to encompass the autonomous development of skill hierarchies through *intrinsically motivated learning*. An agent's activity is said to be intrinsically motivated if the agent engages in it for its own sake rather than as a step toward solving a specific problem.

Our approach builds on existing research in machine learning, with input from recent advances in the neuroscience of brain reward systems as well as classical and contemporary psychological theories of motivation. Not all of our ideas are new, having antecedents in many different areas, including some in machine learning and RL as we outline below. However, we argue that *recent theoretical and computational advances in RL provide important components for making these ideas work efficiently in artificial agents.*

2. Background

Psychologists distinguish between *extrinsic motivation*, which means being moved to do something because of some specific rewarding outcome, and *intrinsic motivation*, which refers to being moved to do something because it is inherently enjoyable. Intrinsic motivation leads organisms to engage in exploration, play, and other behavior driven by curiosity in the absence of explicit reward. In a classic paper, White [32] argued that intrinsically motivated behavior is essential for an organism to gain the competence necessary for autonomy. A system that is competent in this sense has *a broad set of reusable skills* for controlling its environment. The activity through which these broad skills are learned is motivated by an intrinsic reward system that favors the development of broad competence rather than being directed to more specific externally-directed goals. But these skills act as the "building blocks" out of which an agent can form solutions to specific problems that arise over its lifetime. Instead of facing each new challenge by trying to create a solution out of low-level primitives, it can focus on combining and adjusting higher-level skills, greatly increasing the efficiency of learning to solve new problems.

Psychology—A large collection of psychological literature inspires our approach. In 1959 White [32] influentially reviewed the evidence that the (even then) classical Hullian view of motivation in terms of reducing drives related to the biologically primary needs for food, water, sex, and escape was not sufficient to account for an animal's exploratory behavior. Ample evidence existed—and has been greatly augmented since then—that the opportunity to explore a novel environment can itself act as reward. Moreover, not only exploration incited by novelty, but also manipulation, or just activity itself, can be rewarding. This is supported by experimental evidence showing that these activities are not always secondary reinforcers: their motivational significance is built-in rather than being acquired through association with a standard primary reinforcer. The modern expression of these views is most clearly seen in developmental and educational psychology, where a distinction is drawn between intrinsic and extrinsic motivation [5].

The psychology literature is less helpful in specifying the concrete properties of experience that incite intrinsically motivated behavior, although there have been many hypotheses. Berlyne [2] probably had the most to say on these issues, suggesting that the factors underlying intrinsic motivational effects involve novelty, surprise, incongruity, and complexity. He also hypothesized that moderate levels of novelty have the highest hedonic value because the rewarding effect of novelty is overtaken by an aversive effect as novelty increases. This is consistent with many other views holding that situations intermediate between complete familiarity (boredom) and complete unfamiliarity (confusion) have the most hedonic value. Another hypothesis about what we find satisfying in exploration and manipulation is that we enjoy "being a cause" [9], which is a major component of Piaget's theory of child development [20]. In this paper, we use only the degree of surprise of salient stimuli as intrinsic reward, but this is merely a starting point.

Neuroscience—The neuromodulator dopamine has long been associated with reward learning and rewarded behavior, partly because of clear evidence of its key role in drugs of addiction [6]. The original observation [12, 8, 18, 26] that the activity of dopamine cells in the monkey midbrain in reward-learning tasks closely follows the form of a key training signal in RL (the temporal difference prediction error) is an important backdrop for our approach.

Recent studies [15, 3] have focused on the idea that dopamine not only plays a critical role in the extrinsic motivational control of behaviors aimed at harvesting explicit rewards, but also in the intrinsic motivational control of behaviors associated with novelty and exploration. For instance, salient, novel sensory stimuli inspire the same sort of phasic activity of dopamine cells as unpredicted rewards [25, 11]. However, this activation extinguishes more or less quickly as the stimuli become familiar. This may underlie the fact that novelty itself has rewarding characteristics [21]. Theoretical treatments [14, 15] have directly related dopamine activity with mechanisms for controlling exploration in RL such as exploration and shaping bonuses [27, 4, 19]. Although space here does not permit development of these connections, they form key components of our approach to intrinsically motivated RL.

Computational Models of Intrinsic Motivation—Although there have been previous computational studies related to intrinsic motivation, most relevant is recent work from the epigenetic robotics community, some of which discusses the important role of novelty and curiosity in intelligent behavior (e.g., [13, 16]). However, this work does not build upon the mathematical framework of RL and does not use the recently-developed RL methods that we employ. Closely related RL research is that of Schmidhuber (e.g., [23, 24]) on curiosity and exploration. While some promising initial results were demonstrated, this work was left in a very preliminary state, and it also predates the new RL methods that we use.

Interestingly, the most closely related recent computational work comes from the field of architecture and design. In a study of artificial creativity, Saunder's recent thesis [22] presents a system that includes intrinsic motivation based on novelty and surprise following Berlyne's [2] theories. We find this work inspiring, though it focuses on searching design spaces rather than the development of reusable sequential skills.

3. Intrinsic Motivation in Reinforcement Learning

RL is a very active area of machine learning, with considerable attention also being received from decision theory, operations research, and control engineering. RL algorithms address the problem of how a behaving agent can learn to approximate an optimal behavioral strategy, usually called a *policy*, while interacting directly with its environment. In the terms of control engineering, RL consists of methods for the on-line approximation of closed-loop solutions to stochastic optimal control problems, usually under conditions of incomplete knowledge of the system being controlled. One can think of a problem's optimality criterion as defining a primary reward function, and one can think of an approximate solution as the skill of expertly controlling the given system according to this optimality criterion.

In what follows, we describe the elements of the standard RL framework that our approach builds upon, and then we describe a preliminary simulation we have produced that shows how these elements can be exploited for intrinsically motivated learning.

Internal and External Environments—According to the "standard" view of RL (e.g., [28]) the agent-environment interaction is envisioned as the classical interaction between a controller (the agent) and the controlled system (the environment), with a specialized reward signal coming from the environment to the agent that provides at each moment of time an evaluation (usually with a scalar reward value) of the agent's ongoing behavior. The component of the environment that provides this evaluation is usually called the "critic" (Fig. 1A). The agent learns to improve its skill in controlling the environment in the sense of learning how to increase the total amount of reward it receives over time from the critic. With appropriate mathematical assumptions, the problem faced by the learning agent is that of approximating an optimal policy for a Markov Decision Process (MDP).

Sutton and Barto [28] carefully point out that the scheme in Fig. 1A is quite abstract and that one should not identify this RL agent with an entire animal or robot. An animal's reward signals are determined by processes within its brain that monitor not only external events through exteroceptive systems but also the animal's internal state, which includes information pertaining to critical system variables (e.g., blood-sugar level) as well as memories and accumulated knowledge. The critic is in an animal's head. Fig. 1B makes this more explicit by "factoring" the environment of Fig. 1A into an *external environment* and an *internal environment*, the later of which contains the critic which determines primary reward. Notice that this scheme still includes cases in which reward can be thought of as an external stimulus (e.g., a pat on the head or a word of praise). These are simply stimuli transduced by the internal environment so as to generate the appropriate level of primary reward.

Because Fig. 1B is a refinement of Fig. 1A (that is, it is the result of adding structure rather than changing it), the standard RL framework already encompasses intrinsic reward. In fact, according to this model, *all* reward is intrinsic, and what psychologists would call extrinsic reward is just intrinsic reward that is directly triggered by external events. But the point of departure for our approach is to note that the internal environment contains, among other things, the organism's motivational system, *which needs to be a sophisticated system that should not have to be redesigned for different problems*. In contrast, the usual practice in applying RL algorithms is to formulate the problem one wants the agent to learn how to solve (e.g., win at backgammon) and define a reward function specially tailored for this problem (e.g., reward = 1 on a win, reward = 0 on a loss). Sometimes considerable ingenuity is required to craft an appropriate reward function. In effect, a different special-purpose motivational system is hand-crafted for each new problem. This should be largely unnecessary.

Skills—Autonomous mental development should result in a collection of reusable skills. But what do we mean by a skill? Recent RL research provides a concrete answer to this question, together with a set of algorithms capable of improving skills with experience. To combat the complexity of learning in difficult domains, RL researchers have turned to principled ways of exploiting "temporal abstraction," where decisions are not required at each step, but rather where each decision invokes the execution of a temporally-extended activity which follows its own closed-loop policy until termination. Substantial theory exists on how to plan and learn when temporally-extended skills are added to the set of actions available to an agent. Since a skill can invoke other skills as components, hierarchical control architectures and learning algorithms naturally emerge from this conception of a skill. Specifically, our approach builds on the theory of *options* [29].

Briefly, an option is something like a subroutine. It consists of 1) an *option policy* that directs the agent's behavior for a subset of the environment states, 2) an *initiation set* consisting of all the states in which the option can be initiated, and 3) a *termination condition*, which specifies the conditions under which the option terminates. It is important to note that an option is not a sequence of actions; it is a closed-loop control rule, meaning that it is responsive to on-going state changes. Theoretically, when options are added to the set of admissible agent actions, the usual MDP formulation of RL extends to semi-Markov decision processes (SMDPs), with the one-step actions now becoming the "primitive actions." All of the theory and algorithms applicable to SMDPs can be appropriated for decision making

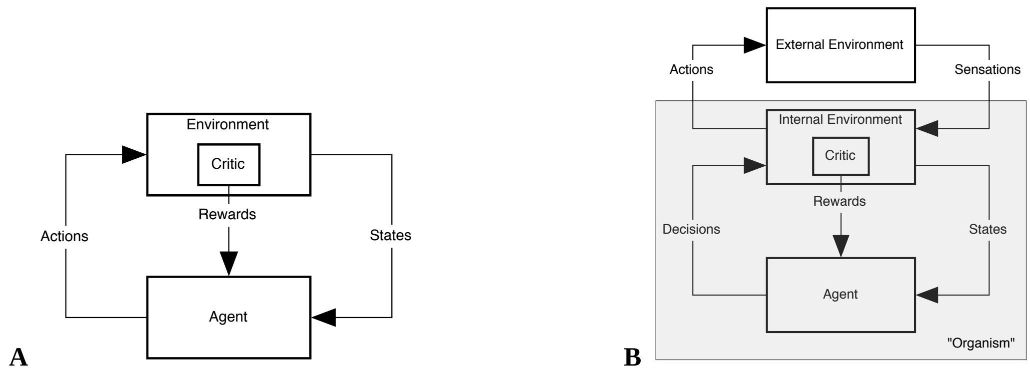


Figure 1. *Agent-Environment Interaction in Reinforcement Learning.* **A***: Reward is supplied to the agent from a "critic" in its environment.* **B***: An elaboration of Panel A in which the environment is factored into and internal and external environment, with reward coming form the former. The shaded box corresponds to what we would think of as the "organism."*

and learning with options [1, 29].

Two components of the the options framework are especially important for our approach:

1. *Option Models*: An option model is a probabilistic description of the effects of executing an option. As a function of an environment state where the option is initiated, it gives the probability with which the option will terminate at any other state, and it gives the total amount of reward expected over the option's execution. Option models can be learned from experience (usually only approximately) using standard methods. Option models allow stochastic planning methods to be extended to handle planning at higher levels of abstraction.

2. *Intra-option Learning Methods*: These methods allow the policies of many options to be updated simultaneously during an agent's interaction with the environment. If an option *could have* produced a primitive action in a given state, its policy can be updated on the basis of the observed consequences even though it was not directing the agent's behavior at the time. Intra-option methods essentially "multiplex" experience to greatly increase the efficiency of learning [29].

In most of the work with options, the set of options must be provided by the system designer. While an option's policy can be improved through learning, each option has to be predefined by providing its initiation set, termination condition, and the reward function that evaluates its performance. Many researchers have recognized the desirability of automatically creating options, and several approaches have recently been proposed (e.g., [7, 10, 17]). For the most part, these methods extract options from the learning system's attempts to solve a particular problem, whereas our approach creates options outside of the context of solving any particular problem.

Developing Hierarchical Collections of Skills—It is clear that children accumulate skills while they engage in intrinsically motivated behavior, e.g., while at play. When they notice that something they can do reliably results in an interesting consequence, they remember this in a form that will allow them to bring this consequence about if they wish to do so at a future time when they think it might contribute to a specific goal. Moreover, they improve the efficiency with which they bring about this interesting consequence with repetition, before they become bored and move on to something else. *We claim that the concepts of an option and an option model are exactly appropriate for developing analogs of this type of behavior in artificial agents*. An option model is not a passive model of environment dynamics; it is conditioned on the agent's activity. An option model basically says that "If I begin this behavior in this situation, then this is what is likely to happen." When stored appropriately, the agent will effectively know that it has the means to efficiently bring about these consequences, which is what the agent needs to know to both learn higher-level skills (that use lower-level skills as building blocks) and to learn how to solve specific tasks as they arise.

All skills acquired in this way do not have to be useful. Later learning in the context of specific tasks will assign values to skills depending on how useful they turn out to be. We already know how to do this using recently-developed hierarchical RL algorithms. The major computational challenge is to develop and cache a set of skills that is rich in skills that are likely to be widely useful. Intrinsic reward does not have to infallibly identify useful activities, but it has to do a reasonable job of identifying good candidates—and it shouldn't miss too much. If we speculate about the evolution of intrinsic motivational systems in animals, it is plausible that they have been tuned through evolution to do exactly this, resulting in the kind of "drive for mastery" that has been discussed by psychologists for at least half a century.

What kind of intrinsic reward function do we propose to implement? While there are several sources of inspiration for this as discussed above, in this work we focus on the striking connection between computational RL algorithms and the activity of dopamine neurons. In particular, we will illustrate how we use a kind of "surprise" analogous to the so-called novelty responses of dopamine neurons to implement one form of intrinsic reward.

Whatever the details of how intrinsic reward is defined, it should diminish with continued repetition of the activity that generates it. For example, continued exercise of causal influence on the environment should effectively lose its rewarding quality after becoming sufficiently "routine" (i.e., the agent gets bored). As a result, the agent moves on to learn another skill based on its discovery of another mode of controlling its environment, and so on. Similarly, exploration of regions about which the agent is not yet ready to learn should be aversive to the agent. Skills formed through earlier experience are available as action choices in this RL process. Policies for new skills have the potential of invoking existing skills. This will allow the construction of hierarchically organized collections of skills that become more sophisticated as the agent continues to accumulate experience. This process will naturally produce what Utgoff and Stracuzzi [30] called "many-layered" learning in which the agent learns what is easy to learn first, then uses this knowledge to learn harder things. This results in a generative power that is absent from current machine learning systems.

4. An Example

To make our discussion above more concrete, we briefly describe an example implementation of some of these ideas in a simple artificial "playroom" domain shown in Fig. 2A. In the playroom are a number of objects: a light switch, a ball, a bell, two movable blocks that are also buttons for turning music on and off, as well as a toy monkey that can make sounds. The agent has an eye, a hand, and a visual marker (seen as a cross hair in the figure). At any time step, the agent has the following actions available to it: 1) move eye to hand, 2) move eye to marker, 3) move eye one step north, south, east or west, 4) move eye to random object, 5) move hand to eye, 6) move hand to marker, 7) move marker to eye, and 8) move marker to hand. In addition, if both the eye and and hand are on some object, then natural operations suggested by the object become available, e.g., if both the hand and the eye are on the light switch then the action of flicking the light switch becomes available, and if both the hand and eye are on the ball, then the action of pushing the ball become available (the ball when pushed moves in a straight line to the marker), etc. Finally, there is a visual-search action that moves the eye to a random object in the room.

The objects in the playroom all have potentially interesting characteristics. The bell rings once and moves to a random adjacent square if the ball is kicked into it. The light switch controls the lighting in the room. The color of any of the blocks in the room is only visible if the light is on, otherwise they appear similarly gray. The blue block if pressed turns music on, while the red block if pressed turns music off. Either block can be pushed and as a result it moves to a random adjacent square. The toy monkey makes frightened sounds if simultaneously the room is dark and the music is on and the bell is rung. These objects were designed to have varying degrees of difficulty to engage. For example, to get the monkey to cry out requires the agent to do the following sequence of actions: 1) get its eye to the light switch, 2) move hand to eye, 3) push the light switch to turn the light on, 4) find the blue block with its eye, 5) move the hand to the eye, 6) press the blue block to turn music on, 7) find the light switch with its eye, 8) move hand to eye, 9) press light switch to turn light off, 10) find the bell with its eye, 11) move the marker to the eye, 12) find the ball with its eye, 13) move its hand to the ball, and 14) kick the ball to make the bell ring. Notice that if the agent has already learned how to turn the light on and off, how to turn music on, and how to make the bell ring, then those learned skills would be of obvious use in simplifying this process of engaging the toy monkey.

For this simple example, the agent has built-in notions of salience of stimuli. In particular, changes in light and sound intensity are considered salient by the playroom agent. The agent behaves by choosing actions according to an ϵ-greedy policy with respect to its value function [28]. Because the initial value function is uninformative, the agent starts by exploring its environment randomly. Each first encounter with a salient event initiates the learning of an option and an option-model for that salient event. For example, the first time the agent happens to turn the light on, it initiates the data-structures necessary for learning and storing the light-on option, including the initiation set, the policy, the termination probabilities, as well as for storing the light-on option-model including the terminal-state probabilities and the expected reward until termination. As the agent moves around the world, all the options and their models are simultaneously updated using intra-option learning algorithms. Initially, of course, the light-on option and its model will be nearly empty.

The agent's intrinsic reward is generated in a way suggested by the novelty response of dopamine neurons. The

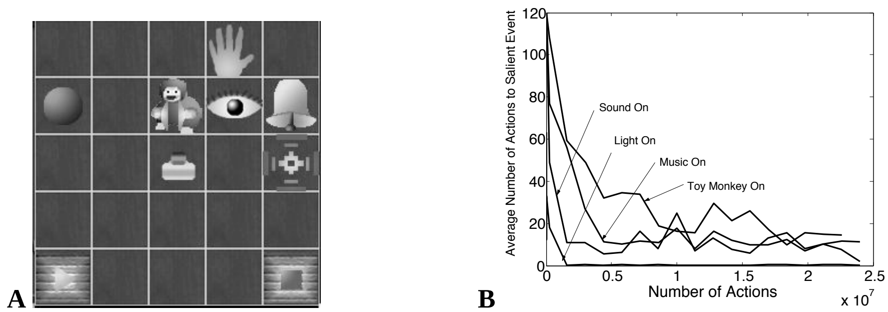


Figure 2. **A.** *Playroom domain. See text for details.* **B.** *Speed of learning of various skills. See text for details*

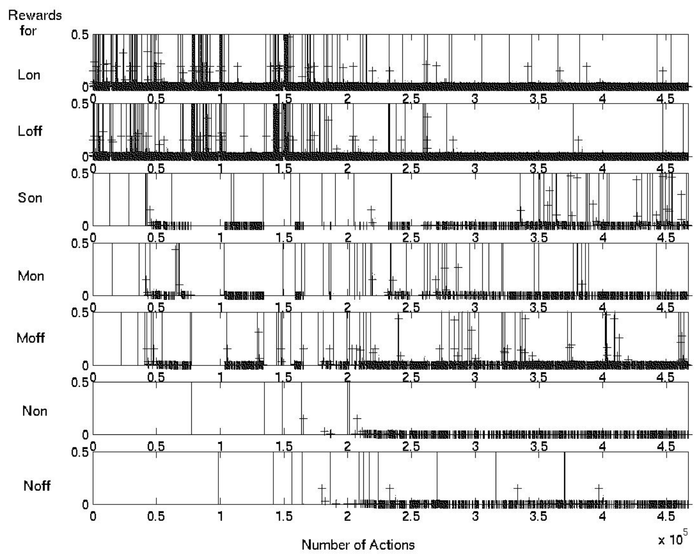


Figure 3. *Occurrence and magnitude of rewards for the salient events. See text for details.*

intrinsic reward for each salient event is proportional to the error in its prediction of that salient event according to the learned option model for that event. The intrinsic reward is used to update the value function the agent is using to determine its behavior in the playroom. As a result, when the agent encounters an unpredicted salient event a few times, its updated value function drives it to repeatedly attempt to achieve that salient event. There are two interesting side effects of this: 1) as the agent tries to repeatedly achieve the salient event, learning improves both its policy for doing so and its option-model that predicts the salient event, and 2) as its option policy and option model improve, the intrinsic reward diminishes and the agent gets "bored" with the associated salient event and moves on. Of course, the option policy and model become accurate in states the agent encounters frequently. Occasionally, the agent encounters the salient event in a state (set of sensor readings) that it has not encountered before, and it generates intrinsic reward again (it is "surprised").

A summary of results is presented in Fig. 3. Each panel of the figure is for a distinct salient event. The graph in each panel shows both the time steps at which the event occurs and the intrinsic reward associated by the agent to each occurrence. Each occurrence is denoted by a vertical bar whose height denotes the amount of associated intrinsic reward. Note that as one goes from top to bottom in this figure, the salient events become harder to achieve and, in fact, become more hierarchical. Indeed, the lowest one for turning on the monkey noise (Non) needs light on, music on, light off, sound on in sequence. A number of interesting results can be observed in this figure. First note that the salient events that are simpler to achieve occur earlier in time. For example, Lon (light turning on) and Loff (light turning off) are the simplest salient events, and the agent makes these happen quite early. The agent tries them a number of times (determined by the learning rate parameter and details of the agent's current value function) before getting bored and moving on to other salient events. The reward obtained for

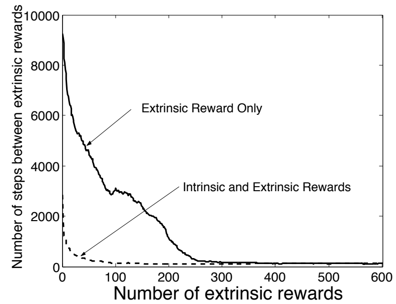


Figure 4. *The effect of intrinsically motivated learning when extrinsic reward is present. See text for details.*

each of these events diminishes after repeated exposure to the event. Thus, automatically, the skill of achieving the simpler events are learned before those for the more complex events.

Of course, the events keep happening despite their diminished capacity to reward because they are needed to achieve the more complex events. Consequently, the agent continues to turn the light on and off even after it has learned this skill because this is a step along the way toward turning on the music, as well along the way toward turning on the monkey noise. Finally note that the more complex skills are learned relatively quickly once the required sub-skills are in place, as one can see by the few rewards the agent receives for them. The agent is able to bootstrap and build upon the options it has already learned for the simpler events. The fact that all the options are learned is also seen in Fig. 2B, which shows how the time it takes the agent to bring about each option's target event changes with the agent's experience (there is an upper cutoff of 120 steps). This figure also shows that the simpler skills are learned earlier than the more complex ones.

An agent having a collection of skills learned through intrinsic reward can learn a wide variety of extrinsically rewarded tasks more easily than an agent lacking these skills. To illustrate, we looked at a playroom task in which extrinsic reward was available only if the agent succeeded in making the monkey cry out. This requires the 14 steps described above. This is difficult for an agent to learn if only the extrinsic reward is available, but much easier if the agent can use intrinsic reward to learn a collection of skills, some of which are relevant to the overall task. Fig. 4 compares the performance of two agents in this task. Each starts out with no knowledge of task, but one employs the intrinsic reward mechanism we have discussed above. The extrinsic reward is always available, but only when the monkey cries out. The figure, which shows the average of 100 repetitions of the experiment, clearly shows the advantage of learning with intrinsic reward.

5. Discussion

While the experiment and results described above serve as a concrete illustration of our basic ideas, they are merely a starting point in our study of intrinsically motivated learning. One of the key aspects of the Playroom example is that intrinsic reward is generated only by unexpected salient events. But this is only one of the simplest possibilities and has many limitations. It cannot account for what makes many forms of exploration and manipulation "interesting." In the future, we intend to implement computational analogs of other forms of intrinsic motivation as suggested by the psychological and neuroscience literatures and guided by the statistical

Despite the "toy" nature of this domain, these results are among the most sophisticated we have seen involving intrinsically motivated learning. Moreover, they were achieved quite directly by combining a collection of existing RL algorithms for learning options and option-models with a simple notion of intrinsic reward. The idea of intrinsic motivation for artificial agents is certainly not new, but we hope to have shown that the elaboration of the formal RL framework in the direction we have suggested, together with the use of recently-developed hierarchical RL algorithms, provides a fruitful basis for developing competently autonomous agents.

Acknowledgement Andrew Barto's research was funded by NSF grant CCF 0432143 and by a grant from DARPA's IPTO program. Satinder Singh and Nuttapong Chentanez were funded by NSF grant CCF 0432027 and by a grant from DARPA's IPTO program. The authors thank Peter Dayan for his essential input regarding the neuroscience of intrinsic reward systems.

References

[1] A. G. Barto and S. Mahadevan. Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamcal Systems: Theory and Applications*, 13:341–379, 2003.

[2] D. E. Berlyne. *Conflict, Arousal. and Curiosity*. McGraw-Hill, N.Y., 1960.

[3] P. Dayan and B. W. Balleine. Reward, motivation and reinforcement learning. *Neuron*, 36:285–298, 2002.

[4] P. Dayan and T. J. Sejnowski. Exploration bonuses and dual control. *Machine Learning*, 25:5–22, 1996.

[5] E. L. Deci and R. M. Ryan. *Intrinsic Motivation and Self-Determination in Human Behavior*. Plenum Press, N.Y., 1985.

[6] G. Di Chiara. Drug addiction as dopamine-dependent associative learning disorder. *European Journal of Pharmacology*, 375(1-3):13–30, 1999.

[7] B. Digney. Learning hierarchical control structure from multiple tasks and changing environments. In *From Animals to Animats 5: The Fifth Conference on Simulation of Adaptive Behavior*, Cambridge, MA, 1998. MIT Press.

[8] K. J. Friston, G. Tononi, G. N. Reeke, O.Sporns, and G. M. Edelman. Value-dependent selection in the brain: Simulation in a synthetic neural model. *Neuroscience*, 59:229–243, 1994.

[9] K. Groos. *The Play of Man*. D. Appleton, N.Y., 1901.

[10] B. Hengst. Discovering hierarchy in reinforcement learning with HEXQ. In *Maching Learning: Proceedings of the Nineteenth International Conference on Machine Learning*, pages 243–250, San Francisco, CA, 2002. Morgan Kaufmann.

[11] J. C. Horvitz, T. Stewart, and B. Jacobs. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Research*, 759:251–258, 1997.

[12] J. C. Houk, J. L. Adams, and A. G. Barto. A model of how the basal ganglia generates and uses neural signals that predict reinforcement. In J. C. Houk, J. L. Davis, and D. G. Beiser, editors, *Models of Information Processing in the Basal Ganglia*, pages 249–270. MIT Press, Cambridge, MA, 1995.

[13] X. Huang and J. Weng. Novelty and reinforcement learning in the value system of developmental robots. In C. G. Prince, Y. Demiris, Y. Marom, H. Kozima, and C. Balkenius, editors, *Proceedings of the Second International Workshop on Epigenetic Robotics : Modeling Cognitive Development in Robotic Systems*, pages 47–55, Edinburgh, Scotland, 2002. Lund University Cognitive Studies.

[14] S. Kakade and P. Dayan. Dopamine bonuses. In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems 13*, pages 131–137. MIT Press, 2001.

[15] S. Kakade and P. Dayan. Dopamine: Generalization and bonuses. *Neural Networks*, 15:549–559, 2002.

[16] F. Kaplan and P.-Y. Oudeyer. Motivational principles for visual know-how development. In C. G. Prince, L. Berthouze, H. Kozima, D. Bullock, G. Stojanov, and C. Balkenius, editors, *Proceedings of the Third International Workshop on Epigenetic Robotics : Modeling Cognitive Development in Robotic Systems*, pages 73–80, Edinburgh, Scotland, 2003. Lund University Cognitive Studies.

[17] A. McGovern. *Autonomous Discovery of Temporal Abstractions from Interaction with An Environment*. PhD thesis, University of Massachusetts, 2002.

[18] P. R. Montague, P. Dayan, and T. J. Sejnowski. A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience*, 16:1936–1947, 1996.

[19] A. Ng, D. Harada, and S. Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In *Proceedings of the Sixteenth International Conference on Machine Learning*. Morgan Kaufmann, 1999.

[20] J. Piaget. *The Origins of Intelligence in Children*. Norton, N.Y., 1952.

[21] P. Reed, C. Mitchell, and T. Nokes. Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. *Animal Learning and Behavior*, 24:38–45, 1996.

[22] R. Saunders. *Curious Design Agents and Artificial Creativity: A Synthetic Approach to the Study of Creative Behaviour*. PhD thesis, University of Sydney, 2002.

[23] J. Schmidhuber. A possibility for implementing curiosity and boredom in model-building neural controllers. In *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, pages 222–227, Cambridge, MA, 1991. MIT Press.

[24] J. Schmidhuber and J. Storck. Reinforcement driven information acquisition in nondeterministic environments, 1993. Technical report, Fakultat fur Informatik, Technische Universit at Munchen.

[25] W. Schultz. Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80:1–27, 1998.

[26] W. Schultz, P. Dayan, and P. R. Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1598, March 1997.

[27] R. S. Sutton. Integrated modeling and control based on reinforcement learning and dynamic programming. In R. P. Lippmann, J. E. Moody, and D. S. Touretzky, editors, *Advances in Neural Information Processing Systems: Proceedings of the 1990 Conference*, pages 471–478, San Mateo, CA, 1991. Morgan Kaufmann.

[28] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[29] R. S. Sutton, D. Precup, and S. Singh. Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112:181–211, 1999.

[30] P. E. Utgoff and D. J. Stracuzzi. Many-layered learning. *Neural Computation*, 14:2497–2539, 2002.

[31] J. Wang, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental develoopment by robots and animals. *Science*, 291:599–600, 2001.

[32] R. W. White. Motivation reconsidered: The concept of competence. *Psychological Review*, 66:297–333, 1959.

An imaging study on human action selection using hierarchical rules

Hidefumi Funakoshi
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma,
Nara 630-0192, Japan
hidefu-f@is.naist.jp

Wako Yoshida
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma,
Nara 630-0192, Japan
CREST, Japan Science and Technology Agency
wako-y@is.naist.jp

Shin Ishii
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma,
Nara 630-0192, Japan
CREST, Japan Science and Technology Agency
ishii@is.naist.jp

Abstract

To adapt to a dynamic environment, appropriate behavioral switching is necessary. In most real-world problems there are numerous possible actions, and it is often impossible to select the optimal action by evaluating all of them. Even in such a situation, humans can select an action efficiently by searching only a subspace of the whole action space. In this study, we design a Multi Feature Sorting Task in which the behavioral rules have a hierarchical structure, and conduct an fMRI experiment using the task. This task consists of two kinds of rule switches: a higher-order switch to search for a rule across different subspaces, and a lower-order switch to change a rule within the same subspace. The results of our imaging study show that the left inferior frontal gyrus is involved in the higher-order switch, and the right fronto-polar and right dorsolateral prefrontal cortex are significantly activated with the lower-order switch. We also suggest a functional model for the prefrontal cortex which explains the hierarchical rule-switching mechanism.

1 Introduction

In the real environment around us, there are numerous possible behaviors in each situation, and it may be impossible to immediately make an appropriate decision by evaluating all of them. To adapt to a dynamic environment, humans must seek action candidates efficiently and select the best one within a limited time. Recent studies in the engineering field suggest that a hierarchical structure of action candidates is useful for effective action selection [2][18], and an analogous method may be performed in human behavioral decisions. For instance, when we search for a lost article in the house, we will first check the most likely places rather than search the house uniformly. Such a searching scheme uses a hierarchical structure of available information; the whole search space is divided into subspaces, and a local optimization problem in one subspace is first solved. This hierarchical approach is effective as a computational algorithm and reasonable for a human behavioral model. However, it has been unclear how such a hierarchical mechanism operates in the real brain.

Assuming a human selects an action according to behavioral rules, he/she should switch between rules in response to environmental changes. The Wisconsin Card Sorting Task (WCST) [7] is one of the best-known tasks for studying such a rule-switching process. In WCST, the subject is required to discover a hidden correct rule from multiple possible rules using true/false feedback given correspondingly to the selected rule. Since the correct rule often changes without notice, the subject should try a new rule if he/she receives a false feedback. Many imaging and lesion studies have shown that prefrontal cortex is closely involved in solving WCST [1][6][9][12][19]. One study using a modified WCST with a variable number of rules revealed that the bilateral rostral inferior frontal sulcus (BA45/44) was activated when a subject switched rules due to environmental changes (correct rule changing) [9]. In contrast, another imaging study using a categorization task suggested that the bilateral fronto-polar prefrontal cortex (BA10) and left superior frontal sulcus (BA9/10) were related to rule switches [17]. Although both of these tasks needed rule-switch processes, different regions of the prefrontal cortex were reported as being engaged in rule-switch functions; however, functional segregation of these regions has yet to be clarified.

Existing studies based on rule-switching tasks assumed that possible rules were independent of each other and that a feedback for the used rule had no clue about the new correct rule; thus, a subject should examine all rule candidates uniformly. Since we aim in this study at specifying the brain regions involved in a hierarchical rule searching mechanism, we have designed a Multi Feature Sorting Task in which the behavioral rules have a hierarchical structure. All rules in our task are categorized into two meta-rules, and hence there are two kinds of rule switches: a meta-rule switch (higher order switch) to search for a rule belonging to the other meta-rule class (subspace), and a rule switch (lower order switch) to change rules within the same meta-rule class subspace. Using this newly devised task, we conducted an fMRI experiment which showed that the different regions were activated during higher and lower order rule switching. This result suggests that different regions in the prefrontal cortex may cooperate to solve action selection problems in complicated situations.

2 Methods

2.1 Subjects

Eight normal subjects (7 males and 1 female) who were graduate students participated in this experiment. Before scanning, all subjects were instructed about the aims and procedures of the experiment, and gave their written informed consent which was reviewed and approved by the ethical committee of Advanced Telecommunications Research Institute International (ATR). Each subject was paid a fixed monetary reward regardless of task performance. To acquire proficiency in the task, all subjects practiced a training task equivalent to the scanning one on the day before scanning.

2.2 Multi Feature Sorting Task

In this study, we designed a Multi Feature Sorting Task in which the subject was required to sort three figures with multiple features using a rule (Fig.1). The purpose of the subject was to find out a hidden correct rule using feedback information of past behaviors. Three figures were displayed on a screen in the MRI device, and the subject sorted them by pushing the corresponding three buttons one by one.

Table 1. Six rules and two meta-rules

"Feature rule" (FR)
1. The first sorting is "number of vertices" and the second sorting is "size", or vice versa
2. The first sorting is "number of vertices" and the second sorting is "brightness", or vice versa
3. The first sorting is "size " and the second sorting is "brightness", or vice versa

"Order rule" (OR)
1. The first sorting is "ascending order" and the second sorting is "descending order"
2. The first sorting is "descending order" and the second sorting is "ascending order"
3. The first and second sorting orders are the same as each other

When the subject pushed a button, a red marker above the corresponding figure was turned on. There were three features: "number of vertices", "size" and "brightness", and each figure was categorized as "large", "middle" or "small" for each of these features. For example, a large dark square would be represented as "number of vertices: middle; size: large; brightness: small". Since the features of each figure did not overlap with other figures displayed simultaneously, it is possible to sort the figures according to a single feature in ascending or descending order. For example, "descending order for the number of vertices" corresponds to the sorting order pentagon, square, and triangle. Subjects performed such sorting twice using the same or different way within a single trial (Fig.1). Namely, after the subject sorted three figures (stimulus 1) by pushing three buttons, the next three figures (stimulus 2) were displayed to sort once again. This defined the subject's behavior within one trial. At the center of the screen, a fixation cross was displayed which was red for 2 sec after appearance of a stimulus, to encourage a response, and was yellow thereafter. Subjects were instructed to sort three figures by pushing buttons three times as quickly as possible during the red fixation period; if the subject could not complete a sorting task within the red fixation period for the first and/or second stimuli in a trial, it was regarded as a mis-trial and fed a caution message but no point feedback. Since a set of three figures can be sorted in two ways, "ascending order" and "descending order", to every three features, there are six sorting options in one sorting. So, 36 sorting options existed within a single trial which consisted of two sorting behaviors. To make a hierarchical structure for the sorting rules, a favorable set of two sorting behaviors was integrated into six rules without overlap, and these six rules were categorized into two meta-rules (Table 1). One meta-rule was a "feature rule" which focuses only on the combination of features ("number of vertices", "size" or "brightness") in a set of two sorting behaviors and not on sorting order ("as-

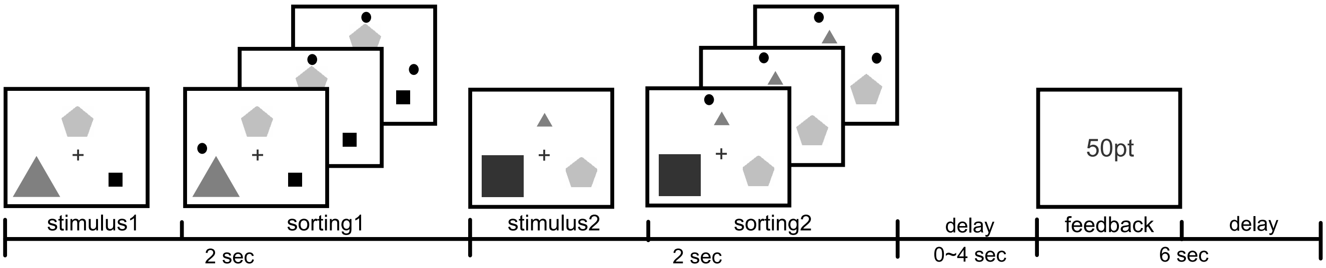


Figure 1. Stimulus and time design of a single trial in Multi Feature Sorting Task.

cending" or "descending"). A "feature rule" included 8 out of 36 sorting options, since the combination of two features had two patterns, and each combination pairing "ascending" and "descending" had four patterns. The other meta-rule was an "order rule" which focused only on permutations in the two sorting behaviors regardless of the features used in sorting. Since an "order rule" allowed three features used in the combination, the first and second "order rule" in Table 1 each with one possible order pattern consisted of 3 sorting options, while the third one with two possible order patterns consisted of 6 sorting options. It is noted that the number of sorting options of a "feature rule" was more than that of an "order rule". However, since the subjects were likely to use few favorites among allowable sorting options in each rule, this difference in the numbers of sorting options could be ignored; actually, the entropy of sorting options selected by subjects was small enough within a single meta-rule (data not shown).

For each trial, there was a hidden correct rule selected from the six rules in Table 1. After the subject finished a set of two sorting behaviors, feedback was displayed in the center of the screen according to the used sorting rule and the correct rule. For this task, we designed a probabilistic feedback. When the used rule agreed with the correct rule, the subject was given 50 points with 90% probability but 0 points with 10% probability. When the used rule was different from the correct rule, the subject was given 0 points with 90% probability but 50 points with 10% probability. All subjects were informed that the feedback was to be given in a probabilistic manner, but the rate of probability was not revealed. In case of deterministic feedback, if the negative feedback was given for the used rule, the subject just excludes it from rule candidates. In case of probabilistic feedback, however, since the feedback information is ambiguous, the subject may determine the priority for every rule to search for the correct rule. Accordingly, the subject was required to perform "exploration", i.e., searching for a new rule, or "exploitation", i.e., continuing with the same rule as the previous one, based on the outcome of previous tasks and these decisions are an introspective one.

When the rule used agreed with the correct rule in any three among four successive trials, the correct rule was changed to another one without notification to the subject. Otherwise, the correct rule was the same as that in the previous trial. When the correct rule was changed, the new rule was selected with a higher probability (about 70%) from the same meta-rule class than from the other meta-rule class. Moreover the correct rule change from one meta-rule class to another one did not occur continuously. Subjects were told that this would be the case but the changing frequency was not revealed.

A control task was conducted to determine the baseline of imaging analysis. In the control task, the basic experimental procedure, consisting of stimuli and the requirement for subject's behaviors, was the same as the main sorting task, while the fixed correct rule for all trials was given as a visual message at the beginning of the control task. Thus, subjects did not need to select a rule themselves. One session consisted of the first main task (45 trials), a control task (5 trials) and the second main task (45 trials), and each subject performed 3 sessions in the experiment.

2.3 Scanning Procedures and Imaging Analysis

Using a whole-brain 1.5-tesla scanner (Magnetic Eclipse; Shimadzu-Marconi, Kyoto, Japan), functional images were obtained with T2*-weighted echo planner imaging (EPI), with blood oxygen-level depletion (BOLD) contrast. The volumes were acquired continuously every 2.0 sec (TR) with 20 slices of 5 mm thickness (TE: 48 msec, FA: $80°$, FOV: 192 mm, matrix size: 64×64). The first six (12 sec) EPI images in each session were excluded from the analysis to avoid the effect of T1 equilibrium. During one session, 560 EPI images were acquired. To investigate anatomical localization, T1-weighted three-dimensional images were acquired (TR: 12 msec, TE: 4.5 sec, FA: $20°$, FOV: 256 mm, matrix size: 256×256, thickness: 1 mm, 191 slices).

Imaging data were analyzed using Statistical Parametric Mapping 99 (SPM99; Wellcome Department of Cognitive Neurology, London, UK). All functional images from each subject were realigned with the first image, using rigid transformation, and then the slice timing was corrected. After that, EPIs were registered to the individual anatomical image. The EPI images were normalized using parameters such that anatomical T1 images were normalized to the MNI (Montreal Neurological Institute) template. The normalized EPIs were spatially smoothed with a Gaussian kernel of 10 mm (FWHM).

We excluded imaging data for mis-trials. We then defined an event-block as 10 sec, such that the onset was the feedback of the previous trial and the end was the finish time of the second sorting in the current trial. For the analysis, three kinds of events-blocks were extracted from all event-blocks according to the subject's behaviors in the corresponding trial. The first was "meta-rule switch (MSW)", in which a subject tried a new sorting rule whose meta-rule class was different from the previous one. The second was "rule switch (RSW)", in which a subject used a different rule within the same meta-rule class as the previous one. Note that the difference between rule and meta-rule switches could be distinguished based only on the subject's behaviors. The third was "exploitation (EXP)", in which a subject used the same rule as in the previous trial, and the rule was the correct one. Event-blocks not categorized in any of these three kinds were excluded from the analysis. The MSW and RSW were the same in terms of selecting one rule out of the six possible rules, but differed in that the MSW required a switch process between two meta-rules. The EXP differed from MSW and RSW because no rule switches were necessary. All event-blocks were convolved with a homodynamic response function (HRF), and the control task was designed as an epoch which defined base activation. Six realignment parameters were also designed as regressors to eliminate moving artifacts. The data were high-pass-filtered using a low-frequency cosine function with a cut-off time of 60 sec. To account for inter-subject variability and to allow statistical inference at the population level, one sample t-test for statistical significance of group random effects was used. For comparison between MSW and RSW, the threshold at the voxel level was set to $p<0.01$ (uncorrected), and for that between RSW and EXP, and between MSW and EXP, it was set to $p<0.001$ (uncorrected). After that, cluster level analysis was applied with $p<0.05$ (corrected). We also conducted a time-course analysis of regions found to be significantly activated in the group analysis. The activation level for each region was represented as the average of signal intensities of all voxels within the region. These time course data were smoothed using a high-pass filter with a cut-off time of 100 sec and a low-pass filter with a cut-off time of 8 sec.

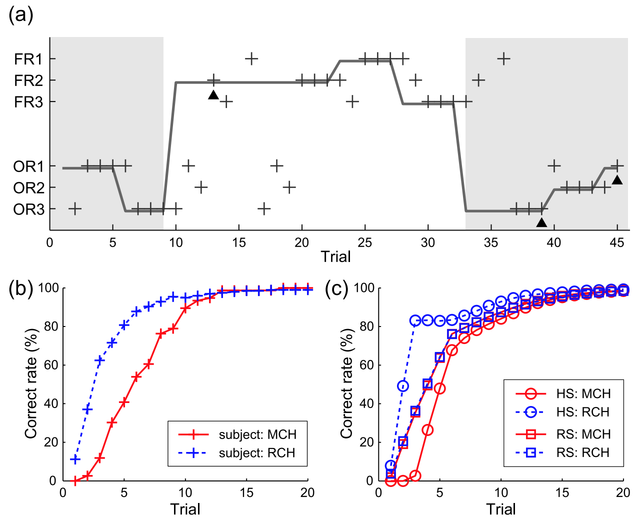


Figure 2. Behavioral results. (a) An example of a subject's profile. (b) The correct rate after meta-rule and rule change for subjects. (c) The same plot as (b) for two computer agents.

3 Results

3.1 Behavioral Results

To show the efficiency of hierarchical searching in our task, we conducted a computer simulation using two kinds of computer agents. One was a hierarchical-search agent (HS) which explores within the current meta-rule class with a higher priority than within the other meta-rule class, and the other was a random-search agent (RS) which searches over five possible rules except for the current one randomly. In the simulation, these two agents performed 100,000 sessions of the identical task to the fMRI task. When a positive feedback (50 point) was given, both agents chose the previously used rule at the next trial. The average point acquired by the HS was significantly higher than the RS, and it was shown that the searching scheme using the hierarchical structure of rules is effective to solve our task.

In the Multi Feature Sorting Task, the correct sorting rule was changed depending on the subject's behavior, and there were two types of rule changes: a "meta-rule change (MCH)" in which a new correct rule is selected from the other meta-rule class, and a "rule change (RCH)" in which a new correct rule is selected from the current meta-rule class. The number of correct rule changes varied among subjects; the average number was 40 ± 5, consisting of 13 ± 2 meta-rule changes and 27 ± 3 rule changes. The Fig.2(a) shows an example of behavioral profile in a certain sorting task block, where the abscissa denotes the number of trials. The line and crosses in this figure denote correct rule transition and subject's used rule in each trial, respectively. Triangles indicate the trials in which the subject received feedback stimuli contradictory to the used rule (10%) due to the probabilistic nature of feedback stimuli. The gray and white backgrounds show the periods of "order rule" and "feature rule", respectively. When the MCH occurred at the 10th trial, the subject first explored all rules within "order rule" to which the previously correct rule belonged, and then switched to "feature rule". The opposite-directional behaviors were also observed after the 33th trial. These behaviors show that the subject was likely to explore one meta-rule exhaustively before switching to the other meta-rule.

Fig.2(b) shows the time course of choosing correct rule after MCH and RCH for subjects. The abscissa denotes the number of trials in which the zero value is the timing of correct rule changing. The solid and dot lines correspond to the rate of correct trials after MCH and RCH, respectively. The lower correct rate immediately after MCH than RCH ($p<0.001$) supported that the subjects first explored within the current meta-rule, and then explored the other meta-rule. Similar properties were also examined for the two agents, RS and HS, described above (Fig.2(c)). Since the RS searched for correct rules randomly, the correct rates after MCH and RCH behaved similarly. In the HS, the correct rate after MCH was lower than that after RCH, supporting that the subject utilized the hierarchical rule structure to search for correct rules quickly and efficiently.

Subject behaviors in the main task (other than excluded ones) can be classified into three kinds of conditions, MSW, RSW and EXP. Each subject was required to make prompt and accurate responses in the experiment, and their reaction time (RT, the time interval between presentation of feedback stimulus in the previous trial and initiation of a response to the stimulus in the current trial) was examined. Since subjects decided the current rule based on previous trial feedback, this RT may reflect the cognitive process of interest events. The RT was significantly longer in MSW than in RSW ($p<0.05$), implying that MSW needs heavier cognitive processing than RSW. Since RTs between RSW and EXP showed no significant difference, it is considered that cognitive processing inherent to RSW had been completed within the 6 sec period between the previous feedback presentation and the current stimulus presentation.

3.2 Imaging results

Brain areas significantly activated in the MSW condition and the RSW condition were compared. Group analysis showed significant activation of the left prefrontal lobe, especially in the inferior frontal gyrus (BA11, 45, 47) and insula (BA13), and these statistics are summarized in Table 2. Furthermore, we divided all voxels of the left inferior frontal gyrus into three areas, BA47, BA11 and BA45, and applied a time-course analysis to each of these three areas (Fig.3, lower panels). In each lower panel of Fig.3, the ordinate and abscissa denote the BOLD signal changing rate and the time elapsed since the feedback presentation in the previous trial, respectively. 0 sec and the line at 6 sec denote the timing of the feedback and the stimulus 1 presentation, respectively. In the MSW condition, BA47 had a clear activation peak after feedback presentation, while BA11 and BA45 showed significant activation related to switch events but no distinct peaks.

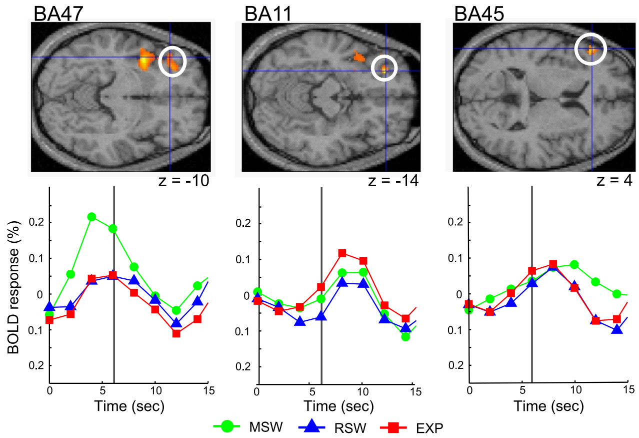


Figure 3. Imaging results for MSW vs RSW. Major activated regions and time courses of these BOLD responses.

We next compared brain images between the RSW condition and the EXP condition, and the statistics of activated regions are also summarized in Table 2. Figure 4 shows the right cortical hemisphere and the areas activated in the RSW condition: the right superior frontal gyrus (BA10), right middle frontal gyrus (BA9/46,6) and superior parietal lobule (BA7,40). The time courses of signal intensities in these areas (Fig.4, lower panels) reveal that the superior frontal gyrus (BA10) showed a marked activation peak compared with the other conditions. Although BA9/46 also showed an activation peak in the RSW condition compared with the EXP condition, this peak also occurred in the MSW condition. In BA6 of the middle frontal gyrus, although the overall activation level was higher in the MSW and RSW conditions than in the EXP condition, the time courses resembled each other in all three conditions.

In addition, the MSW condition and the EXP condition were compared, and the statistics of activated regions in the frontal cortex are summarized in Table 2. This comparison shows that regions in left BA47 and right BA46 were significantly activated in the MSW condition.

Table 2. Statistics of significantly activated regions

Condition	Region	Left / Right	Brodmann	Talairach[x,y,z]			*Z*-value
MSW > RSW	Inferior PFC	L	11/45/47	-28	38	-14	3.52
	Extra-Nuclear	L	13	-38	5	-7	3.99
	Insula	L	13	-36	17	-1	3.03
	Superior Temporal	L	22/38	-45	5	-8	2.78
RSW > EXP	Fronto-Polar PFC	R	10	32	58	3	3.86
	Dorsolateral PFC	R	9/46	50	20	24	4.39
	Middle Frontal Gyrus	R	6	32	2	50	3.95
	Precuneus	R	7	6	-64	46	5.54
	Precuneus	L	7	-6	-61	56	4.44
	Superior Parietal	R	7	30	-58	49	4.38
	Inferior Parietal	R	40	42	-54	45	3.77
	Supramarginal Gyrus	L	40	-40	-43	37	4.97
MSW > EXP	Inferior PFC	L	47	-39	21	-3	4.48
	Lateral PFC	R	46	51	24	23	4.49

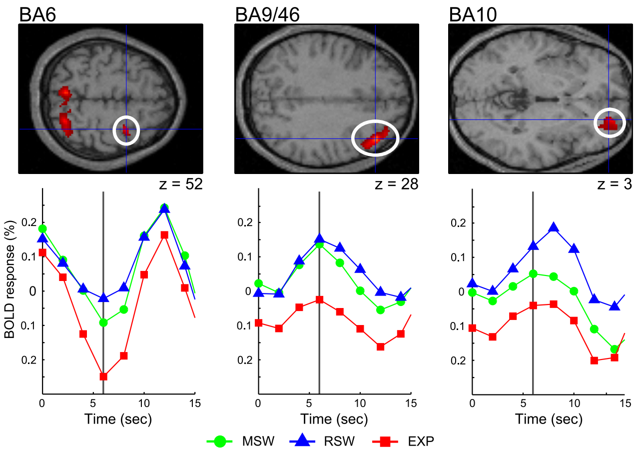


Figure 4. Imaging results for RSW vs EXP. Mainly activated regions and time courses of these BOLD responses.

4 Discussion

4.1 Meta rule switch: higher order

Because there is a hierarchical structure of rules in our task, an appropriate search consists of a higher-order switch, i.e., a switch to a rule of the other meta-rule class, and a lower switch, i.e., a switch to a rule within the current meta-rule class. We consider this hierarchical structure introduces a ‘context’ to the exploration strategy for correct rules.

In the MSW condition, the left inferior frontal gyrus, consisting of BA11, 45 and 47, was significantly activated. Analysis of the three Brodmann areas revealed that each area has a different time course. BA47 showed a temporal increase of the signal intensity in MSW which was not observed in either RSW or EXP. In both the MSW and RSW conditions, subjects switched their rule because they were given 0 points as feedback in the previous trial. Since the activation of BA47 was observed only in MSW, however, this region was not involved in the detection of erroneous feedback. We also found that the activation in both BA11 and BA45 exhibited similar time courses in all three conditions; thus, BA47 is closely related to the meta-rule switch process in the left inferior frontal gyrus.

According to recent imaging studies, the left inferior frontal gyrus plays an important role in the retrieval process for episodic memory [5][10][16]. It was suggested that one of the cognitive processes in episodic memory retrieval is the systematic analysis of possible semantic relations between a stimulus and the known characteristics of potential information sources, which would be helpful for recollecting contextual details of the encountered stimulus [3][13]. To isolate this cognitive process, Dobbins et al. [5] used a task in which the subject recalled a word-class after having performed a semantic classification of many words. This study revealed that the left inferior frontal gyrus (BA47) was concerned with the information retrieving process for word stimuli. Other studies have also shown that almost the identical region was involved in recollections related to the recognized stimulus [8][15]. We consider that meta-rules (higher-order components) are intensive information representations of lower-order rules in our task. Thus, restricting the search space based on meta-rules may exploit a cognitive process that performs efficient information retrieval for episodic memory, i.e., contextual information. Although Dobbins et al.’s task was a linguistic one while ours is a diagrammatic one, and they have different modalities, these results suggest that the left inferior prefrontal gyrus plays an important role in manipulating aggregated information.

4.2 Rule switch: lower order

In the RSW condition, we found that the right superior frontal gyrus (BA10) and right middle frontal gyrus (BA9/46,6) were significantly activated.

A previous study using a categorization task, in which the subject was required to search for a hidden correct rule by trial and error, suggested that the right superior frontal gyrus (BA10) was involved in seeking rules which was induced by changing the correct rule [17]. However, this region was not active when subjects sought a correct rule in WCST [9][12]. In both tasks, subjects had to switch rules based only on feedback; hence, they knew it was necessary to switch rules if they received a 'false' feedback. In the categorization task, each stimulus could be compatible with multiple rules. If a false feedback was given, therefore, the subject could eliminate not only the used rule but also several other rules, whereas a true feedback did not necessarily indicate that the used rule was correct. Thus, the subject would maintain more than one possible rule candidate. In the WCST, however, feedback was given according only to the rule used by the subject. If a false feedback was given, therefore, the subject simply removed the used rule from the candidates. In contrast, probabilistic feedback was given in our task. If an unfavorable feedback was given, the subject would be expected to reduce the probability (likelihood) that the used rule was correct, and to increase the probability of the other rule candidates being correct; if a favorable feedback was given, the subject would be expected to perform the contrary. The common feature of the categorization task and our task, in both of which activation of BA10 was found, is that two or more rule candidates should be handled in response to a given feedback because feedback in both cases was not explicit. BA10 may be activated when the subject estimates the hidden correct rule from given feedbacks and updates the likelihoods of rule candidates so as to redefine the rules' priorities. Moreover, according to our time-course analysis of this region, a more prominent activation was found in RSW than in MSW (although the activation was larger in MSW than in EXP). This result can be interpreted as follows. In RSW, likelihoods can be updated because removal of a used rule reduces the number of possible rule candidates. In MSW, in contrast, removal of a used rule does not reduce the number of possible rule candidates because it does not yield any knowledge on rule candidates belonging to the other meta-rule class; thus, there is no need to update the prioritized weights.

In the lower-order RSW, the right dorsolateral prefrontal cortex (BA9/46) was significantly activated. This region was previously found to be activated in WCST regardless of whether a true or false feedback was given [12]. Although subjects with lesions in this region could discover the first correct rule, they could not adapt to changes in the correct rule because they clung to this rule [4][11][14]. It is therefore thought that BA9/46 is involved in monitoring and/or updating the information stored in working memory. In our experiment, the activation intensity of this region in the EXP condition was lower than that in the MSW and RSW conditions; this is consistent with current understanding as outlined above.

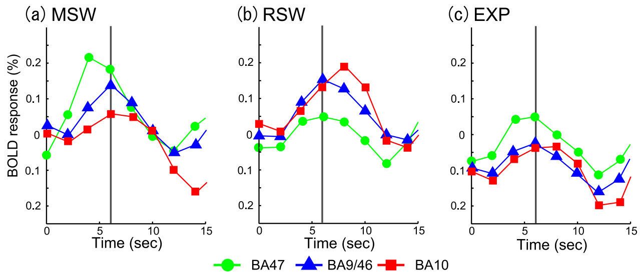


Figure 5. The BOLD responses in each event condition. (a) Meta-rule switching; (b) rule switching; (c) exploitation.

4.3 Information processing hypothesis

Time-course analysis of significantly activated regions indicated that the timing of activation was different in each of the three activated regions in the prefrontal cortex. Based on these results, we suggest a brain information processing model which explains the behaviors of rule switching.

The time courses of signal intensities in the three regions, discriminated by the three behavioral conditions, are shown in Fig.5. These different time courses can be interpreted as meaning that the subject first limits the searching space in BA47, then loads rule candidates in working memory in BA9/46, and finally determines the priorities of these candidates in BA10. In the MSW condition (Fig.5(a)), the activation increase in the left inferior frontal gyrus (BA47) was followed by activation in the right dorsolateral prefrontal cortex (BA9/46), whereas the fronto-polar prefrontal cortex (BA10) did not show any distinct activation. Since the subject arbitrarily selects one rule from the loaded candidates, activation of BA10 is not required because of the absence of any priorities among the loaded candidate rules. In the RSW condition (Fig.5(b)), activation of BA47 does not occur because the searching subspace is already determined. The subject first refers to the candidates held in BA9/46, and the priority of each candidate is then assigned in BA10. In the EXP condition (Fig.5(c)), none of the regions are markedly activated because the sub-processes above are not necessary.

5 Concluding remarks

We designed a Multi Feature Sorting Task in which the behavioral rules have a hierarchical structure, and conducted an fMRI experiment using this task. In our task, subjects were required to apply two kinds of rule switches which correspond to the retrieval of different hierarchies. The left inferior frontal gyrus (BA47) was specifically activated in higher-order meta-rule switches. It is considered that this region restricts the searching space by handling intensive information, in agreement with previous studies suggesting that BA47 is involved in recollecting information from episodic memory; this recollection function is useful for limiting rule candidates, as required in our task. The right fronto-polar prefrontal cortex (BA10) was specifically activated in lower-order rule switches; this region may thus be involved in prioritizing rules, in agreement with previous work suggesting that BA10 is involved in predictive rule switching tasks. Our results suggest that humans can effectively represent information as a hierarchical rule structure which can be operated efficiently by incorporating contextual information.

References

[1] KF Berman, JL Ostrem, C Randolph, J Gold, TE Goldberg, R Coppola, RE Carson, P Herscovitch, and DR Weinberger. Physiological activation of a cortical network during performance of the wisconsin card sorting test: a positron emission tomography study. *Neuropsychologia*, 33, 1995.

[2] J Bryson. Hierarchy and sequence vs. full parallelism in reactive action selection architectures. *From Animals to Animats*, 6:147–156, 2000.

[3] PW Burgess and T Shallice. Confabulation and the control of recollection. *Memory*, 4:359–411, 1996.

[4] R Dias, TW Robbins, and AC Roberts. Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*, 380:69–72, 1996.

[5] IG Dobbins, H Foley, DL Schacter, and AD Wagner. Executive control during episodic retrieval: multiple prefrontal processes subserve source memory. *Neuron*, 35:989–96, 2002.

[6] TE Goldberg, KF Berman, K Fleming, J Ostrem, JD Van Horn, G Esposito, VS Mattay, JM Gold, and DR Weinberger. Uncoupling cognitive workload and prefrontal cortical physiology: a PET rCBF study. *Neuroimage*, 7:296–303, 1998.

[7] DA Grant and EA Berg. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Exp Psychol*, 38:404–411, 1948.

[8] RN Henson, MD Rugg, T Shallice, O Josephs, and RJ Dolan. Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *J Neurosci.*, 15:3962–72, 1999.

[9] S Konishi, K Nakajima, I Uchida, M Kameyama, K Nakahara, K Sekihara, and Y Miyashita. Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nat Neurosci.*, 1:80–4, 1998.

[10] M Lepage, M Brodeur, and P Bourgouin. Prefrontal cortex contribution to associative recognition memory in humans: an event-related functional magnetic resonance imaging study. *Neurosci Lett.*, 346:73–6, 2003.

[11] B Milner. Effects of different brain lesions on card sorting. *Arch. Neurol*, 9:90–100, 1963.

[12] O Monchi, M Petrides, V Petre, K Worsley, and A Dagher. Wisconsin card sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J Neurosci.*, 21:7733–41, 2001.

[13] M Moscovitch and B Melo. Strategic retrieval and the frontal lobes: evidence from confabulation and amnesia. *Neuropsychologia*, 35:1017–1034, 1997.

[14] RE Passingham. Non-reversal shifts after selective prefrontal ablations in monkeys (Macaca mulatta). *Neuropsychologia*, 10:41–46, 1972.

[15] MD Rugg, LJ Otten, and RN Henson. The neural basis of episodic memory: evidence from functional neuroimaging. *Philos Trans R Soc Lond B Biol Sci.*, 29:1097–110, 2002.

[16] MD Rugg and EL Wilding. Abstract retrieval processing and episodic memory. *Trends Cogn Sci.*, 4:108–115, 2000.

[17] BA Strange, RN Henson, KJ Friston, and RJ Dolan. Anterior prefrontal cortex mediates rule learning in humans. *Cereb Cortex*, 11:1040–1046, 2001.

[18] T Tyrrell. Computational Mechanisms for Action Selection. *PhD thesis, University of Edinburgh. Centre for Cognitive Science.*

[19] L Wang, R Kakigi, and M Hoshiyama. Neural activities during wisconsin card sorting test–MEG observation. *Brain Res Cogn Brain Res.*, 12:19–31, 2001.

MESO: Perceptual Memory to Support Online Learning in Adaptive Software*

E. P. Kasten and P. K. McKinley
Software Engineering and Network Systems Laboratory
Department of Computer Science and Engineering
Michigan State University
East Lansing, Michigan 48824
{kasten,mckinley}@cse.msu.edu

Abstract

Adaptive and autonomic systems often must be able to detect and respond to errant behavior or changing conditions with little or no human intervention. Decision making is a critical issue in such systems, which must learn how and when to invoke corrective actions based on past experience. This paper describes the design, implementation and evaluation of a perceptual memory system, called MESO, that supports online decision-making in adaptive systems. MESO uses clustering and pattern classification methods while addressing the needs of online, incremental learning.

Keywords: adaptive software, decision making, imitative learning, machine learning, pattern classification, perceptual memory.

1 Introduction

Increasingly, software needs to adapt to dynamic external conditions involving hardware components, network connections, and changes in the surrounding physical environment [5, 10, 23]. For example, to meet the needs of mobile users, software integrated into hand-held, portable and wearable computing devices must balance several conflicting concerns, including quality-of-service, security, energy consumption, and user preferences. Moreover, the promise of autonomic computing systems [16], that enable software to dynamically self-heal and self-manage, appeals to system's administrators and users alike.

In adaptive applications, future decisions should benefit from past experience, helping to improve the fitness of the software with respect to its environment and/or function. However, learning from experience requires that a system remember appropriate responses to sensed environmental context. *Perceptual memory*, a type of long-term memory for remembering external stimulus patterns [7], plays an important role in supporting context-aware, adaptive software.

The main contribution of this paper is to present a perceptual memory system, called MESO[1], that applies pattern classification and clustering techniques [6] to online, dynamic learning systems. The benefits of MESO include: rapid incremental training, rapid reorganization of an existing classifier tree, high recall accuracy, lack of dependence on a priori knowledge of adaptive actions, and support for data compression. Each of these benefits is important to constructing a dynamic decision-maker. For instance, incremental training enables a system to learn over time, addressing changing user requirements or environments. Limiting the impact of a growing population of training patterns can be addressed using data compression, reducing the memory and processor requirements needed for managing large data sets. We show how MESO can be used to enable software decision-making by presenting preliminary results of experiments with an audio streaming application that can *imitatively learn* [1, 13] how to adapt to changing network conditions. Due to space limitations, many details of this study are omitted here, but can be found in the accompanying technical report [15].

2 Background and Related Work

An adaptive software system must include a decision-making component to realize adaptive behavior. Learning-based approaches [11, 22] show substantial promise for addressing the needs of decision makers. By observing its

* This work was supported in part by the U.S. Department of the Navy, Office of Naval Research under Grant No. N00014-01-1-0744, and in part by National Science Foundation grants EIA-0000433, EIA-0130724, and ITR-0313142.

[1] The term MESO refers to the tree algorithm used by the system (Multi-Element Self-Organizing tree).

environment, an application can determine if it is operating within acceptable limits. For instance, if a network application perceives a high packet loss rate, it might interpret this condition as detrimental to quality of service and decide to increase the level of error correction. Once invoked, this response is evaluated and if acceptable, assimilated into the decision maker's experience for possible future use in similar situations. *That is, adaptive systems need to remember and recall past experience.*

We explore pattern classification and clustering methods for associating adaptive responses with observed or sensed data pertinent to application quality-of-service. The embodiment of this method is a *classifier* [6] that uses supervised learning to produce a model of environmental stimuli. Comprising the operation of a classifier are two basic functions: *training* and *testing*. During training, patterns are added to the classifier, enabling the construction of an internal model of the training data. Each training pattern is an array of continuous, binary or nominal values labeled as belonging to a specific, real-world category. Once the classifier has been trained, the system can be queried using an unlabeled pattern. The classifier then tests this pattern and returns a label, indicating the category to which the tested pattern most likely belongs.

Clustering and pattern classification research is an active field of study. Recently, a number of projects have addressed clustering and classification of large data sets, a characteristic of decision making for autonomic software. Tantrum et al. [21] consider model-based refractionation for clustering large data sets. Yu et al. [24] use an hierarchical approach to clustering using support vector machines (SVMs). Kalton et al. [14] address the growing need for clustering by constructing a framework that supports many clustering algorithms. Methods for online clustering and classification have also been explored [4, 17]. Such online methods might be used as the basis for a perceptual memory system similar to MESO.

Like our project, other works have explored the use of statistical methods and pattern classification and clustering techniques in learning systems. Some have used *developmental learning* algorithms that enable a system to learn online through interaction with the physical world. For example, Hwang and Weng [9] developed hierarchical discriminant regression (HDR) and applied it successfully as part of the developmental learning process in humanoid robots. In addition, Ivanov and Blumberg [11] developed the layered brain architecture [11], which was used for the construction of synthetic creatures, such as a "digital dog." That project used clustering and classification methods to construct perceptual models as part of the dog's developmental learning system. A notable aspect of that work is the use of compression schemes to limit the impact of large training sets on memory consumption and processing power requirements.

Imitative learning, where a learner acquires skills by observing and remembering the behavior of a teacher, has also been studied in the context of providing humanoid robots with motor skills. Amit and Matarić [1] used hidden Markov models (HMMs) to learn aerobic-style movements. The ability of the system to reconstruct motion sequences is encouraging, demonstrating the potential importance of imitative learning. Jebar and Pentland [13] conducted imitative learning experiments using a wearable computer system that included a camera and a microphone. A human subject was observed by the system during interactions with other people. The observed training data was used to train an HMM. Later the system was allowed to respond autonomously when presented with visual and audio stimuli, demonstrating a limited ability to reproduce correct responses. However, since learning by observing real human behavior is very complex, even limited recognizable response is significant and promising.

The key hypothesis of our project is that clustering and classification methods can be extended to support decision making in *adaptive software*. We focus in this paper on applications that operate over lossy wireless networks and thereby must accommodate periods of high packet loss. Since error correction or retransmission often consumes additional bandwidth and increases packet delay, applications must balance these concerns while correctly choosing a response for current conditions. By measuring the loss rate, bandwidth and other network and system behavior, a pattern can be constructed that enables a decision maker to "remember" an appropriate response.

3 MESO

As a first step in our study, we developed MESO, a perceptual memory system for adaptive software. At a basic level, MESO functions as a pattern classifier and can be used to incrementally classify environmental stimuli or other data while accomodating very large data sets. Indeed, in early experiments we used HDR [9] to classify environmental stimuli related to application quality of service. The insight gleaned from those experiments led to our design of MESO.

Basic Approach. As depicted in Figure 1, a unique feature of MESO's design is the use of small agglomerative clusters, called *sensitivity spheres*, that aggregate similar training patterns. In adaptive software, training patterns comprise observations related to quality of service or environmental context, such as network bandwidth or physical location. The quantity of training patterns collected while a system executes may be very large, requiring more memory and processing resources as new patterns are added to the classifier. However, many training patterns may be very similar, enabling their aggregation into a much smaller number of sensitivity spheres while helping limit resource consumption.

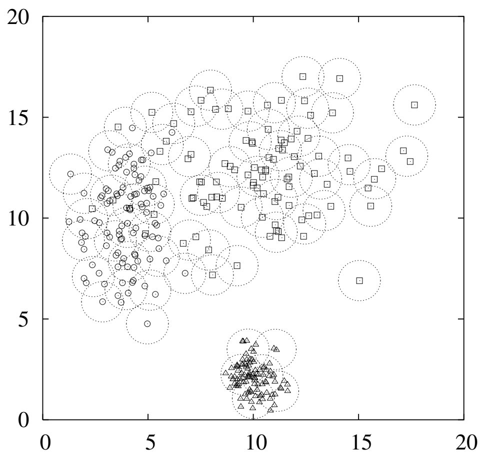


Figure 1. Sensitivity spheres for three 2D-Gaussian clusters. Circles represent the boundaries of the spheres as determined by the current δ. Each sphere contains one or more training patterns.

The size of the sensitivity spheres is determined by a δ value that specifies the sphere radius in terms of distance (e.g. Euclidean distance) from the sphere's center. A sphere's center is calculated as the mean of all patterns that have been added to that sphere. The δ is a ceiling value for determining if a training pattern should be added to a sphere, or if creation of a new sphere is required.

As with many other classifiers, MESO uses a hierarchical data structure, or tree, to organize training patterns for efficient retrieval. Figure 2 shows the organization of a MESO tree. The MESO tree is built starting with a root node that comprises the set of all sensitivity spheres. The root node is then split into subsets of similar spheres, producing child nodes. Each child node is further split into subsets until each child comprises only one sphere. Consolidating similar patterns into sensitivity spheres enables construction of a tree using only spheres rather than agglomerating individual patterns at tree nodes. Moreover, many clustering algorithms construct a tree by agglomerating individual patterns into large clusters near the root of the tree, and then split these clusters at greater tree depths. In such a structure, the tree must be reorganized using the training patterns directly. Conversely, MESO can be reorganized using only existing sensitivity spheres. The use of sensitivity spheres enables a MESO tree to be more rapidly reorganized than approaches that require direct consideration of training patterns.

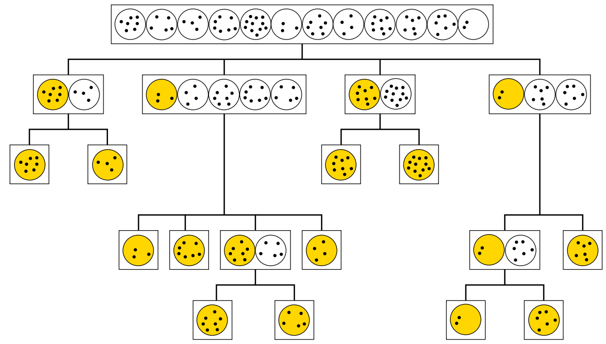


Figure 2. MESO tree organization. The rectangles are partitions and the shaded spheres are partition pivots. The root partition is split successively until a leaf is formed where a partition contains only one sphere.

As shown in Figure 2, sensitivity spheres are partitioned into sets during the construction of a tree. Each node of the tree comprises a collection of sensitivity spheres called a *partition*, defined as a subset of similar spheres. A partition may have one or more children, each comprising a subset of the parent's sensitivity spheres. A *pivot sphere* is selected for each partition and used to assign other spheres, nearest to the pivot, as members of the partition. Smaller partitions provide finer discrimination and better classification of test patterns. Moreover, the partitioning of sensitivity spheres produces a hierarchical model of the training data. That is, each partition is an internal representation of a subset of the training data that is produced by collecting those spheres that are most similar to a pivot sphere. At greater tree depths, parent partitions are split, producing smaller partitions of greater similarity. For each partition, classification proceeds by comparing a test pattern with each child's pivot and following the branch to the child containing the pivot that is most similar. In this way, the search continues at each tree depth. At a leaf node, a label is returned indicating the category to which the test pattern most likely belongs.

By leveraging this hierarchical structure, a relatively simple distance metric, even Euclidean distance, can be used to achieve high recall accuracy. As shown in Section 4, another advantage of this approach is that it requires a minimum amount of calculation, enabling rapid classifier training and testing. In addition, MESO supports incremental training, which enables construction of a MESO tree by adding one pattern at a time. As such, MESO can be trained and tested during concurrent interaction with users or other system components.

Sensitivity Sphere Size. An important consideration in building an efficient MESO tree is the appropriate value of

δ for constructing sensitivity spheres. For each data set, a δ value needs to be calculated. If δ is too small, training time increases dramatically. If too large, testing time becomes unacceptable. Moreover, data set compression requires a proper δ value to balance the tradeoff between compression rate and recall accuracy.

Since online learning is an incremental process, it is possible to adjust δ incrementally as MESO is trained. Key to incrementally calculating a good δ is determining a δ *growth function* that balances sphere creation with sphere growth. That is, rapid δ growth during early training produces few spheres, with very large δ's, that agglomerate many patterns. The result is a coarse-grained, inefficient tree where most testing time may be spent in a linear search of a single sphere. Conversely, overly slow δ growth produces a large number of very small spheres that contain only a few patterns. Having many small spheres is expensive, in terms of both space and time, to organize as a MESO tree.

MESO's δ growth function is defined as:

$$grow_\delta = \frac{(d-\delta)\frac{\delta}{d}f}{1+\ln(d-\delta+1)^2},$$

where d is the distance between the new pattern and the nearest sensitivity sphere. The $\frac{\delta}{d}$ factor scales the result relative to the difference between the current δ and d. Intuitively, the denominator of $grow_\delta$ limits the growth rate based on how far the current δ is from d. That is, if d is close to δ then δ will grow to be nearly d. However, if d is much larger than δ, the increase will be only a small fraction of $d-\delta$.

The sigmoid-curve activation function, f, inhibits sensitivity sphere growth when the number of spheres is small compared to the number of patterns. This function is defined as:

$$f = \frac{1}{2} + \frac{\tanh(\frac{3r}{c}-3)}{2},$$

where $r = \frac{spheres}{patterns}$ and c is a configuration parameter in the range $[0, 1.0]$. Increasing c moves the center of the sigmoid curve to the right, suppressing sphere growth and encouraging the production of new spheres. The $\tanh()$ function is the hyperbolic tangent.

Our $grow_\delta$ function balances sphere production with sphere growth, producing good spheres for a wide range of values for c. Only for very large values of c is growth inhibited sufficiently to significantly impact training time. The $grow_\delta$ function promotes the production of trees that are comparable with good choices for fixed δ values.

Compression. Adaptive systems often must continue to function for long periods while addressing changing user preferences and the sensed environment. Such an enormous amount of data requires substantial processor and storage resources, potentially inhibiting timely response by decision makers or impacting application performance. Thus, perceptual memory systems may need to "forget" less informative training samples in favor of important or novel observations. Compression techniques eliminate training patterns while attempting to minimize information loss.

In MESO, compression takes place on a per sensitivity sphere basis. That is, rather than trying to compress the entire data set using a global criterion, the patterns in each sensitivity sphere are compressed separately. Moreover, compression is further limited so that all existing pattern labels are not eliminated from a sphere. We implemented three types of compression, the evaluation of which is discussed in Section 4.

Means compression reduces the set of patterns in each sensitivity sphere to the mean pattern vector for each label. This is the most aggressive and simple of the compression methods. Moreover, the computational requirements are quite low.

Spherical compression is a type of boundary compression [11] that treats patterns on the boundaries between sphere's as most important to the classification of test patterns. For each sphere, the feature values are converted to spherical coordinates. Along a given vector from the sphere center, only those patterns farthest from the sphere center are kept.

Orthogonal compression removes all the patterns that are not used for constructing an orthogonal representation of a sphere's patterns. The idea is to keep only those patterns that are most important as determined by their orthogonality. Samples that represent parallel vectors in m-dimensional space are removed.

Complexity. Table 1 shows the space and time complexities for training MESO and some well-known clustering algorithms [12]. In this table, n is the number of patterns, k is the number of clusters and l is the number of iterations to convergence. Without compression, MESO has a worst case space complexity of $O(n)$, comparable to the shortest spanning path algorithm. MESO's memory consumption can be significantly reduced with compression.

Intuitively, time complexity for training can be considered in terms of locating the sensitivity sphere nearest to a new pattern and adding the pattern to that sphere. If a sufficiently close sphere cannot be found, a new sphere is created. Locating the nearest sphere is an $O(\log_q k)$ operation. This search must be completed once for each of n patterns. Each pattern must also be added to a sensitivity sphere, and k sensitivity spheres must be created and added to the MESO tree. This process yields a complexity of $O(n\log_q k)+O(n)+O(k)+O(k\log_q k)$ which reduces to $O(n\log_q k)$. Assuming an appropriate δ selection and a data set of significant size, MESO has a time complexity better than the leader algorithm.

Table 1. Space and time complexities for several clustering algorithms [12].

Algorithm	Space	Time
MESO	$O(n)$	$O(n \log_q k)$
leader	$O(k)$	$O(kn)$
k-means	$O(k)$	$O(nkl)$
ISODATA	$O(k)$	$O(nkl)$
shortest spanning path	$O(n)$	$O(n^2)$
single-line	$O(n^2)$	$O(n^2 \log n)$
complete-line	$O(n^2)$	$O(n^2 \log n)$

The search complexity for classifying a test pattern using MESO is $O(\log_q k) + O(\bar{s})$ for a balanced tree, where q is the maximum number of children per node, $\bar{s}$ is the average number patterns agglomerated by a sensitivity sphere, and k represents the number of sensitivity spheres produced. The $\bar{s}$ component represents the number of operations required to assign a category label once the most similar sensitivity sphere has been located. Thus, the worst case search complexity occurs when only one cluster is formed and the search algorithm degenerates into a linear search of $O(n)$. Conversely, a best case search complexity of $O(\log_q k)$ occurs when one sensitivity sphere is formed for each training pattern.

4 MESO Assessment

We evaluated MESO as a pattern classifier using several standard data sets in cross-validation experiments. The results illustrate the recall accuracy of MESO. We also compare MESO with the IND classifier [3] to provide a benchmark for MESO performance.

The Data Sets. Table 2 lists the data sets used to assess MESO. The number of patterns and features per pattern are shown for each data set. All but one of the data sets were retrieved from the UCI [2] and KDD [8] machine learning repositories. The ATT faces [20] data set was acquired from AT&T Laboratories Cambridge. Most of these data sets comprise continuous, integer or binary feature measurements. For instance, the Cover Type data set contains continuous features, measuring features such as elevation or slope, and binary values indicating whether a pattern is a particular soil type. However, the Mushroom data set consists entirely of nominal values encoded as alpha characters converted to their ASCII equivalent for processing by MESO. In contrast, the ATT Faces data set comprises image pixel values of human faces.

The Japanese Vowel data set requires further description. This data set comprises 640 time series blocks where each block consists of a set of records. Each record comprises 12 continuous measurements of utterances from nine male speakers. The 9,859 patterns are produced by treating each record as an independent pattern and randomizing the data set. As such, no understanding of utterance order is retained. Thus, the classification task is to identify the speaker of each utterance independent of its position in a time series.

Table 2. Data set sizes and feature counts.

Data Set	Size	Features	Classes
Iris	150	4	3
ATT Faces	360	10,304	40
Mult. Feature	2,000	649	10
Mushroom	8,124	22	2
Japanese Vowel	9,859	12	9
Letter	20,000	16	26
Cover Type	581,012	54	7

Experimental Method. We tested MESO using cross-validation experiments as described by Murthy et al. [19]. The experiment proceeds as follows:

1. Randomly divide the training data into p equal-sized partitions.
2. For each partition, train the classifier using all the data outside of the selected partition. Test the classifier using the data in the selected partition.
3. Calculate the classification accuracy by dividing the sum of all correct classifications by the total number of patterns tested.
4. Repeat the preceding steps t times, and calculate the mean and standard deviation for the t iterations.

In our tests we divide each data set into 10 equal-sized partitions and repeat the test 10 times. Thus, MESO is trained and tested 100 times for each mean and standard deviation calculated.

Experimental Results. Table 3 shows MESO's cross-validation accuracy and standard deviations for each of the data sets. For comparison, accuracy using the IND classifier is also presented. IND can be used to build a classifier in several modes. Here we include results using CART, ID3 and Bayesian IND modes. As shown, MESO is competitive with IND, exhibiting better accuracy for most data sets. Moreover, MESO has good accuracy for data sets of different size and pattern feature count. The NC designation indicates that IND could not complete a particular test. In the case of ATT Faces, lack of memory prevented IND from completing a data set encoding process, which must be completed before IND can be trained.

It is noteworthy that MESO shows high accuracy for the Mushroom data set, since this data set consists entirely of nominal values. Such pattern values have no comparative numeric value but rather indicate characteristics, such as cap shape, by name. IND addresses such nominal values explicitly by designation of some features as nominal. MESO does not explicitly address nominal values, but still

Table 3. MESO accuracy compared to IND.

Data set	MESO	IND CART	IND ID3	IND Bayesian
Iris	95.1%±0.0%	92.8%±0.3%	93.5%±0.7%	94.2%±1.1%
ATT Faces	94.1%±1.0%	NC	NC	NC
Mult. Feature	94.0%±0.6%	93.1%±0.6%	94.2%±0.2%	94.4%±1.1%
Mushroom	100.0%±0.0%	99.9%±0.0%	100.0%±0.0%	100.0%±0.0%
Japanese Vowel	91.9%±0.2%	82.3%±0.3%	84.2%±0.3%	84.7%±0.3%
Letter	90.1%±0.2%	84.4%±0.3%	87.9%±0.1%	88.6%±0.2%
Cover Type	96.0%±0.0%	93.9%±0.9%	95.2%±0.2%	94.4%±0.3%

Table 4. Results when using compression.

Data set		Uncompressed	Means	Spherical	Orthogonal
Iris	Accuracy%	95.1%±0.0%	95.7%±1.0%	96.0%±1.3%	96.2%±2.27%
	Compression%	0.0%	0.02%±0.0%	0.0%±0.0%	1.9%±0.07%
ATT Faces	Accuracy%	94.1%±1.0%	93.2%±1.1%	93.7%±1.5%	93.9%±1.7%
	Compression%	0.0%	0.0%±0.0%	0.0%±0.0%	0.0%±0.0%
Mult. Feature	Accuracy%	94.0%±0.6%	94.2%±0.6%	94.3%±0.5%	94.2%±0.5%
	Compression%	0.0%	0.3%±0.0%	0.3%±0.0%	0.3%±0.0%
Mushroom	Accuracy%	100.0%±0.0%	99.9%±0.0%	99.7%±0.0%	99.9%±0.0%
	Compression%	0.0%	90.3%±0.0%	71.9%±0.5%	90.1%±0.0%
Japanese Vowel	Accuracy%	91.9%±0.2%	81.0%±0.4%	90.1%±0.5%	80.6%±0.7%
	Compression%	0.0%	93.9%±0.1%	26.5%±1.3%	93.9%±0.0%
Letter	Accuracy%	90.1%±0.2%	87.8%±0.3%	90.2%±0.4%	87.5%±0.4%
	Compression%	0.0%	88.9%±0.2%	22.0%±1.0%	89.0%±0.2%
Cover Type	Accuracy%	96.0%±0.0%	82.5%±0.7%	95.1%±0.0%	82.1%±0.3%
	Compression%	0.0%	98.2%±0.2%	48.0%±0.9%	98.3%±0.1%

accurately classifies these nominal patterns. The high recall accuracy may be due to MESO's ability to capture m-dimensional structure. Determining how MESO addresses nominal and mixtures of nominal and continous values warrants further exploration.

Compression. Table 4 shows MESO results using the three data compression methods described earlier. All results were generated using cross-validation. Means compression provides high compression and good accuracy. This result can be attributed to applying compression on a per sensitivty sphere basis. As such, the ability of MESO to capture the m-dimensional structure of the training data can be leveraged to limit information loss during compression. For all compression methods, small data sets are compressed very little. Limited compression is expected since δ growth is inhibited during early training, and spheres contain only a few samples. However, since processor and storage usage is low for small data sets, compression is of limited importance. Larger data sets are compressed significantly while recall accuracy remains high.

5 The Network Application

We explored the use of MESO to support learning in adaptive software by applying it to the implementation of an audio streaming network application, called XNetApp, that can adapt to changes in packet loss rate in a wireless network. A stationary workstation transmits an audio data stream to a mobile receiver over the wireless network. Our goal is to enable the mobile device to adapt to the wireless packet loss encountered as a user roams about a wireless cell. One way to address the high loss rates of wireless channels is to insert redundant information into the data stream, enabling a receiver to correct some losses without contacting the sender for retransmission. This study focuses on *erasures* of packets. An (n, k) *block erasure code* [18] converts k source packets into n encoded packets, such that any k of the n encoded packets can be used to reconstruct the k source packets.

The XNetApp's decision maker uses MESO for "remembering" user preferences for balancing packet loss with bandwidth consumption. By autonomously modifying (n,k) settings and packet size, the decision maker can adapt the XNetApp based on current environmental conditions. In our experiments, the decision maker learns about a user's preferences through imitative learning. Thus, a user shows the XNetApp how to adapt to a rising loss rate by selecting an (n,k) setting with greater redundancy. If the new setting reduces the perceived loss rate to an acceptable level, the user reinforces the new configuration (e.g., by pressing a particular key), and the XNetApp "remembers" the sensed environment and current configuration by storing it using MESO. When operating autonomously, the decision maker senses current environmental conditions and calculates time-sampled and Fourier features, constructing a pattern. Using this pattern, the XNetApp queries MESO for a system configuration that most likely addresses current conditions. Then, the decision maker emulates the user's actions and adapts the XNetApp, changing the configuration to match that returned from MESO.

As shown in Table 5, 56 environmental features are sensed or calculated and used as input to the decision making process. The first 4 features are instantaneous measurements. Perceived features are observed from the application's viewpoint. That is, perceived packet loss represents the packet loss as witnessed by the application following error correction, while real packet loss is the number of packets actually dropped prior to error correction. For each of the first four features, 7 time-sampled measurements and 6 Fourier spectrums are calculated.

Table 5. Features used for training and testing the XNetApp.

#	Feature Description
1–4	Instantaneous measurements: bandwidth, perceived packet delay, perceived loss and real loss.
5–32	Time-sampled measurements: median, average, average deviation, standard deviation, skewness, kurtosis and derivative.
33-56	Fourier spectrum of the time-sampled measurements: median, average, average deviation, standard deviation, skewness and kurtosis.

6 Results

For testing, we studied the ability of the XNetApp to autonomously balance real packet loss and bandwidth consumption. The IP multicast protocol was used for transmission of the data stream. Data was collected as a user roamed about a wireless cell carrying a notebook computer running an XNetApp receiver.

We experimented with the XNetApp using a 1.5GHz AMD Athlon workstation for data transmission. A 500MHz X20 IBM Thinkpad notebook was used as a mobile receiver. All systems run the Linux operating system. We tested atop an 11Mb 802.11b wireless network as a user roamed about a wireless cell. The XNetApp was trained using an emulated loss rate in the range $[0, 0.6]$.

Figure 3 shows a comparison of real versus perceived loss observed as the user roamed about the cell. The XNetApp was able to adapt to changing real loss rates, and minimize application loss. The center plot depicts the redundancy ratio, defined as $\frac{(n-k)}{n}$, showing the responsiveness of the XNetApp on a real wireless network. Greater redundancy is required during periods of high loss. However, this redundancy consumes greater network bandwidth. The XNetApp did not simply choose a high (n, k) ratio, but changed parameters to correspond with the changing real loss rate.

Table 6 shows results from running cross-validation tests using the data acquired during XNetApp training. This

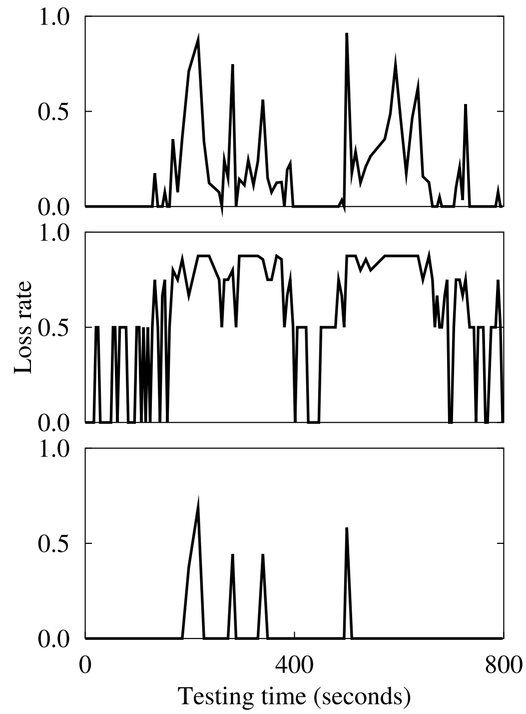


Figure 3. Comparison of real loss (top), redundancy ratio (center) and perceived loss (bottom).

data was produced during training for autonomous XNetApp operation on a real wireless network. The final training set contained 32,709 patterns in 10 classes. This table shows recall accuracy, with and without compression, helping quantify how well the XNetApp can be expected to imitate a user. In all cases, the XNetApp was highly accurate at "remembering" a user's preferences for balancing loss rate with bandwidth consumption.

Table 6. XNetApp results with and without compression.

Uncompressed	Accuracy%	94.1%±0.2%
	Compression%	0.0%
Means	Accuracy%	87.7%±0.2%
	Compression%	91.8%±0.1%
Spherical	Accuracy%	92.4%±0.7%
	Compression%	5.8%±0.2%
Orthogonal	Accuracy%	87.3%±0.4%
	Compression%	91.8%±0.1%

7 Conclusions

We have presented a perceptual memory approach, called MESO, that uses pattern classification and clustering techniques while addressing issues important to support online developmental learning. We showed that, when used as a pattern classifier, MESO can accurately classify patterns from standard data sets. We also implemented an application that imitatively learns how to make decisions through interaction with a user. Preliminary results show that the imitative learning approach, used by the XNetApp, has promise. We postulate that such software, which can be trained to make good decisions, may simplify the integration of software into new or pervasive computing environments.

Further information. This work is part of the RAPIDware project, which addresses the design of high-assurance adaptive software. Related papers and software packages can be found at `http://www.cse.msu.edu/rapidware`.

Acknowledgements. The authors would like to thank Professor Juyang Weng, Xiao Huang and Dave Knoester at Michigan State University for their contributions to this work.

References

[1] R. Amit and M. Matarić. Learning movement sequences from demonstration. In *Proceedings of the 2nd International Conference on Development and Learning*, pages 165–171, Boston, MA, June 2002.

[2] C. L. Blake and C. J. Merz. UCI repository of machine learning databases. http://www.ics.uci.edu/∼mlearn/MLRepository.html, 1998.

[3] W. Buntine. Tree classification software. In *Third National Technology Transfer Conference and Exposition*, Baltimore, MD, December 1992.

[4] K. Crammer, J. Kandola, and Y. Singer. Online classification on a budget. In *Proceedings of Advances in Neural Information Processing Systems (NIPS)*, Cambridge, MA, 2003.

[5] Proceedings of the Fourth International Workshop on Distributed Auto-adaptive and Reconfigurable Systems (DARES'04), held in conjunction with the 24nd International Conference on Distributed Computing Systems (ICDCS'04). Tokyo, Japan, March 2004.

[6] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification, Second Edition*. John Wiley and Sons, Incorporated, New York, NY, 2001.

[7] J. M. Fuster. *Memory in the Cerebral Cortex: An Empirical Approach to Neural Networks in the Human and Nonhuman Primate*. The MIT Press, Cambridge, MA, 1995.

[8] S. Hettich and S. D. Bay. UCI KDD archive. http://kdd.ics.uci.edu, 1999.

[9] W.-S. Hwang and J. Weng. Hierarchical discriminant regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), November 2000.

[10] Proceedings of the International Conference on Autonomic Computing (ICAC'04). New York, NY, May 2004.

[11] Y. A. Ivanov and B. M. Blumberg. Developmental learning of memory-based perceptual models. In *Proceedings of the 2nd International Conference on Development and Learning*, pages 165–171, Boston, MA, June 2002.

[12] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Computer Surveys*, 31(3):264–323, September 1999.

[13] T. Jebara and A. Pentland. Statistical imitative learning from perceptual data. In *Proceedings of the 2nd International Conference on Development and Learning*, pages 191–196, Boston, MA, June 2002.

[14] A. Kalton, P. Langley, K. Wagstaff, and J. Yoo. Generalized clustering, supervised learning, and data assignment. In *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 299–304, San Francisco, CA, August 2001.

[15] E. P. Kasten and P. K. McKinley. MESO: Perceptual memory to support online learning. Technical Report MSU-CSE-04-15, Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, April 2004.

[16] J. O. Kephart and D. M. Chess. The vision of autonomic computing. *IEEE Computer*, pages 41–50, January 2003.

[17] J. Kivinen, A. J. Smola, and R. C. Williamson. Online learning with kernels. In *Proceedings of Advances in Neural Information Processing Systems (NIPS)*, Cambridge, MA, 2002.

[18] A. J. McAuley. Reliable broadband communications using burst erasure correcting code. In *Proceedings of ACM SIGCOMM*, pages 287–306, September 1990.

[19] S. Murthy, S. Kasif, and S. Salzberg. A system for induction of oblique decision trees. *Journal of Artificial Intelligence Research (JAIR)*, 2:1–32, 1994.

[20] F. Samaria and A. Harter. Parameterisation of a stochastic model for human face identification. In *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, FL, December 1994.

[21] J. Tantrum, A. Murua, and W. Stuetzle. Hierarchical model-based clustering of large datasets through fractionation and refractionation. In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 183–190, Edmonton, Alberta, CA, July 2002.

[22] J. Weng and W.-S. Hwang. An incremental learning algorithm with automatically derived discriminating features. In *Proceedings Asian Conference on Computer Vision*, pages 426–431, Taipei, Taiwan, January 2000.

[23] Proceedings of the ACM Workshop on Self-Healing, Adaptive and self-MANaged Systems (SHAMAN), held in conjunction with the 16th Annual ACM International Conference on Supercomputing. New York, NY, June 2002.

[24] H. Yu, J. Yang, and J. Han. Classifying large data sets using SVMs with hierarchical clusters. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 306–315, Washington, D.C., August 2003.

Neuromodulation and open-ended development

Frédéric Kaplan
Sony Computer Science Laboratory Paris
Developmental Robotics Group
6 rue Amyot 75005 Paris France
kaplan@csl.sony.fr

Pierre-Yves Oudeyer
Sony Computer Science Laboratory Paris
Developmental Robotics Group
6 rue Amyot 75005 Paris France
py@csl.sony.fr

Recent discoveries showing a convergence between patterns of the activity in the midbrain dopamine neurons and computational model of reinforcement learning have lead to an important amount of speculations about learning activities in the brain [5]. In particular actor-critic reinforcement learning architectures have been presented as relevant models to account for functional and anatomical subdivisions in the midbrain dopamine system. Central to some of these models is the idea that dopamine cells report the error in predicting expected reward delivery and that this information is used in two different ways. The value system learned by the critic is associated with projections from the ventral tegmental area to the amygdala and the orbitofrontal cortex. The action selection scheme of the actor is thought to be realized by dopamine pathways initiated in the substantia nigra pars compacta and projecting to the striatium, thus controlling the choice of actions during cortico-striato-thalamo-cortical loops. This popular model has raised some amount of controversies (e.g. [1]) but has definitively shown that artificial learning paradigms could lead to interesting new interpretations of neurophysiological data.

We want to emphasize the inspiring role that research in developmental robotics can play in this context. One of the important goal of this new research field is to understand which dynamics can lead to open-ended development *i.e.* how robots can be designed to continuously learn new skills of increasing complexity. Paradigms based on conditioning and external rewards have difficulties to account for the active nature of development and exploratory behaviors. Children in the first years of their life actively choose in which learning task they take part, avoiding situations that are too difficult for them or that have become too predictable. This suggests the existence of intrinsic motivations structuring learning activities. Proposing models for such motivations has become a major challenge for developmental robotics.

We argue that in order to realize autonomous mental open-ended development, reinforcement learning models could be interestingly associated with an internal reward system based on the maximization of learning progress. Several preliminary computational and robotic experiments show how intrinsic motivations enable the development of novel behaviors of increasing complexity (e.g. [3, 4]). These new models naturally lead to investigate how the basic actor-critic paradigm could be extended to account for an architecture capable of evaluating its own ”learning progress”. Studies suggesting that dopamine responses could be interpreted as reporting ”prediction error” (and not only ”reward prediction error”) [2] may be taken into consideration for formulating new hypotheses about neural processes that could account for a system of intrinsic motivations.

References

[1] P. Dayan and W. Belleine. Reward, motivation and reinforcement learning. *Neuron*, 36:285–298, 2002.

[2] J-C. Horvitz. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*, 96(4):651–656, 2000.

[3] F. Kaplan and P-Y. Oudeyer. Maximizing learning progress: an internal reward system for development. In F. Iida, R. Pfeifer, L. Steels, and Y. Kuniyoshi, editors, *Embodied Artificial Intelligence*, LNAI 3139, pages 259–270. Springer-Verlag, 2004.

[4] P-Y. Oudeyer and F. Kaplan. Intelligent adaptive curiosity: a source of self-development. In Luc Berthouze, Hideki Kozima, Christopher G. Prince, Giulio Sandini, Georgi Stojanov, G. Metta, and C. Balkenius, editors, *Proceedings of the 4th International Workshop on Epigenetic Robotics*, volume 117, pages 127–130. Lund University Cognitive Studies, 2004.

[5] W. Schultz, P. Dayan, and P.R. Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1599, 1997.

A Model of Frame and Verb Compliance in Language Acquisition

Rutvik Desai*
Department of Computer Science and the Cognitive Science Program
Indiana University, Bloomington, IN 47405
rudesai@indiana.edu

Abstract

Researchers studying word learning have discovered that the syntactic frame in which a word appears plays an important role in the interpretation of the word, and this importance diminishes gradually with the increase in age. The interpretation of the sentences based on the frame and the verb is known as frame and verb compliance respectively. Here, a connectionist model is presented that learns a miniature language by associating sentences with the corresponding "scenes." In doing so, when the input to the network is changed to reflect the increasing linguistic experience of children, it exhibits a shift from frame to verb compliance. It is argued that these phenomena can be attributed to the increasingly combinatorial linguistic experience and representations that change with learning, and it is not necessary to invoke specialized mechanisms or principles.

1. Introduction

Children learn new words rapidly. A common-sense explanation for vocabulary acquisition is that word meanings are learnt by observing real-world contingencies of their use. The meaning of *jump* is learnt from noticing that it occurs in the presence of jumping events. However, this simple explanation has several difficulties when attempting to account for acquisition of meaning of all words. Many of these problems are listed in [1]: (a) This theory fails to account for the fact that children with radically different exposure conditions (*e.g.*, the blind and the sighted) acquire much the same meanings, (b) many verbs are used for the same events and only provide a perspective on an event (*e.g.*, *chase* and *flee*), and (c) many verbs only differ in the level of specificity at which they describe single events (*e.g.*, *see*, *look*, *orient*).

In light of these problems, it has been suggested that children use another rich source of information, namely the syntactic context in which the words occur. This proposal is known as *syntactic bootstrapping* [1, 2, 3]. According to this hypothesis, children can use the knowledge of syntax to predict meanings of words. The learner observes the real world situations and also observes the language structures in which various words appear. If there is a correlation between meanings and a range of syntactic structures, the meaning (or some components of the meaning) of an unknown word can be predicted when it appears in a familiar structure.

1.1. Verb Compliance and Frame Compliance

One way to study the effect of syntax on the acquisition of word meaning is to use familiar words in a different or incorrect syntactic context and examine the effect on the interpretation of the word. For example, we can insert a transitive verb in an intransitive frame and examine how children interpret the sentence. If children are still learning about a verb, then they may more readily accept its occurrence in an incorrect frame. They are more likely to reject an incorrect frame when they have fully acquired the verb. If children interpret the sentence in accord with the frame, they are said to be *Frame Compliant*. If the interpretation fits more with the verb, they are *Verb Compliant.*

Frame and Verb Compliance are interesting for another theoretical reason. While children's verb use is overwhelmingly correct, a major exception to this appears somewhere around the age of 3. As reported by Bowerman [4, 5], children sometimes use verbs in incorrect sentence frames, as in **Don't fall that on me* (to protest the impending dropping of an object by someone). Thus, children overgeneralize, *e.g.*, they use a verb transitively when only intransitive use is allowed, or vice versa. Children must learn eventually which uses are "licensed" for which verbs. For example, they must learn that *sink* can be used either transitively or intransitively, but *fall* and *go* allow only noncausal interpretation. How children overcome these overgeneralizations is a major question in language acquisition. This question is essentially the same as asking why children become Verb Compliant at some stage. When children show Verb

*Currently at the Department of Neurology, Medical College of Wisconsin, Milwaukee, WI 53226.

Compliant behavior, they have sufficient confidence in their knowledge of verb meaning and syntax that they reject contradictory cues, which is exactly the requirement for eliminating overgeneralizations.

Now we look at some empirical evidence for compliance effects.

1.2. The Data

Naigles and colleagues [6, 7] conducted experiments involving the approach described above. They asked 120 children, from 2.5 to 12 years of age, as well as adults, to enact grammatical and ungrammatical sentences using "Noah's Ark" and wooden toy animals as props. Ungrammatical sentences were constructed by placing transitive verbs (*bring*, *take*, *push*, *put*) in intransitive frames (e.g., *T*he lion puts in the ark*, **The zebra brings*). Similarly, intransitive verbs (*come*, *go*, *fall*, *stay*) were inserted in transitive frames (e.g., **The elephant comes the giraffe*). The children's enactment was deemed to be Frame Compliant if they modified the meaning of the verb to conform to the frame in which it was encountered (e.g., the elephant pushing or carrying the giraffe). It was considered Verb Compliant if they followed the restrictions of the verb (*e.g.*, the elephant moving independently of giraffe).

Their overall results indicated that younger children, especially the 2-year-olds, were more Frame Compliant, enacting the ungrammatical sentences according to the demands of the frame and altering the meaning of the verb. They allowed the novel frames to influence the interpretation of the familiar verbs. Older children, and especially the adults, were more Verb Compliant, following the restrictions of the verb and repairing the sentence. Children at the intermediate ages were en route to the adult state, showing intermediate levels of Frame and Verb Compliance.

Similar experiments have been conducted with children with Down Syndrome (DS) [8]. The linguistic skills of children with DS are split in an interesting way. Relative to their syntactic knowledge (often measured by measured MLU or auxiliary use) their vocabulary growth is advanced. It was reported [8] that children with DS who had a "vocabulary age" of 6 years were syntactically like 3-year-olds. While children with DS were more Frame Compliant than their chronological-age mates, they also exhibit the move from Frame to Verb Compliance. Adolescents with DS show more Verb Compliance than gradeschoolers with DS. Thus, with the advance in syntactic knowledge, DS children also move toward Verb Compliance.

In this paper I present a connectionist model that attempts to explain the mechanisms by which this shift occurs. First, a network is presented that learns a miniature language by associating simple sentences to the corresponding "scenes." The network's behavior with respect to the compliance effects is then examined. Then, various theories of compliance and the implications of network's behavior are discussed. We end with a discussion of the nature of representations and input in the network.

2. The Network

The architecture of the network is shown in Figure 1. It contains recurrent connections in the hidden layer as in a Simple Recurrent Network [9] to handle temporal sequences of words. Recurrent connections on the output layer make it easier for it to remember what has been already learned from the earlier portion of the sentence.

Then input to the network consists of sentences or noun phrases (henceforth called "utterances") describing one or two objects and optionally an action, generated by the grammar shown below:

S → NP | NP1 | NP is IV | NP1 are IV | NP is TV NP
NP → DET N | DET SIZE N
NP1 → NP and NP
N → boy | girl | dog | mouse
IV → jumping | dancing | running | walking
SIZE → large | small
TV → pushing | holding | hugging | kicking
DET → a

One can divide the utterances generated by this grammar into five basic types: (a) N, (b) NN, (c) NV, (d) NNV, and (e) NVN. With optional adjectives describing the size, utterances such as *a girl and a big dog are jumping* or *a small dog and a big mouse* are obtained. These utterances are presented to the network sequentially, one word at each time step. Words are represented in a localist manner by turning on a single bit in the input layer. Also, *ing* is treated as a separate word, with the assumption that it can be discerned from the word stem as a separate unit. An end-of-utterance marker, STOP, is presented after the last word of each utterance, at which point all context units are reset.

On the output or the semantic end, the descriptions of scenes corresponding to the input utterance are presented as a 30-bit fixed-width vector. There are two slots for objects, and one for the action or the event taking place. Each object slot is divided into two slots of 4 and 6 units each, which represent the attribute large (1100) or small (0011) and type of object respectively. In the 10-bit event slot, the first 4 bits indicate whether the action is causal or noncausal (with activations 1100 and 0011, respectively), and the remaining 6 bits describe other features of the action. A distributed representation for each individual object and event is generated by turning on 3 randomly chosen bits in its slot. If each bit is viewed as representing a feature, this creates representations with partially overlapping features. The slots for an attribute, object, or action not described in the scene are set to 0.

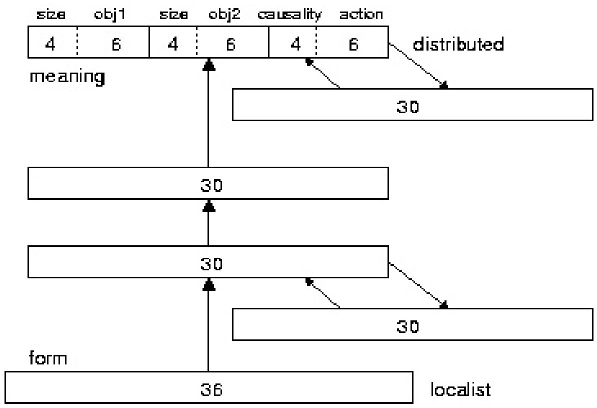


Figure 1. The network architecture.

3. Comprehension

We first test if the network is capable of performing the basic task of producing the correct scene corresponding to an input utterance. The set of all utterances generated by the grammar was probabilistically divided into two parts, one for training and the other for testing generalization. There is significant variability in the total number of different types of utterances generated by the above grammar: There are 12 utterances of type (a), 144 utterances of type (b), 48 utterances of type (c), and 576 utterances each of type (d) and (e). Hence, these utterances were included in the training set with differing probabilities: 1.0 for type (a), 0.4 for type (b), 0.7 for (c), and 0.2 for (d) and (e). To give more representation to utterances that have a lower frequency, and to ensure that they are not overwhelmed by other more frequent patterns, type (a) utterances were included thrice in the training set while type (c) utterances were included twice. The entire scene is presented as the target for every word. If there is a large difference in the type or token frequency between different types of utterances (*e.g.*, many more transitives than intransitives), this can lead to local optima, since the network attempts prediction from incomplete information. The purpose of the chosen probabilities is only to maintain a significant representation for each of the utterance and verb types in the input; other than that the precise values are not critical to the results.

The network was trained using backpropagation on the utterances in the training set. The initial weights were sampled from a uniformly random distribution between -0.2 and 0.2. The complete target was held constant for the duration of the entire utterance. This ensures that no words are given a special status and encourages the network to process the words as soon as they arrive. A learning rate of 0.0005 and no momentum were used. The weights were updated at the end of each epoch. Training was continued till there was no significant improvement in the error.

To assess the performance of the network, if the activation of an output unit was less than 0.5, it was considered OFF, and it was taken as being ON otherwise. An utterance was declared to be processed correctly if, at the end of the utterance, all output units had the desired ON or OFF activations. With this criterion, in a average of 5 runs, 100% accuracy was achieved on a training set of 322 (different) utterances and 96% of utterances were processed correctly in the remaining 1034 utterances of the testing set, which the network had not seen during training. The network was, then, largely successful in this task of producing semantics given an utterance, or comprehension. Next, we look at the experiments regarding frame and verb compliant behavior in the network.

4. Frame and Verb Compliance in the Network

To qualitatively simulate the increasing vocabulary and linguistic experience of children, the network was trained in stages with increasing numbers of nouns, rather than with the entire vocabulary as in the basic comprehension task described in the previous section. Four transitive and four intransitive verbs were used in all stages. The number of nouns used at each stage was increased gradually from 1 to 7. No adjectives were used. The set of utterances at each stage was again divided probabilistically into training and testing sets. Utterances in the testing set were not part of the training set at any stage. The network started with the weights from the previous stage and was trained to near-perfect accuracy on the training set using a learning rate of 0.001.

Two types of ungrammatical utterances were generated to test the network. The first was an NVN transitive sentence with a known intransitive verb (e.g., *a dog is dancing a boy*) while the second was an NNV intransitive sentence with a known transitive verb (e.g., *a dog and a boy are holding*). Two transitive and two intransitive verbs were chosen and four sentences were generated with each verb using two nouns. Transitive verbs were inserted in intransitive sentences, and vice versa. We are interested in examining whether the network interprets these verbs in incompatible frames as depicting causal events or noncausal ones. Recall that there are four units in the network's output indicating causality, where the pattern 1100 stands for a causal meaning while 0011 stands for a noncausal interpretation. To assess the network's response, a variable δ is defined as the mean activation of the first two units minus the mean activation of the last two units. A positive δ indicates a transitive response, while a negative δ suggests an intransitive interpretation of the verb. The mean value of δ, calculated from the eight transitive and intransitive sentences at the end of each stage, is shown in Figure 2.

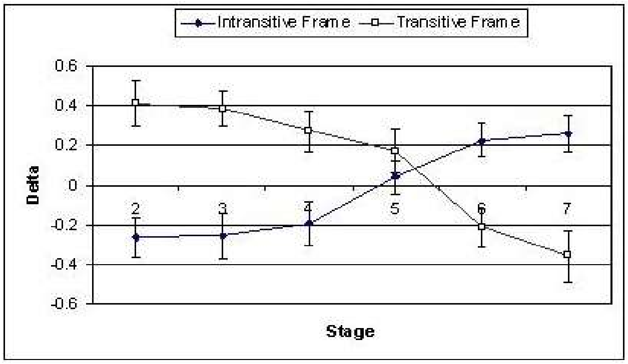


Figure 2. The value of the parameter δ across different stages. A positive δ implies a causal interpretation of the verb while a negative δ indicates noncausal interpretation.

The conflict arising from the mismatch between the frame and the verb is indicated by the values of δ. All four units receive some activation in most cases. In the initial stages, however, the frame tends to "win," as indicated by the higher activation of the units for causality (positive δ) in the case of transitive frame and units for noncausality (negative δ) in the case of intransitive frame. In other words, we get a Frame Compliant response in earlier stages with a small number of nouns. This behavior changes gradually with the increase in the number of nouns. With 7 nouns, there is still a conflict due to the mismatch, but now the response is more in accordance with the type of the verb. For the transitive verb, the output is closer to 1100 and for the intransitive verb it is closer to 0011, suggesting a Verb Compliant response.

5. Discussion

Frame and verb compliant behavior is closely related to the well-known overgeneralization errors made by children, and their recovery from those errors. We can consider some of the theories related to overgeneralization and frame and verb compliance, and ask what the implications of the model are.

5.1. Maturation

A maturation-based account is offered in [10]. Very briefly, verbs become organized into semantic subclasses known as *narrow range subclasses* as their representations are refined. The semantics of the verb class determines whether the verbs allow alternations (*e.g.*, causative and noncausative use in transitive and intransitive frames) or not. When the representation of a verb matches that of another verb in the same subclass that is known to alternate between causal and noncausal, the former is allowed to alternate as well. For example, motion verbs that encode path (e.g., *bring*, *take*, *go*) can be used either causally or noncausally, but not in both ways. On the other hand, motion verbs that encode manner, like *roll* and *bounce* can be used both transitively and intransitively. Overgeneralization occurs because a verb is used in the same manner as other semantically similar verbs in a subclass.

The shift to Verb Compliance may occur because the verb representations are elaborated to the extent that they have formed grammatically relevant narrow range subclasses. Some verbs no longer allow causal interpretation because they do not fit the semantic specification of the subclasses that are causal. At the time of puberty, those subclasses of verbs for which there has been no evidence of alternation become fixed or "closed." After that, no new information about the verb is accepted, resulting in Verb Compliant behavior. For example, since *come* and *go* do not encode manner of motion, they do not match the specification of the alternating subclass of motion verbs (that includes *roll* and *bounce*). This subclass is closed at maturation, so *come* and *go* no longer allow causative interpretation.

As pointed out in [6, 8], there are factors other than age that appear to affect compliance behavior and present a serious problem for this account. If the maturation-based account was correct, one would expect to observe an across-the-board shift from Frame to Verb Compliance, for all verbs and frames. But this is not the case. Some verbs elicit more Frame Compliance than others. For example, in the Naigles, Fowler, and Helm (1992) study, in the NVN frame, *come* and *go* elicited significantly more Verb Compliance than *stay* and *fall*. *Stay* and *fall* also differed from each other significantly. In the NV frame, *bring*, *take*, and *put* showed significantly more Frame Compliance than *push*. Secondly, some frames induce Frame Compliance to a later stage than others. For example, the shift to Verb Compliance for the NV frame is effectively complete at age 5. On the other hand, even 12-year-olds and adults continue to exhibit Frame Compliance for their NVNPN frame. Intransitive frames shifted earlier than transitive frames. Furthermore, the move towards Verb Compliance can occur at different ages, anywhere between 2 to 7 years of age.

In the study of children with DS [8], it was found that although these children were more Frame Compliant than their chronological age mates, they exhibit this shift also. Since their maturational progress is dissociated, one would expect that prepubescent children with DS will be no less Verb Compliant than their adolescent counterparts if the maturational account is correct. This appears not to be the case and children with DS also move to become more Verb Compliant.

5.2. Mutual Exclusivity

Another proposal about recovery from overgeneralizations is termed The Mutual Exclusivity Principle (also called Contrast, Uniqueness, or Pre-emption) [5, 11, 12, 13]. In brief, this principle is that children will allow only one lexical entry to occupy a semantic niche. When two words are determined to have similar meanings, one of them is pre-empted and removed from the lexicon. For example, causative *come* is basically equivalent to *bring*. Using Bowerman's [5] example, during the period in which overgeneralized (causative) *come* is frequent in production, *bring* is practically nonexistent. When *bring* becomes more frequent eventually, the causative *come* declines. This can explain why some verbs elicit Verb Compliance. For example, transitive *bring* and *take* pre-empt causative uses of *come* and *go* respectively. We do not discuss all the details of Mutual Exclusivity here (see [5, 6] for a more detailed discussion), but note that while Mutual Exclusivity may have some role to play in recovering from overgeneralizations, it does not account for all the effects found in the data. For example, it does not explain why intransitive *push* causes Verb Compliance earlier than intransitive *bring* or *take*. This principle also does not work for all the verbs, since for some verbs, it is difficult to find a similar meaning verb that can pre-empt its use in the right way.

5.3. Lexical Knowledge

A different account based on lexical knowledge is offered in [7]. This account relies on children's knowledge of individual verbs. Children's conjectures about verb meanings are refined by ongoing events as well as the structures in which they appear. At early stages of vocabulary acquisition, open-minded children assume that not all structures have as yet been heard and therefore certain properties of verbs (such as whether they encode causality) may be unknown to them. In this case they make use of the structural information provided by the frames. At some point, however, older children and adults feel warranted to believe that all the relevant information about the meaning has been obtained. Then they would perceive a novel structure as simply ill-formed, causing verb compliant behavior.

This theory can explain various effects in the data well. For example, the shift towards Verb Compliance is a function of individual verbs and individual frames because different amount of knowledge is accrued for them due to their differing frequencies in the input. While this account is supported by data, some important details remain unclear. One might ask where the so-called "open-mindedness" in the initial stages and the confidence about the meaning at later stages come from. After hearing a verb in a certain number of contexts, exactly what makes a child more or less open-minded to accept new meanings? An answer is not offered in [7], but one possibility is to invoke some type of innate parameter or threshold that allows children to determine whether a certain amount of experience with a verb is enough to warrant confidence in the meaning of that verb.

5.4. Lexical Knowledge and Innate Principles

More recently, Lidz, Gleitman, and Gleitman [14] offer an explanation that involves both lexical knowledge and innate principles. It is best summarized by the following quote:

> "The deduction of verb meaning based on an analysis of the surface structure is a learning heuristic. The learning device is asking itself, in effect: Assuming Principles [the *Theta Criterion* and the *Projection Principle*], what could be the meaning of the verb now heard, such that these principles projected this observed (surface) structure for it? Such a deductive procedure will be invoked only when the learner does not have secure knowledge of the verb in question." (p. 37)

Thus, when children know that they have secure knowledge of a verb, they assume that anyone who uses it otherwise has misspoken, resulting in Verb Compliance. Otherwise they invoke innate principles that state that participants in an event will line up one-to-one with noun-phrases in the clause [15], and make a decision based on that.

This account has a problem similar to that of the Lexical Knowledge account: Exactly how do children determine whether they have "secure knowledge" of the verb? An all-or-nothing decision about knowing a verb also seems to be involved in this account. A child either doesn't know the verb (and invokes the principles), or does (and rejects the frame). However, the knowledge of a verb is likely to be graded. Subjects even show different compliance effects for the same verb in different frames. The spontaneous remarks of subjects in [6] indicate that both children and adults are ambivalent about the sentences they are asked to act out. They have conflicting information and varying degrees of confidence in their knowledge, and hence it seems unlikely that are cleanly choosing one path over the other.

5.5. Emergence

The network simulation suggests an alternative to the previous theories that does not rely on innate principles or overt determination of the knowledge of a verb. It can be viewed as an extension of the Lexical Knowledge theory. The explanation of various effects, such as some verbs inducing more verb compliance than others and some frames continuing to elicit frame compliance till a later stage, is identical to the one offered by the Lexical Knowledge theory: They are input driven. The verbs that are experienced frequently and in multiple contexts, tend to elicit Verb Compliance early.

This account differs from the Lexical Knowledge theory the in *use* of the input. The network's Frame Compliance in the initial stages is a result of representations being more context-bound at those stages. The context-sensitivity of the representation is in turn a consequence of the memory, or the number of connections in the network and the number of patterns stored. With a certain amount of memory available, it is possible, and easier, to simply memorize entire syntactic-semantic patterns as wholes as long as there are relatively few patterns. This gives rise to context-bound representations in the hidden layer (see [16] for a more detailed discussion of context bound representations). Since frames and context have more weight in the representation, the network gravitates towards an interpretation based on the frame since the incompatibility of one word has a relatively small effect. The “open-mindedness” of the network, so to speak, in early stages is a result of the fact that the context plays an important role in early representations.

As more words are encountered in varied contexts, it is no longer feasible to store entire patterns individually because it entails excessive demands on memory. As a result, various components in an utterance are separated, and the words are gradually de-contextualized. As verbs (along with other words) attain their own separate representation, the effect of context is diminished and the relevant form/meaning mappings are strengthened. The words in various frames are encountered with many other words describing causal and noncausal events, and therefore do not exert significant influence with respect to the causality of an event. The network learns the remaining consistent correlations between groups of verbs and causal or noncausal events. This results in Verb Compliant behavior. Stated another way, the network exhibits open-mindedness in early stages and confidence about the meaning in later stages, but it is not a result of reasoning about the number of contexts in which words were encountered, or the confidence in the knowledge of verb meaning. Rather, Frame and Verb Compliant behavior is an “epiphenomenon,” or an emergent consequence of the task, the input, the learning procedure, as well as the size and architecture of the network.

5.6. The Nature of Representations

The claim here is that the shift in compliance behavior is due to the diminishing role of context with increasing linguistic experience. This may raise two questions: (1) What independent evidence is there that the network's early representations are context-bound and more become context-free later? (2) What is the evidence that context plays a large role in *children's* early representations, and this role diminishes with age?

Representations in the network

One way to gain some insight into the representations used by the network is to test its generalization ability at various stages. If the network has memorized or rote learned entire training patterns, it should perform poorly when the same words are encountered in a different context. On the other hand, it it has learned context-free word-level form/meaning mappings then they should be recognized regardless of the sentential context. The generalization performance for five networks on the respective testing sets at the end of different stages is shown in Figure 3.

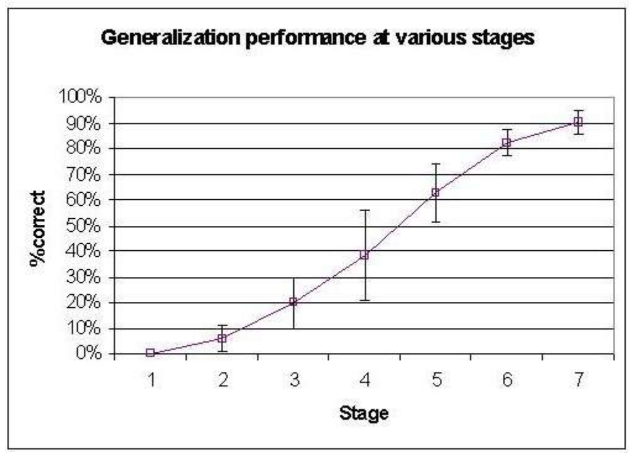


Figure 3. Performance on the testing set across stages.

The results suggest that in early stages, the network has not developed individuated representations of various components of the input utterances, and they develop after increasingly combinatorial input. St. John [17] found a similar pattern of results in his story-processing network. The network performed well at generalization only when it was exposed to highly combinatorial input, with many participants in the various slots in the representation of events.

Another method of examining the nature of network's representations involves probing the hidden unit activations directly as the utterances are processed. One way to examine the activation space is Principal Components Analysis (PCA), which can measure the underlying dimensions of variation in the hidden unit activations. If two complex patterns are classified by a network with a simple strategy such as, for example, the presence or absence of one or two features, then one would expect few significant principal components in the hidden layer space because there would be few underlying dimensions of variation. For a more complex decision-making process involving combinations of multiple features at different time-points, one would expect a high-dimensional principal component space as more components would be needed to account for the variance.

Here, the set of utterances used in the first stage is passed through the network after each stage, and the 30 unit activations of the first hidden layer are recorded after each word. PCA is performed on this 30-dimensional set of vectors after each stage.The results of this analysis, as mean eigenvalues of five networks, are shown in Figure 4.

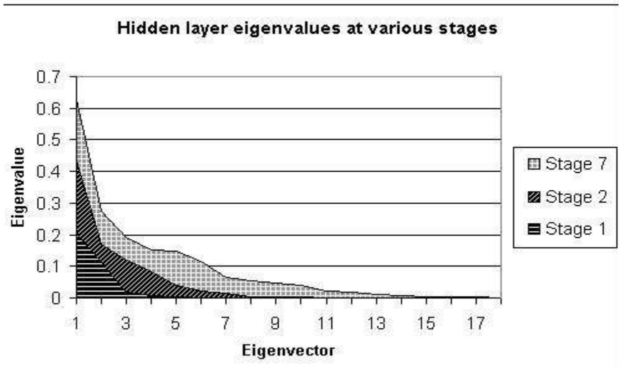


Figure 4. The mean eigenvalues of the hidden layer activation space at different stages, when processing the same set of utterances.

After the first stage, only 3 of the 30 eigenvalues are 0.01 or more on average. This number increases to 7 after stage 2, and to 13 after the last stage. The same set of utterances is processed differently, resulting in more eigenvectors of higher magnitude. During the processing of an utterance, in the early stages, once the network finds cues sufficient to distinguish one item or utterance from another, it does not need to analyze it further since it can predict the rest. The hidden layer activations remain relatively constant after that, resulting in few dimensions of variation overall. With context-free representations, activations change with each input word, as the network accomodates the newly available information, resulting in more dimensions of variation [16].

Representations in children

From studies of productions of children, there is evidence that children's early representations are context-bound too, and they gradually become de-contextualized. A compelling collection of evidence that children’s early language is highly item-based, or based on specific linguistic items and expressions they comprehend and produce, is provided in [18, 19, 20]. There is also evidence that early productions are bound not only to linguistic context, but also to non-linguistic context such as actions, social routines, or salient events that occur frequently in the experiences of the young child [21].

5.7. The Nature of the Input

An important feature of the network is the changing input through different stages. The size of the training set, as well the vocabulary, is increased gradually.[1] However, one can view the training sets at different stages as being of the same size, since training involves cycling repeatedly over the same set. The stages are differentiated by increasing type frequency and decreasing token frequency, not by the number of utterances.

With respect to the increase in vocabulary, older children are more likely to have experienced higher number of words and word combinations simply by virtue of having more linguistic experience over a longer period of time, even if they experience the same linguistic environment. Although the vocabulary is explicitly changed here, it can be viewed as gradually sampling more utterances from a fixed set. The approach in [22] is noteworthy, where selective attention is used on a fixed input to model children's increasing experience with words to achieve the same effect as explicit addition of words to the vocabulary in a model of past-tense learning.

Secondly, there is evidence that the linguistic environment of children is not constant but changes with age. Child-directed speech (CDS) to younger children is syntactically and semantically simplified, is less diverse, and contains more high-frequency words [16, 20, 23, 24, 25]. Caretakers restrict their vocabularies when talking to young children [23, 24], and the type-to-token ratio increases with age in CDS [23].

The simulation used in this work is admittedly small, and the utterances are simple. It should be noted that the sentences used in the experiments with 2- and 3-year-old children are also simple, and an attempt was made to use identical or very similar utterances in the simulations. It would also be desirable to use natural CDS from a database, rather than an artificial grammar. The use of raw CDS, however, is more suitable in paradigms where no form/meaning associations are involved, and the task of the network is (typically) to predict the next word in the input. One can choose utterances from the CDS that are suitable for the task and relevant to the phenomenon under consideration, but that defeats the purpose of using natural CDS to some extent.

6. Conclusions

[1] Note that this is unlike early models of inflectional morphology, which were criticized for sudden changes in the input.

A connectionist network was presented that learns to comprehend utterances of a miniature language by associating them with the corresponding scenes. When the training set of the network is varied to reflect the increasing linguistic experience of children, the network exhibits frame and verb compliance effects. The account of the shift from frame to verb compliance resulting from the network's behavior is similar to the lexical knowledge theory in that these effects are attributed to increasing experience with words in varied contexts. However, this account does not entail explicit rules or reasoning mechanisms. There is nothing in the network designed specifically to produce these effects; they emerge as a result of the network attempting to efficiently accomplish the task of associating utterances with scenes. The network's behavior is a consequence of various low-level parameters such as the number of weights, the number of units in each layer, the number of input patterns, and so forth. Children's compliance behavior, similarly, may change automatically when they have developed relatively context-free representations and sufficiently strong individual form-meaning mappings so that conflicting information is ignored, without making any explicit decision to do so. This work supports the view that specific mechanisms or behaviors can arise as a result of the nature of the task and the general characteristics of the tools employed to perform the task, without the presence of dedicated mechanisms.

References

[1] Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 3-55.

[2] Gleitman, L. & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science, 1,* 31-35.

[3] Landau, B. & Gleitman, L. (1985). *Language and Experience: Evidence From the Blind Child.* Cambridge, MA: Harvard University Press.

[4] Bowerman, M. (1977). The acquisition of rules governing "possible lexical items:" Evidence from spontaneous speech errors. In *Proceedings of Research in Child Language Development, 13*, 148-156.

[5] Bowerman, M. (1982). Starting to talk worse: Clues to language acquisition from children's late speech errors. In S. Strauss (Ed.), *U-shaped Behavioral Growth.* New York: Academic.

[6] Naigles, L., Fowler, A., & Helm, A. (1992). Developmental shifts in the construction of verb meanings. *Cognitive Development, 7*, 403-428.

[7] Naigles, L., Gleitman, H., & Gleitman, L. (1993). Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), *Language and Cognition: A Developmental Perspective.* Norwood, NJ: Ablex.

[8] Naigles, L., Fowler, A., and Helm, A. (1995). Syntactic bootstrapping from start to finish with special reference to Down Syndrome. In M. Tomasello, and W. Merriman (Eds.), *Beyond Names of Things: Young Children's Acquisition of Verbs.* Hillsdale, NJ: Erlbaum.

[9] Elman, J.L. (1990). Finding structure in time. *Cognitive Science, 14,* 179-211.

[10] Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure.* Cambridge, MA: MIT Press.

[11] Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition.* Hillsdale, NJ: Erlbaum.

[12] Clark, E. V. (1991). Acquisitional principles in lexical development. In S. Gelman and J. Byrnes (Eds.), *Perspectives on Thought and Language: Interrelations in Development*. Cambridge: Cambridge University Press.

[13] Markman, E. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization.* Cambridge: Cambridge University Press.

[14] Lidz, J. Gleitman, H., & Gleitman, L. (2001*). Kidz in the 'Hood: Syntactic Bootstrapping and the Mental Lexicon*. (Tech. Rep. IRCS-01-01). Philadelphia, PA: University of Pennsylvania, Institute of Research in Cognitive Science.

[15] Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lecture*s. Foris: Dordrecht.

[16] Desai, R. (2002). Item-based Language Learning in Children and Connectionist Networks. In *Proceedings of the 38th Annual Conference of the Chicago Linguistic Society*.

[17] St. John, M. F. (1992). The story gestalt: A model of knowledge-intensive processes in text comprehension. *Cognitive Science, 16,* 271-306.

[18] Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences, 4*(4), 156-163.

[19] Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209-253.

[20] Cameron-Faulkner, T., & Tomasello, M. (2003). A construction based analisys of child directed speech. *Cognitive Science, 27*, 843-873.

[21] Barrett, M.D. (1986). Early semantic representations and early word usage. In S.A. Kuczaj & M.D. Barrett (Eds.), *The Development of Word Meaning.* New York: Springer.

[22] Jackson, D., Constandse, R. M., and Cottrell, G. W. (1996). Selective attention in the acquisition of the past tense. In *Proceedings of the 18th Cognitive Science Conference.* San Diego, CA: Erlbaum.

[23] Chapman, R. S. (1981). Mother-child interaction in the second year of life: its role in language development. In R. L. Schiefelbusch and D. Bricker (Eds.), *Early Language: Acquisition and Intervention.* Baltimore: University Park Press.

[24] Hu, Q. (1994). A study of some common features of mothers' vocabularies. In J. L. Sokolov and C. E. Snow (Eds.), *Handbook of Research in Language Development Using CHILDES.* Hillsdale, NJ: Erlbaum.

[25] Pine, J. (1994). The language of primary caregivers. In C. Gallaway, and B. J. Richards (Eds.), *Input and Interaction in Language Acquisition.* Cambridge, UK: Cambridge University Press.

On-Line Cumulative Learning of Hierarchical Sparse n-grams

Karl Pfleger *
Computer Science Department
Stanford University
kpfleger@cs.stanford.edu

Abstract

We present a system for on-line, cumulative learning of hierarchical collections of frequent patterns from unsegmented data streams. Such learning is critical for long-lived intelligent agents in complex worlds. Learned patterns enable prediction of unseen data and serve as building blocks for higher-level knowledge representation. We introduce a novel sparse n-gram model that, unlike pruned n-grams, learns on-line by stochastic search for frequent n-tuple patterns. Adding patterns as data arrives complicates probability calculations. We discuss an EM approach to this problem and introduce hierarchical sparse n-grams, a model that uses a better solution based on a new method for combining information across levels. A second new method for combining information from multiple granularities (n-gram widths) enables these models to more effectively search for frequent patterns (an on-line, stochastic analog of pruning in association rule mining). The result is an example of a rare combination—unsupervised, on-line, cumulative, structure learning. Unlike prediction suffix tree (PST) mixtures, the model learns with no size bound but using less space than the data. It does not repeatedly iterate over data (unlike MaxEnt feature construction). It discovers repeated structure on-line and (unlike PSTs) uses this to learn larger patterns. The type of repeated structure is limited (e.g., compared to hierarchical HMMs) but still useful, and these are important first steps towards learning repeated structure in more expressive representations, which has seen little progress especially in unsupervised, on-line contexts.

1 Introduction

For purposes here, *on-line learning* means the learner sees data a little at a time and cannot remember all data nor repeatedly iterate over it. This paper takes as a given the importance of on-line learning and will not dwell on motivating it. It suffices to note that (a) on-line learning is critical for long-lived autonomous agents in complex environments, (b) in many specific applications data sizes overwhelm storage capacities [1], and (c) a vast array of literature has argued for on-line learning (e.g., [1, 2, 3, 4]). To quote [2], "in a broad sense, online learning is essential if we want to obtain *learning* systems as opposed to merely *learned* ones." This paper further assumes the importance of *Cumulative learning* (also called layered or hierarchical learning), which involves using the results of prior learning to facilitate further learning (e.g., building new knowledge structures from experience by combining previously learned structures).

We present a model, *hierarchical sparse n-grams*, for on-line cumulative learning of frequently occurring patterns from unstructured unsegmented data, along with a related subcomponent model, *sparse n-grams*. Language, music, spatial configurations, event chronologies, action sequences, and many other types of data exhibit repeated substructure. For intelligent autonomous agents, identifying repeated substructure or frequent patterns in data has many benefits, including prediction of unseen information, improving short-term memory capacity and thus information processing capability generally, and facilitating communication and further learning. In addition, frequent patterns serve as building blocks for higher-level knowledge representation. General methods for identifying frequent patterns would greatly aid automated selection of higher-level representational units that are tuned to the environment. Frequent patterns in unbounded, unsegmented data streams can be identified on-line by simply noticing and remembering patterns that occur often, and using these to search for larger patterns that would otherwise have been more difficult to notice. Putting this into practice involves several challenges, as we will explain. We must emphasize that though our models are based on n-grams, our main goal is not model-class specific, but is the (under-investigated) on-line learning of frequent patterns in data and the cumulative use of existing patterns to help identify larger ones. Our purpose is not to tweak a slight performance improvement on

*The author performed this research at Stanford, but is presently at Google.

standard NLP, compression, or speech community n-gram benchmarks (most of which are batch and all of which are non-cumulative).

This paper is organized as follows. Section 2 presents sparse n-grams, the component model. This section demonstrates the benefits of the sparse representation, explains the method for on-line structure learning in the presence of sparseness, and introduces the mathematical form of novel probability estimates that form the basis for inference in both the basic and hierarchical models. Section 3 introduces hierarchical sparse n-grams and explains how the hierarchical nature of these models dramatically improves both the probability estimates (inference) and pattern selection (structure learning). Experimental results demonstrate that the models do an impressive job of finding frequent patterns demonstrated by their environments despite very sparse sampling of huge pattern spaces. Section 4 discusses related work from a diverse set of research communities. In general, the related models share certain properties but address different fundamental goals. The paper concludes with Section 5.

2 Sparse n-grams

n-grams are considered state-of-the-art for problems involving discrete sequences [5, 6]. Exhaustive n-grams store an occurrence count for every pattern of width n. We introduce *sparse n-grams* (or SNGs), which keep only some counts. They trade prediction quality for space, but wider SNGs can out-predict narrower exhaustive n-grams with the same number of patterns, making the sparse models useful when data is plentiful relative to storage. Similar n-grams trained exhaustively and then pruned [7] require increased storage during training and are not easily adaptable to on-line learning. Our goal is not merely to save space during training, however, or to improve n-gram models, but to investigate on-line, hierarchical learning of frequent patterns from unsegmented data. SNGs are a necessary subcomponent of the hierarchical models introduced below, but they are also interesting themselves and many issues are more easily introduced with them.

2.1 Sparse joint distributions and inference

For our purposes here, the learning and inference tasks are the unbounded, sequential analogs of those for standard fixed-feature IID unsupervised learning. A query is a sequence of any width with symbols for some positions specified, some targets to be predicted, and the rest missing. Each position is a symbol from alphabet A. n-grams keep a count C for each of the $|A|^n$ patterns and estimate probabilities as C/N (or a smoothed version, e.g., $\frac{C+1}{N+|A|^n}$), where N sums all counts. An SNG with counts for k patterns, called *tracked* patterns, for fixed k, is a k-n-gram. It estimates tracked patterns as above and distributes the remaining probability mass evenly among the untracked patterns to complete the joint distribution[1], which can be conditioned to make arbitrary predictions.

To demonstrate that sparseness can improve prediction, we trained sparse and exhaustive n-grams on book1 (a Thomas Hardy novel) of the Calgary corpus [6] (stripped of non-letters).[2] Batch training was used here just to demonstrate the potential of SNGs. We examined several inference patterns (forward, backward, and middle prediction and variants with missing values). Accuracy (0/1 loss) and cross-entropy were measured on a held-out test set of the last 10,000 characters. We expected sparseness to impair prediction, expected it to hurt cross-entropy more than accuracy (for which fine distinctions between unlikely events is less important), and expected increased width to mitigate the degradation.

Cross-entropy results were mixed, with n-grams outperforming SNGs in a few cases but not in others. Under accuracy, n-grams were always outperformed by a one-wider SNG with the same number of patterns (Figure 1(a)). This shows that the advantage of an extra predictor variable can outweigh the degradation caused by sparseness. Figure 1(b) shows that there is little degradation until the sparseness becomes severe. Also, any increase in storage can increase SNG performance by increasing k, whether or not it is enough to use a wider exhaustive model (increasingly important as n or $|A|$ rise). Since sparseness does not create inefficiencies[3], SNGs can be useful in some circumstances.

Sparseness degrades accuracy less than cross-entropy since accuracy depends only on correctness of the mode of the model's conditional distribution of the targets, not fine distinctions. The following is a more detailed explanation. When predicting with a full set of n–1 given symbols, the conditional distribution derives from renormalizing $|A|$ entries of the joint distribution. If a well-trained SNG includes at least one of these, it will get the correct mode. Only when *all* are absent will accuracy suffer, and this is the least likely case if the model includes frequent patterns.

[1] n-grams are often stored in a conditional form, but can be undirected. They are interchangeable for exhaustive but not sparse versions. We use the undirected here for simplicity, more flexible sparseness (see Section 4), predicting equally well backwards, and generalizability to 2D data (a long-term goal). ADtrees [8] allow effi cient prediction even from the undirected representation.

[2] Our experiments have concentrated on letters as symbols (standard in the compression community [6]) instead of words. We believe the structure of letters, phonemes, musical notes, etc., is more appropriate for studying identifi cation of frequent patterns (and might play a role in discovering words in the fi rst place). Words have longer-range interactions. Despite their word-level success, n-grams seem more suited to letters (also claimed by [9]). Nonetheless, our models are not limited to the letter level and should work well for word n-grams as well.

[3] Representation of SNGs uses a tree. Each leaf stores the count of a corresponding n-tuple ($O(n)$ lookup as for n-grams).

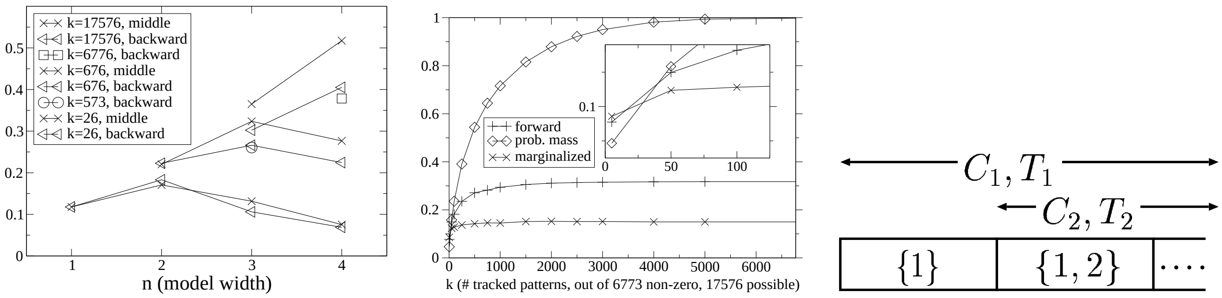


Figure 1. (a) k-n-gram accuracy vs. n for $k = 26^i$ (i=1–3) for backward and middle prediction. Forward behaves similar to backward and is omitted for clarity. Random guessing on this 26-class problem gets only 0.038 accuracy, and majority (a 1-gram) 0.118. The leftmost points of the equi-k lines are exhaustive n-grams, all outperformed by wider SNGs to the right. When k is limited to the number of non-zero counts in the exhaustive (n–1)-gram, performance (circle and square) still exceeds the exhaustive model. (b) k-n-gram accuracy (and total tracked probability mass) vs. k for n=3, for forward (`ggt`) and marginalized forward (`gmt`) prediction (Given, Missing, Target). Degradation is severe only for small k. (c) The counts kept. Time moves to the right. The counts C_2 and T_2 are only counted over the time from the addition of the 2nd pattern.

Cross-entropy will degrade if *any* of these entries are replaced by the untracked average. When there are missing values, the conditional distribution derives from renormalizing sums of entries. Being larger, the tracked entries will usually dominate these sums so that if the model contains any of this much larger set of entries, accuracy will often not suffer. Indeed, Figure 1(b) shows that marginalized prediction degrades relatively less with sparseness. Under a more general loss function that penalizes incorrect answers but allows "skipping questions" for a smaller penalty, sparseness hurts even less since it is obvious to the model when it does not have the relevant counts to make a good guess. These concerns are important because accuracy (or more general loss functions allowing abstention) is more appropriate than entropy-based measures in some situations.

2.2 On-line learning of frequent patterns

The model cannot know the frequent patterns *a priori*, so on-line learning requires structural selection of which counts to include. Structure learning is harder on-line. Batch structure learning often works by optimizing a global metric such as a Bayesian posterior or MDL score over all data. This cannot be done on-line[4]. We identify frequent patterns on-line using a stochastic subset search that incrementally adds and removes patterns. Patterns are added randomly, but only when appearing in training (the add probability for a new instance is a small fixed value). Tracked patterns with low observed frequencies are discarded. Thus, infrequent patterns shuffle in and out of the model. The key that makes learning feasible is that frequent patterns, once added, are identifiable as such. A period of uncharacteristic rarity might cause a truly frequent pattern to be discarded, but if it is truly frequent, it will be added again. As a pattern is tracked longer, the chance of large enough anomalous drought to trigger removal becomes vanishingly small.

Addition and removal can operate independently or one can trigger the other to keep k fixed. If only the least frequent pattern is ever removed, the model will converge. Unfortunately, we do not know the true pattern frequencies but only slowly converging estimates. We created a new version of *Hoeffding races* [11, 1] to decide when to remove a pattern and which to remove. Maron and Moore [11] introduced Hoeffding races to the machine learning community for model selection in supervised learning, and Domingos and Hulten [1] used these races for on-line decision tree construction. We adapt the technique for sparse n-gram pattern removal and introduce a slight improvement specific to the low probabilities of this application.

Hoeffding races are useful when there are several competing entities about which statistical evidence is accumulating and the goal is to find an extreme (high or low) valued example. In the present context, this allows one to smoothly trade off how much less frequent the least-frequent pattern is with how converged the estimates are so that a bigger gap between the frequencies can be acted upon while the error bars are still high, while a close race will require more convergence. Traditional Hoeffding races use the Hoeffding bounds, a member of the broader class of Chernoff bounds. We use a tighter version of the Chernoff bounds, appropriate when the probabilities are close to zero, as they

[4]…unless the model encodes sufficient statistics to summarize past data, which is not possible with structure changes that grow new parameters. Sufficient statistics can be encoded where changes only shrink the models, as in [10] where segmented data allows direct data incorporation steps. Our techniques do not depend on segmented data. Note that adding growth steps (reversing bad merges) breaks the on-line nature of [10].

are for n-grams. The other difference from traditional Hoeffding races is that in this case the race is perpetual, with new competitors continually introduced after the race has started. The exact rule that sparse n-grams employ can be abstractly stated as follows: The model removes a pattern as soon as it can identify one that it is "reasonably sure" is "close" to being the least frequent (probably approximately the least frequent). Specifically, the model finds the pattern whose upper bound (highest possible probability in the confidence interval) is lowest. Then it finds a different pattern whose lower bound is lowest. If the difference between the upper bound of the former and the lower bound of the latter is less than a tolerance parameter τ, the model drops the former pattern. The continual addition and removal process thereby converges to a stochastic equilibrium in which more frequent patterns are more likely to be included.

For each pattern pat_i, the model maintains a count C_i of occurrences since inclusion. Since patterns are added at different times, empirical frequencies have different reliabilities and the model cannot simply use C_i/T as the estimate for every pattern, so it keeps the total number T_i of instances seen since adding pat_i. See Figure 1(c). C_i/T_i is less reliable for new patterns, so probability estimates must be weighted based on age to insure that after adding a new pattern, its probability under the model only slowly diverges from the untracked average $p_{\text{untracked}}$, asymptotically approaching C_i/T_i. To do this, the model keeps patterns sorted by T_i ($T_i > T_{i+1}$). For pat_1, $p_1 = C_1/T_1$, but p_2 has an additional term estimating missed occurrences from when only pat_1 was tracked, called episode_1: $p_2 = \frac{1}{T_1}\left[C_2 + (T_1 - T_2)(1-p_1)\frac{1}{|A|^n - 1}\right]$, where $(T_1 - T_2)(1-p_1)$ estimates the number of non-pat_1 occurrences in episode_1. Appealingly, this is a linear interpolation of $\frac{1-p_1}{|A|^n-1}$ (i.e., $p_{\text{untracked}}$ before adding pat_2) and $\frac{C_2}{T_2}$, weighted by T_2. The general formula adds terms for each episode:

$$p_i = \frac{1}{T_1}\left[C_i + \sum_{j=1}^{i-1}(T_j - T_{j+1}) \cdot q_j \cdot \frac{1}{|A|^n - j}\right] \tag{1}$$

where $q_j \equiv (1 - \sum_{l=1}^{j} p_l)$. These approach C_i/T_i and can be computed in a linear sweep.

We have also investigated Bayesian and EM-based approaches to this problem. Bayesian predictions can be derived for a slight problem variant that assumes knowledge of the counts for each individual episode (a complex likelihood function involving nested interacting sums makes the original problem too difficult). It can be operationalized by approximating the count breakdowns by episode. An EM approach attempts to improve the uniform distribution on Equation 1's right side by instead using the resubstituted solution distribution. The resulting equations can be solved algebraically rather than requiring iteration to a fixed point. Under the modified problem, the solution is equivalent to the Bayesian solution. We implemented this and the above approach and in practice they behave almost identically. The next section improves on them. See [12, Sec.7.3] for more on these approaches and their relationships, and for the tighter "race" bounds.

3 Hierarchical Sparse n-grams

Hierarchical sparse n-grams (HSNGs) consist of multiple sparse n-grams of consecutive widths (possibly an exhaustive n-gram as the smallest) and dramatically improve two aspects of the fixed-width models: the probability calculations and pattern selection. A single tree accommodates all patterns by storing counts in non-leaves for smaller-width patterns (still $O(n)$ lookup). This provides the same variance advantages as traditional multiwidth exhaustive n-grams but uses a new method for combining the models that is more elegant than linear interpolation or backoff models. Slowly growing versions of HSNGs can incrementally add greater widths during on-line training. These models smoothly "surf" the bias/variance curve by fitting parameters for ever-widening joint probability distributions, continually decreasing bias error (by widening the joint) as well as variance error (by converging better parameters).

3.1 Improved Probability Estimation

To estimate untracked occurrences, the EM approach above used the resubstituted width-n distribution, which *is unreliable precisely when needed, when too little data has been seen for the n distribution to have converged*. Smaller widths converge exponentially faster, so each level n in HSNGs uses the $n-1$ distribution to estimate the missing data. Equation 1 becomes:

$$p_{D_n}(\text{pat}_{n,i}) = \frac{1}{T_{\max}}\Big[C_{n,i} + \sum_{j=0}^{i-1}(T_{n,j} - T_{n,j+1}) \cdot q_{n,j} \cdot p_D(\text{pat}_{n,i} \mid \neg\text{pat}_{n,1}, \ldots, \neg\text{pat}_{n,j})\Big] \tag{2}$$

$p_{D_n}(\text{pat}_{w,i})$ is the probability of $\text{pat}_{w,i}$ (the i-th oldest width-w pattern) under the n distribution. $q_{l,j} \equiv (1 - \sum_{p=1}^{j} p_{D_l}(\text{pat}_{n,p}))$, is as before the rest of the probability mass without the j oldest patterns (at level l). For HSNGs, $D = D_{n-1}$ instead of D_n (the EM approach) or uniform D_0 (Equation 1). The conditioning is just renormalization:

$$p_{D_{n-1}}(\text{pat}_{n,i} \mid \neg\text{pat}_{n,1}, \ldots, \neg\text{pat}_{n,j}) = \frac{p_{D_{n-1}}(\text{pat}_{n,i})}{q_{n-1,j}} \tag{3}$$

This all reduces to Equation 1 in the base case.

Of course, the $n-1$ distribution does not specify n-tuple probabilities directly. The trick is to use estimates based on the order-($n-1$) Markov assumption:

$$\begin{aligned} &P(x_1, x_2, \ldots, x_n) \\ &\quad = P(x_1, x_2, \ldots, x_{n-1}) \cdot P(x_n \mid x_1, x_2, \ldots, x_{n-1}) \\ &\quad \approx P(x_1, x_2, \ldots, x_{n-1}) \cdot P(x_n \mid x_2, \ldots, x_{n-1}) \\ &\quad \approx \frac{P(x_1, x_2, \ldots, x_{n-1}) \cdot P(x_2, \ldots, x_{n-1}, x_n)}{P(x_2, x_3, \ldots, x_{n-1})} \end{aligned} \tag{4}$$

The approximation from the first to second lines above, $P(x_n \mid x_1, x_2, \ldots, x_{n-1}) \approx P(x_n \mid x_2, \ldots, x_{n-1})$, is in a sense the best estimate given only order-(n−1) knowledge. This gives us a way to effectively "widen" an n-gram, using two lookups at width $n-1$ and one lookup at width $n-2$ to determine a width-n probability.

For untracked patterns ($i > k_n$, where k_n is the number of width-n patterns), width n falls back on width n–1 directly (and renormalizes):

$$p_{D_n}(\text{pat}_{n,i}) = p_{D_{n-1}}(\text{pat}_{n,i})\frac{q_{n,k_n}}{q_{n-1,k_n}} \quad (5)$$

Using Equation 1, the probabilities for all patterns can be computed in a linear sweep. This works despite that each pattern's probability involves the sum of a linear number of terms since the terms are the same from one pattern to the next. That is not true of Equation 2 due to the rightmost term. Thus, straightforward computation of all probabilities for a level would require computation quadratic in the number of patterns. However, using the expansion of Equation 3 the portion of the rightmost term that changes from pattern to pattern can be factored out, preserving the ability to compute probabilities for all patterns in the model in a linear sweep.

With cached probabilities for the tracked patterns, looking up the probability for an untracked pattern requires three lookups to narrower submodels using Equations 4 and 5. Any of these may in turn be absent, requiring three more lookups. Nonetheless, querying the widest model for a specific probability takes $O(n^2)$ rather than time exponential in n since some of the submodel lookups can be handled by the same traversals down the tree. (Lookup requires up to n tree traversals, each of at most n steps.) n is typically very small relative to the overall size of the model. Furthermore, many inference patterns that require lookup of several patterns can be handled more efficiently since many patterns will be accessible from one traversal. (E.g., looking up all possible single-symbol extensions of a pattern requires only one traversal from the root.) This is a worst case. Normally, each narrower level contains more of the necessary patterns, making typical lookups linear in n. Other optimizations are also possible (see [12, Sec.8.2.2] for further optimizations or more detailed derivations of the equations here.)

Predictions are made using the largest width. Lower-order information filters up through the submodel calls. Smaller widths have larger Ts and more fully converged estimates. Dynamically adding a new level will not jar the distribution as a model having mostly small T values will smoothly fall back on the robust narrow information except where it identifies strong wide patterns in the data. Estimates both within and across widths are naturally weighted by their reliabilities. Authority smoothly transitions to greater widths as more data is seen.

3.2 Improved Pattern Selection

Fixed-width models must search a large space blindly. HSNGs use smaller known-frequent patterns to bias new

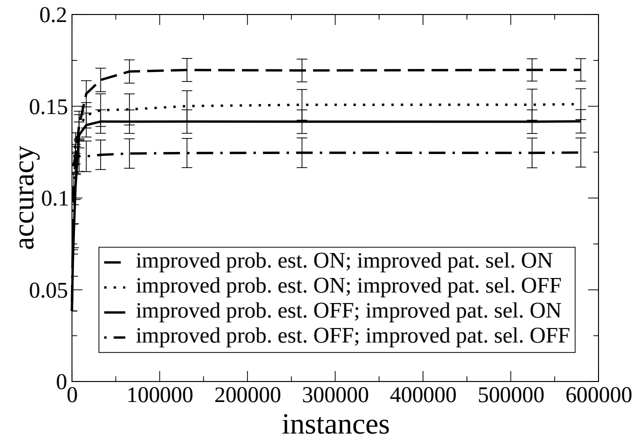


Figure 2. Learning curves for a sparse 2-gram using a pre-trained 1-gram submodel for probability estimation, pattern selection, both, or neither on book1. Each submodel use yields benefi t.

selections. Specifically, the probability for adding any new n-tuple, rather than a small constant, is proportional to its estimate based on D_{n-1}. This can be viewed as a refinement of a simple strategy of combining existing frequent patterns to yield larger likely-frequent patterns. This refinement automatically takes into account all of the possibly many ways to parse the new pattern into existing patterns and the frequency estimates for all of these existing patterns. Since each level seeks frequent patterns, a bias based on their probabilities functions as a bias for compositional chunking of known patterns. This is the key to the cumulative aspect of HSNG learning—the learning at narrower levels is not only combined with the results of learning at the next wider level for better prediction (as in traditional multiwidth n-gram combinations) but also *directly enables* the structure learning at this wider level, which is critical as we will see in the results below.

HSNGs can use fixed k_ns (the number of patterns at each level) or the k_ns and number of levels can grow slowly as more data is seen by applying the add and remove rules independently and always considering addition of a new widest pattern (which creates a new level). Since wider patterns generally have lower absolute probabilities and vastly more patterns, higher levels converge more slowly.

3.3 Results

Figure 2 shows that each use of submodels above provides significant benefit (which can be greater when combined). Similar results are seen for larger n but with the improvement from probability estimation being more dominant when used with fully-converged submodels (because the widening approximation becomes more accurate for larger n). To test frequent pattern identification, hierarchical models were trained on several large datasets using slowly growing k_ns (and numbers of levels). The following tables shows k_n and the top 5 patterns by model probability for each n after training on 30 million letters of Reuters newswire text (1994 section of North American News corpus with non-letters removed.)

n	2	3	4	5	6
k_n	367	1540	1131	433	142
	th	the	tion	ation	nation
	he	ing	said	ofthe	saidth
	in	and	nthe	inthe	ingthe
	er	ion	ther	tions	ations
	an	ent	dthe	ingth	aidthe

n	7	8	9
k_n	38	11	7
	saidthe	official	president
	esident	resident	ternation
	residen	presiden	residento
	nationa	thenatio	theminist
	preside	tsaidthe	andforthe

The learned chunks are reasonable substrings for this text. With spaces removed, frequent word, sub-word, and super-word patterns were all mixed together. The following table shows how many of the most frequent 100 and 1000 patterns (by true corpus frequency) were learned by the model for the smaller widths (larger widths were still far from converging).

	Reuters			
n	2	3	4	5
k_n	367	1540	1131	433
top100	100	100	99	75
top1000	n/a	910	560	238
unique	670	12556	121799	626465

Similar results were achieved with speech data (the TIMIT corpus) and DNA data (chromosome 22 of the human genome, about 34 million base pairs). The DNA results are shown below. Note that our purpose in testing on different data types was not to show particular performance on problems of traditional importance in each of these domains but to demonstrate that our technique for learning compositional patterns and estimating their prevalence is general enough to handle several different kinds of data.

	chromosome 22						
n	5	6	7	8	9	10	11
k_n	886	2094	1629	737	279	114	55
top100	100	100	99	68	44	28	19
top1000	886	991	669	258	128	57	32
unique	1024	4096	16384	65536	261726	1.0mil	3.5mil

The models do an excellent job of identifying the frequent patterns. For Reuters $n = 5$, for example, the model had added 434 patterns and removed 1 (it was still early in convergence at this level), but included 75 of the top 100 out of 626,465 unique 5-tuples that occurred (out of 11.8 million possible)—a dramatic improvement over chance guessing. Further, over half of the patterns added at the 5 level are in the top 1000 (top 0.16%). This is even more impressive considering the $n = 4$ level contained fewer than 1% of the

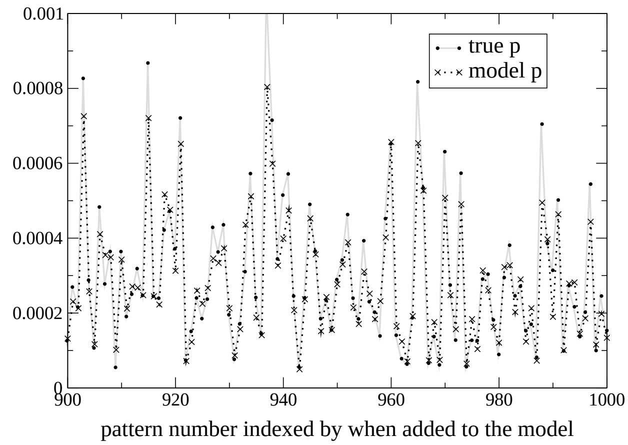


Figure 3. Comparison of the Reuters-trained hierarchical model's estimate vs. the true probability for the 900-1000th oldest 3-grams. The model tracks true probabilities well.

data's unique 4-tuples. This demonstrates that the narrower submodels do an impressive job enabling successful pattern selection at the wider levels (just as those levels will in turn do for even wider levels).

We examined how well the model's probabilities matched true corpus frequencies. Standard summarizations such as relative entropy of the distributions are inappropriate for the sparse representation. Also this metric is dominated by errors in low probabilities, which are specifically sacrificed by these sparse models. Thus, we just inspected all the values directly, and all but the very newest patterns matched quite well. Figure 3 shows a small sample section of patterns. The majority of patterns were older than the sample shown and matched even better than those shown.

4 Related Work

Pruning already-batch-trained n-grams has been investigated [7]. Our sparseness is similar to count cutoffs (widely used in practice). Our models, however, grow rather than scale back their storage and never need excessive space during training. That literature also discusses tradeoffs between space and predictive performance, but this research is the first we are aware of to explicitly address the tradeoff between memory reduction and increased predictive accuracy from increased n and the first to explain that accuracy (0/1 loss) suffers less from missing counts than cross-entropy.

Prediction suffix trees (PSTs) [9, 13, 14] are multiwidth sparse models based on unbalanced trees. Each node represents a string and stores a conditional distribution for the next symbol given the string as preceding symbols. PSTs are grown by adding nodes that (a) represent frequent strings and (b) have distributions sufficiently different from their parent's or have a descendant with this property. Our models differ in several ways. PST learning contains forward-directed inductive bias since it chooses patterns in terms of forward distributions. Thus, PSTs do not make equally good use of information on both sides of a prediction target. Second, our undirected representation allows for greater, more flexible sparseness. PSTs store a full conditional distribution at each node. HSNGs store one probability. A PST with a depth-d node must store counts for all $|A|$ $(d{+}1)$-tuple extensions, while an HSNG can represent any subset of these. The counts saved can be used elsewhere, improving other predictions. Most importantly, PST learning does not utilize repeated substructure. It is not more likely to extend a model by existing sequences with high counts from elsewhere in the tree. PST mixtures [14] use incremental updates, but not on-line structure learning in our sense. Either they are given a depth bound and add all sufficiently small strings that occur, or they have no bound and add all strings that occur. The latter simply reorganizes and remembers all data (often impractical). In the former case, the model cannot grow to recognize patterns wider than the bound and may require memory exponential in the bound. We desire a middle ground where the model grows *slowly*, but *without bound*. A model should be able to include any pattern demonstrated strongly enough by the data.

ADtrees [8], structures for efficient access to discrete multivariate data, are both similar to SNGs and complementary in that they can speed up SNG inference. ADtrees trade space for speed while SNGs trade prediction performance for space. The combination can balance all three. In addition, ADtrees use a sparse representation synergistic with that of SNGs. Applying ADtrees to SNGs allows much greater ADtree pruning (more speed for less space) when compared with application to exhaustive n-grams since all untracked patterns can be pruned.

Models equivalent to SNGs were used as hierarchical Bayesian priors [15]. We have moved the sparseness from the prior to the model itself.

Using the submodel to bias pattern selection is analogous to pruning in association rule mining [16, Section 20.6.2] [17] and in sequence mining (e.g., [18]). The main difference is that our heuristic is on-line and stochastic rather than absolute. Since it does not apply the same threshold to each level, it is more flexible. It seeks the most frequent patterns at each level regardless of their absolute frequencies. Neither the basic pruning ideas nor the related incremental association rule work (based on a very different problem with different assumptions from our work) make clear how to do the job done by HSNGs. For HSNGs, no pattern at the next level is prunable. They must rely on a probabilistic interlevel bias.[5]

Our pattern selection bias (that amounts to a bias for combinations of existing frequent patterns) is similar to the compositional chunking of Sequitur [19] and similar systems, but with finer statistical sensitivity than Sequitur, which is very greedy. Also note that Sequitur remembers all of its input in order to chunk. Therefore, it is not a true on-line algorithm in the sense used here.

Maximum entropy (ME) [5, 20] (and related NLP) techniques use "features" more general than our patterns. Uniform completion of the distribution in SNGs is consistent with ME. In ME, features are added incrementally in a process seemingly similar to ours but actually very different. Each addition requires iterating over all training data. ME, transformation-based techniques [5], and related schemes all require batch training and it is not clear how to adapt them for on-line learning, though our models may suggest some necessary ingredients. ME uses its features as constraints, but it fundamentally assumes all constraints are equally statistically reliable (fine in a batch context but not when new features are introduced after different amounts of data). ME would overweight recently introduced frequency estimates. Some sort of regularization, as introduced here, is required.

Hierarchical HMMs can represent sparse collections of patterns in a more expressive representation, but only recently has structure learning been addressed [21] and only for batch training. Even with complete random access to the data, identification of repeated substructure is problematic. There are other learning models that representationally subsume our models, but in each case the learning problem they address is dramatically different due to use of a domain theory, supervision, batch training, or structured or segmented data (usually many of these, see [12, Sec.8.3, Sec.10.4, Ch.9]). In no case can they accomplish the learning described here.

5 Conclusion

Hierarchical sparse n-grams can be viewed as combining multiwidth n-grams from the NLP and text compression communities, frequent itemset pruning from the KDD community, and on-line learning in the ML and connectionist traditions. (This suggests much cross-fertilization. E.g., interlevel flow of information to help direct search is a key thing lacking from PST work.) The difficulties in combining sparseness with on-line learning are fundamental and require more than subtle changes to exhaustive batch techniques.

Sparse n-grams are useful alternatives to exhaustive n-grams when data overwhelms memory. For autonomous agents in complex environments, this is always the case in

[5] The width-n distribution provides no useful absolute bounds on $n{+}1$ probabilities. If there were a low-frequency n-tuple, the model could avoid its extensions, but the sparse models by their nature do not retain confidently estimated low-frequency patterns.

the long run. Hierarchical sparse n-grams learn frequent patterns using fewer parameters than the number of potential patterns and without remembering all data or repeatedly iterating over it. They use use novel techniques for falling back on narrower distributions. A stochastic bias for compositions of smaller known-frequent patterns facilitates on-line learning of increasingly complex sparse representations. The result is a model capable of finding frequent patterns in huge search spaces and cumulatively constructing ever-larger representations of the frequent patterns demonstrated by the environment.

We have necessarily concentrated on prediction as the primary way to utilize knowledge of frequent patterns (or chunks), but frequent chunks can be very valuable in many other ways within a larger computational system. Learned chunks can act as aids to increase working memory capacity, based on substitution recoding, which improves information processing capacity quite broadly.[22] Frequent patterns are important for developing communication or shared language. Frequent chunks can serve as important features for other types of learning and can enable the automatic formation of associations that would otherwise be impossible to induce. The point is that a single general learning mechanism can build representations that both enable useful predictions and also serve as foundations for improving several aspects of an agent's cognitive behavior.

Sparseness is fundamental in complex knowledge representation. Whether in semantic networks, frames, description logics, ontologies, etc., knowledge is always non-exhaustive for real-world domains. We must continue to investigate domain-independent techniques for choosing which patterns exhibited by the environment to include in sparse representations. For more extensive treatment of all material, including motivations, derivations, a formal description of the learning problem, experimental results, related work comparisons, and explanations of the usefulness of frequent patterns in the future research landscape more detailed than those of the preceding paragraph, see [12].

References

[1] Pedro Domingos and Geoff Hulten. Mining high-speed data streams. In *Knowledge Discovery and Data Mining*, pages 71–80, 2000.

[2] Richard Sutton and Steven Whitehead. Online learning with random representations. In *Proceedings of the 10th International Conference on Machine Learning*, 1993.

[3] Paul E. Utgoff. Many-layered learning. *Neural Computation*, 2002. In press.

[4] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291(5504), 2000.

[5] C. Manning and H. Schutze. *Foundations of Statistical Natural Language Processing*. 1999.

[6] Timothy C. Bell, John G. Cleary, and Ian H. Witten. *Text Compression*. Prentice-Hall, 1990.

[7] Joshua Goodman and Jianfeng Gao. Language model size reduction by pruning and clustering. In *International Conference on Spoken Language Processing*, 2000.

[8] Andrew Moore and Mary Soon Lee. Cached sufficient statistics for efficient machine learning with large datasets. *Journal of Artificial Intelligence Research*, 8:67–91, 1998.

[9] H. Schutze and Y. Singer. Part-of-speech tagging using a variable memory Markov model. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 1994.

[10] A. Stolcke. *Bayesian Learning of Probabilistic Language Models*. PhD thesis, Berkeley, 1994.

[11] Oded Maron and Andrew Moore. Hoeffding races: Accelerating model selection search for classification and function approximation. In *Advances in Neural Information Processing*, 1993.

[12] Karl Pfleger. *On-Line Learning of Predictive Compositonal Hierarchies*. PhD thesis, Stanford University, June 2002. See http://ksl.stanford.edu/~kpfleger/thesis/.

[13] D. Ron, Y. Singer, and N. Tishby. The power of amnesia. *Machine Learning*, 25, 1996.

[14] Fernando Pereira, Yoram Singer, and Naftali Z. Tishby. Beyond word n-grams. In *Proceedings of the Third Workshop on Very Large Corpora*, 1995.

[15] Nir Friedman and Yoram Singer. Efficient Bayesian parameter estimation in large discrete domains. In *Advances in Neural Information Processing Systems*, 1998.

[16] Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom. *Database Systems: The Complete Book*. Prentice Hall, 2002.

[17] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. Verkamo. Fast discovery of association rules. In *Advances in Knowledge Discovery and Data Mining*. AAAI Press, 1996.

[18] H. Mannila, H. Toivonen, and A. Verkamo. Discovering frequent episodes in sequences. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, 1995.

[19] Craig G. Nevill-Manning and Ian. H. Witten. Identifying hierarchical structure in sequences: a linear-time algorithm. *Journal of Artificial Intelligence Research*, 7:67–82, 1997.

[20] S. Della Pietra, V. Della Pietra, and J. Lafferty. Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(4), 1997.

[21] Kevin Murphy and Mark Paskin. Linear time inference in hierarchical HMMs. In *Advances in Neural Information Processing Systems 14*, 2002.

[22] George A. Miller. The magic number seven, plus or minus two: Some limits of our capacity for processing information. *Psychological Review*, 63, 1956.

A Unified Model of Early Word Learning: Integrating Statistical and Social Cues

Chen Yu
Department of Psychology and
Cognitive Science Program
Indiana University
Bloomington, IN, 47405
chenyu@indiana.edu

Dana H. Ballard
Department of Computer Science
University of Rochester
Rochester, NY, 14627
dana@cs.rochester.edu

Abstract

Previous work on early language acquisition has shown that word meanings can be acquired by an associative procedure that maps perceptual experience onto linguistic labels based on cross-situational observation. A new trend termed social-pragmatic theory [27] focuses on the effect of the child's social-cognitive capacities, such as joint attention and intention reading. This paper argues that statistical and social cues can be seamlessly integrated to facilitate early word learning. To support this idea, we first introduce a statistical learning mechanism that provides a formal account of cross-situational observation. The main part of this paper then presents a unified model that is able to make use of different kinds of social cues, such as joint attention and prosody in maternal speech, in the statistical learning framework. In a computational analysis of infant data, the quantitative results of our unified model outperforms the purely statistical learning method in computing word-meaning associations.

1 Introduction

What kinds of learning mechanisms underlie language acquisition? One of the central debates concerns whether the innate or environmental contribution plays a vital role in language development. Learning-oriented theories believe that language is learned and the child's environment plays a crucial role [14, 11, 24]. There is growing evidence that babies do possess powerful statistical learning mechanisms [23]. On the other hand, a nativist view sees linguistic universals as a product of the child's linguistic endowment and suggests that they do not need to be learned, which provide an elegant explanation for cross-linguistic similarities between different human languages [10].

In this paper, we first review two theories of language learning: statistical learning theory and social-pragmatic theory. Then Section 3 proposes our unified model that integrates statistical and social cues in a general system. Section 4 describes the implementation of the statistical learning model of word meaning, which provides a probabilistic framework for further study. Section 5 presents the methods to extract prosodic cues from raw speech and joint attention cues from infant-caregiver interactions. Section 6 provides a comparative study of different methods considering different sets of statistical and social cues.

2 Two Theories of Language Learning

This section reviews two well-known learning-oriented theories of language acquisition. The theory of statistical learning suggests that language acquisition is a statistically driven process in which young language learners utilize the lexical content and syntactic structure of speech as well as non-linguistic contextual information as input to compute distributional statistics. The social-pragmatic theory focuses on mind reading (social cognition) as fundamental to the word learning process. Both theories have been supported by various empirical and computational studies.

2.1 The Theory of Statistical Learning

Human language learners possess powerful statistical learning capacities. That is, the cognitive system in both children and adults is sensitive to features of the input (e.g., occurrence statistics). Saffran, Aslin and Newport [23] showed that 8-month-old infants are able to find word boundaries in an artificial language only based on statistical regularities. Later studies [22] demonstrated that infants are also sensitive to transitional probabilities over tone sequences, suggesting that this statistical learning mechanism is more general than the one dedicated solely to processing linguistic data. Furthermore, statistical language learning includes not only statistical computations to identity words in speech but also algebraic-like computations to learn grammatical structures [18].

In the study of word learning, *associationism* claims that word acquisition is based on statistical learning of co-occurring data from the linguistic modality and non-linguistic context (see a review by [19]). Richards and Goldfarb [21] proposed that children come to know the meaning of a word through repeatedly associating the verbal label with their experience at the time that the label is used. Smith [25] argued that word learning is initially a process in which children's attention is captured by objects or actions that are the most salient in their environment, and then used to associate those objects or actions with acoustic patterns voiced by an adult. Plunkett [19] developed a connectionist model of vocabulary development to associate preprocessed images and linguistic labels. The linguistic behaviors of the network can mimic the well-known vocabulary spurt based on small continuous changes in the connection strengths with and across different processing modalities in the network. In general, the statistical and associative mechanism of word learning divides the word learning task into three subtasks: word discovery, meaning discovery and word-meaning association. The vital part is to use multiple word-meaning pairs collected from different situations to compute co-occurrences and then establish word-to-world mappings [14].

2.2 The Social-Pragmatic Theory

The social-pragmatic theory of language acquisition argued that the major sources of constraints in language acquisition are social cognitive skills, such as children's ability to infer the intentions of adults as adults act and speak to them [1, 27, 7]. These kinds of social cognition are called "mind reading" by Baron-Cohen [4]. Kuhl et al. [16] studied whether phonetic learning of 9-10 month children is simply triggered by hearing language. If so, children should be able to learn by being exposed to language materials via digital video without human interaction. However, the results showed that infants cannot learn phonetics through this way, suggesting that the presence of a live person provides not only social cues but also referential information. Butterworth [9] showed that even by 6 months of age, infants demonstrate sensitivities to social cues, such as monitoring and following another person's gaze. In Baldwin's work [1], the 18-month old infant heard the novel word while his/her attention was focused on one toy and the experimenter looked at another toy. When children heard the same word in a testing phase, they chose the object at which the experimenter had been looking. This suggested that the infants were able to follow the speaker's attention and infer the mental state of the speaker to determine the referent of the novel word. Furthermore, Baldwin et al.[2] proposed that infants give a special weight to the cues of indexing the speaker's referential intent when determining the reference of a novel label. Their experiments showed that infants established a stable link between the novel label and the target toy only when that label was uttered by a speaker who concurrently showed his attention toward the target, and such a stable mapping was not established when the label was uttered by a speaker who showed no signs of attention to the target toy, even if the object appeared at the same moment when that label was uttered and the speaker was touching the object. In addition, their results suggested that children not only attend to referential intentions of a speaker but also actively look for the intention of the speaker when determining whether to associate a novel word with an object.

3 A Unified Model

Bloom [6] argued that children's conceptual biases, intentional understanding and syntactic knowledge are not only necessary for word learning but that they are also sufficient. This claim contrasts with the theory that word learning is based on an associative learning mechanism that is sensitive to statistical properties of the input [19]. The statistical and associative theory suggested that the child's sensitivity to spatio-temporal contiguity is sufficient for word learning, as postulated by associationist models of language acquisition with support by computational implementation [11, 20]. The debate on these two theories has been going on for several years.

Associative learning mechanisms make sense because words are typically spoken at the moment when the child looks at the things that those words refer to. In western cultures, parents provide linguistic labels of objects for their kids when the objects are in the kids' visual fields. Thus, no one doubts that humans can learn co-occurrence relationships and that the easiest way to teach language is to provide linguistic labels at the same time that children focus on them. However, parents do not carefully name objects for their kids in many cultures. Even in western cultures, words are not always used at the moment that their referents are perceived. For instance, Gleitman [13] showed that most of the time, the child does not observe something being opened when the verb "open" is used. Nevertheless, children have no difficulty in learning those words. Associative learning, without some further constraints or additional information, cannot explain this observation.

The theory of mind reading is able to explain many phenomena from the perspective of the inference of a speaker's referential intentions, especially for the cases that words and the corresponding meanings are not co-occurring, or words are temporally correlated with irrelevant meanings. However, the environment in which infants develop does contain the information that is useful for statistical learning mechanisms. Meanwhile, empirical studies (e.g. [23] and [18]) showed that infants can utilize the statistical properties of the input in language acquisition. Taken together, it is very plausible that infants perform statistical computations in language learning.

Fortunately, the theory of statistical learning and social-pragmatic theory are not mutually exclusive. Recently, Hirsh-Pasek, Golinkoff and Hollich [15] proposed a coalition model in which multiple sources, such as perceptual salience, prosodic cue, social eye gaze, social context, syntactic cues and temporal contiguity, are used by children to learn new words. They argued that during the development, the weighting of the cues changes over time while younger children can just detect and make use of only a subset of the cues in the coalition and the older can use a wider subset of cues.

The purpose of this study is to show quantitatively the effects of statistical cross-situational observation and social cues through computational modeling. In early word learning, children need to start by pairing spoken words with the co-occurring possible referents, collecting multiple such pairs, and then figuring out the common elements. Although no one doubts this process, few research has addressed the details of cross-situational observation. This work first introduces a formal model of statistical word learning which provides a probabilistic framework for encoding multiple sources of information. Given multiple scenes paired with spoken words collected from natural interactions between caregivers and their kids, the model is able to compute the association probabilities of all the possible word-meaning pairs. Moreover, we argue that social cues can be naturally integrated in the model as additional constraints in computation. The claim here is that language learners can use social cues, such as gaze direction, head direction, body movement, gesture, intonation of speech and facial expression, to infer speakers' referential intentions. We show how these social cues can be seamlessly integrated in the framework of statistical learning and facilitate word learning. Specifically, we focus on two kinds of social cues: body movement cues indicating the speaker's attention and prosodic cues in speech. This study proposes that those social cues can play a spotlight role (shown in Figure 1) in the statistical learning by causing language learners to focus on certain aspects of a scene. Since every scene is ambiguous and contains multiple possible referents, this spotlight function is crucial in solving the word-to-world mapping problem. The following subsections discuss how those cues might help in detail.

3.1 The Role of Body Movement in Language Acquisition

Ballard et al. [3] argued that at time scales of approximately one-third of a second, orienting movements of the body play a crucial role in cognition and form a useful computational level, termed the embodiment level. At this level, the constraints of the body determine the nature of cognitive operations. This computation provides a language that links external sensory data with internal cognitive programs and motor actions through a system of implicit reference termed deictic, whereby pointing movements of the body are used to bind objects in the world to cognitive programs. Examples of sensorimotor primitives at the embodiment level include an eye movement, a hand movement, or a spoken word.

We apply the theory of embodied cognition in the context of language learning. To do so, one needs to consider the role of embodiment from both the perspective of a speaker (language teacher) and that of a language learner. In the study of speech production, Meyer et al. [17] found that speakers' eye movements were tightly linked to their speech output. When speakers were asked to describe a set of objects from a picture, they usually looked at each new object before mentioning it, and their gazes remained on the object until they were about to say the last word about it. From the perspective of a language learner, Baldwin [1] showed that infants actively gathered social information to guide their inferences about word meanings and they systematically checked the speaker's gaze to clarify his/her reference. In the follow-up studies, Baldwin and Baird [2] proposed that humans gradually develop the skill of mind reading so that ultimately they care little about the surface behaviors of others' dynamic action but focus on discerning underlying intentions based on a generative knowledge system. Summarizing all these ideas on embodied cogni-

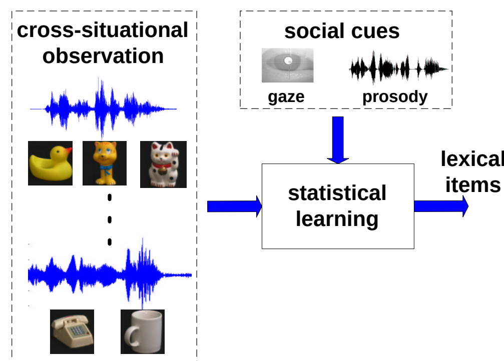


Figure 1. Cross-situational observation and social cues can be seamlessly integrated in a statistical learning model.

tion, speech production and social development, the speakers' body movements, such as eye movements, head movements and hand movements, can reveal their referential intentions in verbal utterances, which, in turn almost certainly could possibly play a significant role in early language development [29]. A plausible starting point of learning the meanings of words is the deployment of speakers' intentional body movements to infer their referential intentions. To support this idea, we provide a formal account of how the intentions derived from body movements, which we term *embodied intention*, facilitate the early stage of vocabulary acquisition. We argue that infants learn words through their sensitivity to others' intentional body movements in a very specific way: They use temporal synchrony between speech and referential body movements to find the referents of language.

3.2 The Role of Prosodic Cue

When talking to human infants, parents use vocal patterns that are different from normal conversation. They speak slowly and with higher pitch and exaggerated intonation contours. Fernald [12] proposed a model consisting of four developmental functions of intonation in speech to infants. The first function is that infants are attentive to intrinsic perceptual and affective salience in the melodic intonation of mothers' speech. At the second level, the exaggerated intonation patterns of mothers' speech would influence both attentional preference and affective responsiveness of infants. The third function is about the inference of the communicative intents of speakers from maternal intonation of speech. Infants are able to interpret the emotional states of others and make predictions about the future actions of others using information available in vocal and facial expressions, which provide reliable cues to the affective state and intentions of speakers. The fourth level focuses on the role of prosodic cues in early language development. Fernald argued that the prosody of speech helps to identify linguistic units within the continuous speech signal. Thus it serves as an attention-focusing device so that mothers use a distinctive prosodic strategy to highlight focused words. Most often, exaggerated pitch peaks are correlated with lexical stress. In light of this, we investigate the role of prosodic cue in early word learning in this paper. Specifically, we focus on the spotlight function of prosody and provide a formal account of how prosodic cues might be used in word learning.

4 A Statistical Model of Cross-Situational Observation

Our study uses the video clips of mother-infant interactions from the CHILDES standard database. These clips contain simultaneous audio and video data wherein a mother introduces her child to a succession of toys stored in a nearby box.

In this kind of natural interaction, the vocabulary is rich and varied and the central items (toy names) are far from the most frequent words. This complex but perfectly natural situation can be easily quantified by plotting a histogram of word frequency which shows that none of the key words – toy names make it into the top 15 items of the list. An elementary idea for improving the ranking of key words assumes that the infants are able to weight the toy utterances more by taking advantage of the approximately coincident body cues. For instance, the utterances that were generated when the infant's gaze was fixated on the toys by following the mother's gaze have more weights than the ones the young child just looked around while not paying attention to what the mother said. We examined the transcript and weighted the words according to how much they were emphasized by such cues, but this strategy does little to help spot the toy names.

Next, we manually labeled visual objects in the context when a spoken utterance was produced, and found what *is* helpful is to partition the toy sequences (contextual information when the speech was produced) into intervals where within each interval a single toy or small number of co-occurring toys is the central subject or meaning, and then categorize spoken utterances using the contextual bins labeled by different toys. The hypothesis is that mothers use temporal synchrony to highlight novel word-referent relations for young infants. That is, presenting information across multiple modalities simultaneously serves to highlight the relations between the two patterns of stimulation. Thus, temporal synchrony can facilitate infants' detection of word-referent relations. Formally, associating meanings (toys, etc.) with words (toy names, etc.) can be viewed as the problem of identifying word correspondences between English and a "meaning language", given the data of these two languages in parallel. With this perspective, a technique from machine translation can address the correspondence problem [8]. The probability of each word is expressed as a mixture model that consists of the conditional probabilities of each word given its possible meanings. In this way, an Expectation-Maximization (EM) algorithm can find the reliable associations of object names and their meanings which will maximize the likelihood function of observing the data set.

The general setting is as follows: suppose we have a word set $X = \{w_1, w_2, ..., w_N\}$ and a meaning set $Y = \{m_1, m_2, ..., m_M\}$, where N is the number of words and M is the number of meanings (toys, etc.). Let S be the number of spoken utterances. All word data are in a set $\chi = \{(S_w^{(s)}, S_m^{(s)}), 1 \leq s \leq S\}$, where each spoken utterance $S_w^{(s)}$ consists of r words $w_{u(1)}, w_{u(2)}, ..., w_{u(r)}$, and $u(i)$ can be selected from 1 to N. Similarly, the corresponding contextual information $S_m^{(s)}$ include l possible meanings $m_{v(1)}, m_{v(2)}, ..., m_{v(l)}$ and the value of $v(j)$ is from 1 to M. Assume that every word w_n can be associated with a meaning m_m. Given a data set χ, We use the machine translation method proposed by Brown et al. [8] to maximize the likelihood of generating the meaning strings given English descriptions:

$$
\begin{aligned}
& P(S_m^{(1)}, S_m^{(2)}, ..., S_m^{(S)} | S_w^{(1)}, S_w^{(2)}, ..., S_w^{(S)}) \\
= & \prod_{s=1}^{S} \sum_a p(S_m^{(s)}, a | S_w^{(s)}) \\
= & \prod_{s=1}^{S} \frac{\epsilon}{(r+1)^l} \prod_{j=1}^{l} \sum_{i=0}^{r} p(m_{v(j)} | w_{u(i)}) \qquad (1)
\end{aligned}
$$

where the alignment a indicates which word is aligned with which meaning. $p(m_{v(j)}|w_{u(i)})$ is the association probability for a word-meaning pair and ϵ is a small constant. The expected number of times that any particular word w_n in a language string $S_w^{(s)}$ generates any specific meaning m_m in the co-occurring meaning string $S_m^{(s)}$ is given by

$$c(m_m|w_n, S_m^{(s)}, S_w^{(s)}) = \frac{p(m_m|w_n)}{p(m_m|w_{u(1)}) + ... + p(m_m|w_{u(r)})} \times \sum_{j=1}^{l} \delta(m_m, v(j)) \sum_{i=1}^{r} \delta(w_n, u(i)) \tag{2}$$

where δ is equal to one when both of its arguments are the same and equal to zero otherwise. Accordingly, the association probabilities are given by

$$p(m_m|w_n) = \frac{\sum_{s=1}^{S} c(m_m|w_n, S_m^{(s)}, S_w^{(s)})}{\sum_{m=1}^{M} \sum_{s=1}^{S} c(m_m|w_n, S_m^{(s)}, S_w^{(s)})} \tag{3}$$

The method sets an initial $p(m_m|w_n)$ to be flat distribution, and then successively compute the counts of all word-meaning pairs $c(m_m|w_n, S_m^{(s)}, S_w^{(s)})$ using Equation (2) and the association probabilities using Equation (3). The technical details of our method can be found in [28]. The results of this statistical learning model are reported in Section 6.

5 The Integration of Social Cues in Statistical Learning

The communication of infants and their caregivers is multisensory. It involves visual information, tactile information as well as auditory information. Besides linguistic information, we believe that social cues encoded in multimodal interaction highlight target word-referent relations for young language learners. In a bidirectional relationship between maternal multimodal communication styles and infants' perception of word-referent relations, mothers synchronize their verbal references and nonverbal body movements (eye gaze, gesture, etc.) for infants. At the same time, infants are able to rely on observing mother's eye gaze and other pointing motions to detect their's referential intentions in speech. Thus, both mothers and infants actively involve into multimodal communication to solve the mapping problem in lexical acquisition. This study provides a quantitative account of how those multimodal social cues can facilitate word learning. Specifically, we focus on two cues: joint attention cues as deictic reference and prosodic cues in maternal speech.

5.1 Visual Spotlight

Children as young as 12-18 months spontaneously check where a speaker is looking when he/she utters a word, and then link the word with the object the speaker is looking at. This observation indicates that joint visual attention (deictic gaze) is a critical factor that should be considered in word learning. When presenting information, that visual spotlight gives maximal processing to that part of the visual field. During natural infant-caregiver interactions, joint visual attention involves detecting a spotlight of a mother's attention to the object in the scene, and then moving the body, head and eyes to acquire the target object with high-resolution focal vision, which is one of the crucial steps to deal with the mapping problem.

transcriptions	attended objects	other objects
– the kitty-cat go meow meow	kitty-cat	baby, big-bird, rattle, book
– ah and a baby	baby	kitty-cat, big-bird, rattle, book
– there's a baby just like my David	baby	kitty-cat, big-bird, rattle, book
– a baby	baby	kitty-cat, big-bird, rattle, book
that's a nice book	book	kitty-cat, big-bird

Table 1. Examples of transcriptions and contextual labels.

In our experiment, we coded visual contexts to study the role of joint attention. As shown in Table 1, we provided two labels to describe visual contextual information for each spoken utterance. One label indicated the objects of joint attention which were attending by both the mother and the kid. The second label represented all the other objects in the visual field of the kid. Figure 3 illustrates two examples of speech-scene pairs in which the shaded meanings are attentional objects and non-shaded meanings are other objects in the scene. In Section 5.3, we describe our method that makes use of this attentional information in word learning.

5.2 Prosodic Spotlight

Snedeker and Trueswell [26] showed that speakers produce and listeners use prosodic cues to disambiguate alternative meanings of a syntactic phrase in a referential communication task. Moreover, previous research suggests that mothers adapt their verbal communication to infants in order to facilitate their language learning. In this work, we analyze maternal speech by extracting low-level acoustic features and using those features to spot the words emphasized by adults. We proposed that perceptually salient prosodic patterns may serve as "spotlights" on linguistic information conveyed by speech. Thus, we focus on the role of prosodic features in word learning, which might help language learners to identify key words from the speech stream.

Fernald [12] suggested that the exaggerated acoustic patterns have evolved to elicit and sustain infants' attention to speech as well as highlight the important parts of the speech stream. In the context of word learning, we observe that prosodically salient words in maternal speech can be categorized into two classes. One group of words serve as communication of intention and emotion. One important role of those words is to attract the kid's attention so that the child would follow what the mother talks about and what she looks at. In this way, both the mother and the language learner share the visual attention, which is a cornerstone in social and language development. The right column in Figure 2 illustrates an example in the video clips in which the mother used high pitch to say *you* to attract the kid's attention. Some other common words and phrases frequently used by the mother are *yeah*, *oh*, *look* and *that's*. The other group of words contain the most important linguistic information that the mother wants to convey. In the context of word learning, most of those words refer to the concepts that are related to visual objects in the physical environment, such as object names, their colors, sizes and functions. An example of words in the second group is the object name *baby* shown in the left column of Figure 2.

In implementation, CMU sphinx speech recognition system was used to align maternal speech and transcriptions. As a result, the timestamps of the beginning and end of each spoken word was extracted. Next, we made three kinds of low-level acoustic measurements on each utterance and word. The prosodic features were extracted based on pitch (f0) information. For each feature, we extracted the values over both an utterance and each word within this utterance.

- **75 percentile pitch** p_{75}: the 75 percentile pitch value of all voiced part of the speech unit.
- **Delta pitch range** p_r: the change in pitch between frames (20ms) was calculated as delta pitch. This measure represents the difference between the highest and the lowest delta pitch values within the unit (utterance or word).
- **Mean delta pitch** p_m: the mean delta pitch of the voiced part of the spoken unit.

We want to obtain prosodically highlighted words in each spoken utterance. To do so, we compare the extracted features from each word with those from each utterance, which indicates whether a word sounds like "highlighted" in the acoustic context. Specifically, for the word w_i in the spoken utterance u_j, we form a feature vector: $[p_{75}^{w_i} - p_{75}^{u_j} \quad p_r^{w_i} - p_r^{u_j} \quad p_m^{w_i} - p_m^{u_j}]^T$, where $p_m^{u_j}$ is the mean delta pitch of the utterance and $p_m^{w_i}$ is that of the word and so on. In this way, the prosodic envelope of a word is represented by 3-dimensional feature vector. We use the support vector clustering (SVC) method [5] to group data point into two categories. One consists of prosodically salient words and the other one includes non-emphasized words. In SVC algorithm, data points are mapped from the data space to a high dimensional feature space using a Gaussian kernel. In this feature space, the algorithm looks for the smallest sphere that encloses the data, and then maps the data points back to the data space and forms a set of contours to enclose them. These contours can be interpreted as cluster boundaries.

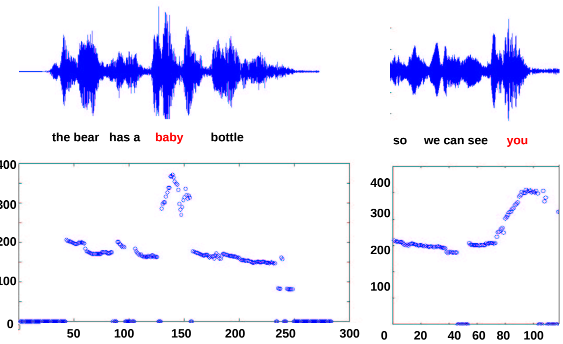


Figure 2. Speech and intonation. The prosodic cues highlight several words. The first column represents speech signals and the second column shows the profiles of fundamental frequency (f0). The word *baby* is highlighted in the left utterance and the word *you* is prosodically distinctive from others in the right utterance.

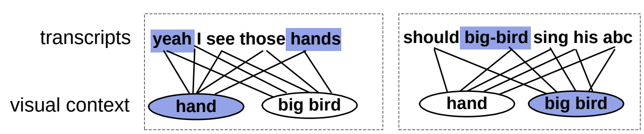


Figure 3. Cross-situational word-meaning association with social cues. The prosodic cues highlight some words in speech and the cues of joint attention highlight attentional objects in visual contexts.

5.3 Modeling the Role of Social Cues in Statistical Learning

We encode social cues in the framework of the statistical learning model as shown in Figure 3. Each word $u(i)$ is assigned with a weight $w_p(i)$ based on its prosodic category. Similarly, each visual object v_j is set with a weight $w_v(j)$ based on whether it is attended by the speaker and the learner. In this way, the same method described in previous section is applied and the only difference is that the estimate of $c(m_m|w_n, S_m^{(s)}, S_w^{(s)})$ now is given by:

$$c(m_m|w_n, S_m^{(s)}, S_w^{(s)}) = \frac{p(m_m|w_n)}{p(m_m|w_{u(1)}) + ... + p(m_m|w_{u(r)})}$$

$$\times \sum_{j=1}^{l} \delta(m_m, v(j) * w_v(j)) \sum_{i=1}^{r} \delta(w_n, u(i) * w_p(i)) \tag{4}$$

In practice, we set the values of $w_v(j)$ and $w_p(i)$ to be 3 for highlighted objects and words. The weights of all the other words and objects are set to be 1.

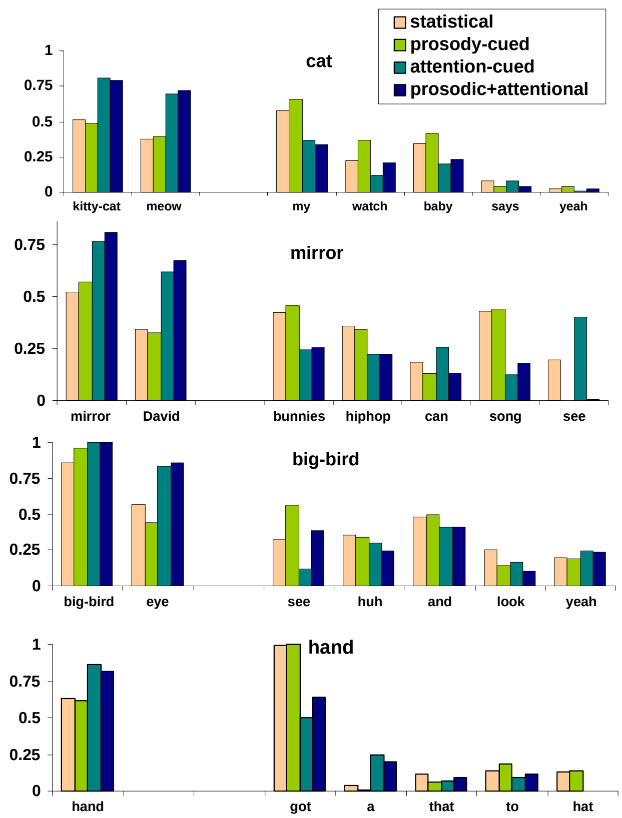


Figure 4. The comparative results of the methods considering different sets of cues. Each plot shows the association probabilities of several words to one specific meaning labeled on the top. The first one or two items are correct words that are relevant to the meanings and the following words are irrelevant.

6 Experimental Results

Our model was evaluated by using two video clips from CHILDES database. We labeled visual contexts in terms of 12 objects that occurred in the video clips. For each object, we selected the correctly associated words based on general knowledge. For instance, both the word *kitty-cat* and *meow* are positive instances because both of them are relevant to the object "cat". Overall, there were 26 positive words for all of the 12 objects. The computational model estimated the association probabilities of all the possible word-meaning associations and then selected lexical items based on a threshold. Two measures were used to evaluate the performance: (1) word-meaning association accuracy (precision) measures the percentage of the words spotted by the model which actually are correct. (2) lexical spotting accuracy (recall) measures the percentage of correct words that the model learned among all the 26 words.

Four methods were applied on the same data and the results are as follows (precision and recall): (1) purely statistical learning (75% and 58%). (2) statisical learning with prosodic cues (78% and 58%). (3) statistical learning with the cues from visual attention (80% and 73%). (4) statistical learning with both attentional and prosodic cues (83% and 77%). Figure 4 shows the comparative results of these four approaches for specific instances. Ideally, we want the association probabilities of the first or second words to be high and others to be low. For instance, the first plot represents the meaning of the object "cat". Both the spoken word *kitty-cat* and the spoken word *meow* are closely relevant to this meaning. Therefore, the association probabilities are high for these two words and are low for all the others words, such as *my*, *watch* and *baby*, which are not correlated with this context. Note that in the meaning of the object "bird", we count the word *eye* as a positive one because the mother uttered it several times during the interaction when she presented the object "bird" to her kid. Similarly, when she introduced the object "mirror", she also mentioned the name of the kid *David* whose face appeared in the mirror.

The results of the statistical learning approach (the first bars) are reasonably good. For instance, it obtains *big-bird* and *eye* for the meaning bird, *kitty-cat* for the meaning "cat", *mirror* for the meaning "mirror" and *hand* for the meaning "hand". But it also makes wrong estimates, such as *my* for the meaning "cat" and *got* for the meaning "hand". We expect that attentional and prosodic constraints will make the association probabilities of correct words higher and decrease the association probabilities of irrelevant words. The method encoding prosodic cues moves toward this goal although occasionally it changes the probabilities on the reverse way, such as increasing the probability of *my* in the meaning "cat". What is really helpful is to encode the cues of joint attention. The attention-cued method significantly improves the accuracy of estimate for almost every word-meaning pairs. Of course, the method including both joint-attention and prosodic cues achieves the best performance. Compared with purely statistical learning, this method highlights the correct associations (e.g., *kitty-cat* with the meaning "cat"), and decreases the irrelevant associations, such as *got* with the meaning "hand". In this method, we can simply select a threshold and pick the word-meaning pairs which are overlapped with the majority of words in the target set. We need to point out that the results here are obtained from very limited data. Without any prior knowledge of the language (the worst case in word learning), the model is able to learn a significant amount of correct word-meaning associations.

7 Conclusion

We believe that in a natural infant-caregiver interaction, the mother provides non-linguistic signals to the infant through her body movements, the direction of her gaze, and the timing of her affective cues via prosody. Previous experiments have shown that some of these non-linguistic signals can play a critical role in infant word learning, but a detailed estimate of their relative weights has not been provided. Based on statistical learning and social-pragmatic theories, this work proposed a unified model of early word learning, which integrates statistical and social cues to enable the word-learning process to function effectively and efficiently. In our model, we explored the computational role of non-linguistic information, such as joint attention and prosody in speech, and provided the quantitative results to compare the effects of different statistical and social cues. We need to point out that the current unified model does not encode any syntactic properties of the language, which definitely play a significant role in word learning, especially in the later stage. Therefore, one natural extension of the current work is to add the syntactic constraints in the current probabilistic framework to study how this knowledge can help the lexical acquisition process and how multiple sources can be integrated in a general system.

References

[1] D. Baldwin. Early referential understanding: Infant's ability to recognize referential acts for what they are. *Developmental psychology*, 29:832–843, 1993.

[2] D. A. Baldwin and J. A. Baird. Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*, 5(4), 2001.

[3] D. H. Ballard, M. M. Hayhoe, P. K. Pook, and R. P. N. Rao. Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, 20:1311–1328, 1997.

[4] S. Baron-Cohen. *Mindblindness: an essay on autism and theory of mind*. MIT Press, Cambridge, 1995.

[5] A. Ben-Hur, D. Horn, H. T. Siegelmann, and V. Vapnik. Support vector clustering. *Journal of machine Learning Research*, 2:125–137, 2001.

[6] P. Bloom. Intentionality and word learning. *Trends in Cognitive Sciences*, 1(1):9–12, 1997.

[7] P. Bloom. *How children learn the meanings of words*. The MIT Press, Cambridge, MA, 2000.

[8] P. F. Brown, S. Pietra, V. Pietra, and R. L. Mercer. The mathematics of statistical machine translation:parameter estimation. *Computational Linguistics*, 19(2), 1993.

[9] G. Butterworth. The ontogeny and phylogeny of joint visual attention. In A. Whiten, editor, *Natural theories of mind: Evolution, development, and simulation of everyday mindreading*. Blackwell, Oxford, England, 1991.

[10] N. Chomsky. *Aspects of the theory of syntax*. MIT Press, 1965.

[11] J. Elman, E. Bates, M. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. *Rethinking innateness: A connectionist perspective on development*. MIT Press, 1996.

[12] A. Fernard. Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In *The Adaptive Mind*. Oxford University Press, 1992.

[13] L. Gleitman. The structural sources of verb meanings. *Language Acquisition*, 1:1–55, 1990.

[14] L. Gleitman, K. Cassidy, R. Nappa, A. Papafragou, and J. Trueswell. Hard words. *Language Learning and Development*, 1, in press.

[15] K. Hirsh-Pasek, R. M. Golinkoff, and G. Hollich. An emergentist coalition model for word learning: mapping words to objects is a product of the interaction of multiple cues. In *Becoming a word learner: a debate on lexical acquisition*. Oxford Press, New York, 2000.

[16] P. K. Kuhl, F.-M. Tsao, and H.-M. Liu. Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *PNAS*, 100(15):9096–9101, July 22 2003.

[17] A. S. Meyer, A. M. Sleiderink, and W. J. Levelt. Viewing and naming objects: eye movements during noun phrase production. *Cognition*, 66:B25–B33, 1998.

[18] M. Pena, L. L. Bonatti, marina Nespor, and J. Mehler. Signal-driven computations in speech processing. *Science*, 2002.

[19] K. Plunkett. Theories of early language acquisition. *Trends in cognitive sciences*, 1:146–153, 1997.

[20] T. Regier. Emergent constraints on word-learning: A computational review. *Trends in Cognitive Sciences*, 7:263–268, 2003.

[21] D. Richards and J. Goldfarb. The episodic memory model of conceptual development: An integrative viewpoint. *Cognitive Development*, 1:183–219, 1986.

[22] J. R. Saffran, E. Johnson, R. Aslin, and E. Newport. Statistical learning of tone sequences by human infants and adults. *Cognition*, 1999.

[23] J. R. Saffran, E. L. Newport, and R. N. Aslin. Word segmentation: The role of distributional cues. *Journal of memory and language*, 35:606–621, 1996.

[24] M. S. Seidenberg. Language acquisition and use: Learning and applying probabilistic constraints. *Science*, 275:1599–1603, March 1997.

[25] L. Smith. How to learn words: An associative crane. In R. Golinkoff and K. Hirsh-Pasek, editors, *Breaking the word learning barrier*, pages 51–80. Oxford: Oxford University Press, 2000.

[26] J. Snedeker and J. Trueswell. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, 48:103–130, 2003.

[27] M. Tomasello. Perceiving intentions and learning words in the second year of life. In M. Bowerman and S. Levinson, editors, *Language Acquisition and Conceptual Development*. Cambridge University Press, 2000.

[28] C. Yu and D. H. Ballard. A multimodal learning interface for grounding spoken language in sensory perceptions. In *Fifth International Conference on Multimodal Interface*. ACM Press, 2003.

[29] C. Yu, D. H. Ballard, and R. N. Aslin. The role of embodied intention in early lexical acquisition. In *Proceedings the Twenty Fifth Cognitive Science Society Annual Meetings*, Boston, MA, July 2003.

On Language and Age of Acquisition

Arturo E. Hernandez
126 Heyne Bldg, Department of Psychology, University of Houston. Houston TX 77204
aehernandez@uh.edu

Abstract

Age of acquistion (AoA) effects have been found to have strong effects in the syntactic domain. The current paper reviews this literature and newer work which suggests that syntactic effects may be present in certain syntactic functions. In addition, work which suggests that AoA also shows effects during semantic processing is presented. It is concluded that AoA effects are pervasive across a wide range of tasks and domains. Theoretical accounts of these effects are discussed.

1. Introduction

A distinguished colleague tells the story of her father, who immigrated to the U.S. from Holland at age 32 and functioned well in English for three decades at home and in his business as an insurance broker. Despite these years of fluency in two languages, he lost the ability to communicate in English (but not Dutch) during hospitalization for a stroke at the age of 63. Reports like the one above are quite dramatic and have led some to propose that each language is represented in different parts of the brain [1]. However, others see this as evidence that each language is differentially sensitive to damage [2, 3]. The variables that modulate neural activity, age of acquisition (AoA) and proficiency (PR), have been discussed for over 100 years [2, 3]. These two variables are also reflected in recent work using functional Magnetic Resonance Imaging (fMRI), a technique which allows researchers to look at the oxygenation level of blood and thereby measure which neural areas are firing more extensively during a particular task. Recent studies using fMRI have found mixed effects with some suggesting that AoA is the primary determinant of neural activity whereas others suggest that proficiency is the primary determinant. AoA has been found to modulate neural activity during sentence comprehension [4] but only when proficiency is NOT taken into account. When early and late bilinguals were equated on proficiency, the differences between these groups disappeared [5]. The importance of proficiency has been supported by studies which find considerable individual differences in the level of proficiency in second language learners, in both early and late learners [6]. Proficiency has also been found to play a role in semantic tasks [7]. Evidence for the importance of proficiency can be found in recent work with populations that are immersed and educated in a second language relatively early in life. Work with Koreans adopted by French families reveal no neural or behavioral trace of the first language even when it was learned as late as age 8 [8]. Second, behavioral work by Hernandez and colleagues suggests that proficiency and not AoA determine naming latencies when L2 acquisition occurs early in life [9-12]. In short, to date there is mixed evidence that AoA is the primary determinant of behavioral and neural asymmetries while performing language tasks.

The fact that AoA seems to play a reduced role in some bilingual research is counterintuitive. AoA is known to be an important factor in a number of domains, especially in phonological processing and production of a second language [13-15]. More importantly, research which has investigated the effects of AoA on language processing has found that tasks involving syntax show larger AoA effects than semantic tasks [16, 17]. In a seminal study, Weber-Fox and Neville [16] asked a group of Chinese-English participants to look at sentences which contained three different types of syntactic violations (phrase structure, specificity constraint, and subjacency constraint) as well as semantic violations. This experiment used event-related potentials (ERP's) a method which provides the means for measuring the brain's electrical activity to a number of linguistic and non-linguistic factors. In the language domain, ERP's have been found to be sensitive to semantic violations [18, 19] and syntactic violations [20]. Results revealed differences in the timing and distribution of the ERP's to syntactic violations in participants who learned English as early as 2. However, differences in the ERP's to semantic violatioins only appeared in participants who learned English after 11. These results are consistent with the view that AoA plays a role in the neural activity associated with grammatical violations.

More recently, Wartenburger et al. [7] asked Italian-German bilinguals to monitor for syntactic violations (number, gender or case) or semantic violations.while being scanned with fMRI. Three groups were tested, early bilinguals with high proficiency in L2 (EAHP in a second language), and late bilinguals with either high (LAHP) or low proficiency (LALP) in L2. Increased brain activity in L2 relative to L1 was seen in all three groups for both semantic and syntactic violations. Furthermore, direct comparisons between groups in L2 yielded an interesting pattern of results. For grammaticality judgements, LAHP subjects showed increased activity in BA 44/6 and BA 44 relative to the EAHP group. BA 44 has been found to be associated with morphosyntactic processing [21] whereas superior BA 44 (near BA 6) is associated with phonological retrieval [22]. Taken together these results suggest that processing of grammatical violations in late learners results in increased motor planning and articulatory effort even when these subjects are matched in proficiency with early learners. Whereas there were also differences between the LAHP and LALP subjects, these were restricted to areas in the temporal-parietal juncture, the inferior parietal lobule and the lingual gyrus. However, there was no increased activity for the LALP subjects relative to the LAHP subjects.

A different pattern emerged for between-group comparisons during semantic processing. In these paradigms, there was increased activity in BA 46 and the fusiform gyrus for the LAHP group. For the LALP group, there was increased acitivity in BA 46/9 and BA 44/6. These results are consistent with the view that the late low proficiency group is engaged in more effortful phonological retrieval (BA 44/6). Furthermore, this increase in phonological retrieval leads to increased activity in BA 46/9, an area that is known to be involved in executive function for both verbal and nonverbal tasks [23]. However, there were no differences between both high proficiency groups. Taken together these results are consistent with the view that syntactic processing is sensitive to differences in AoA whereas semantic processing is sensitive to differences in proficiency. Finally, it suggests that both semantic and syntactic group differences are associated with increased phonological retrieval (BA 44/6) whereas activity associated with brain areas involved in morphosyntactic processing, i.e. using the ends of words to determine their grammatical functions, (BA 44) distinguished between groups that show differences on syntactic tasks. In short, there is some aspect of syntactic processing that leads to activity of areas that are more tightly associated with syntactic processing.

Recent results from the literature open up a number of questions with regard to the finding that syntax is more sensitive to AoA than semantics. First, it is not clear what factors may play a role in the AoA effect. One possibility is that syntactic functions share less across languages than semantic functions (at least the ones tested to date). In addition, it is possible that there is some processing component of syntax which is more sensitive to AoA (a classic third variable problem). Second, it remains to be seen if AoA effects are present for semantic domains.

1.1 Which syntactic functions show AoA effects?

In a first study, a set of Spanish native speakers who had spent less than 2 years in the United States at the time of testing were asked to indicate via button press the grammatical gender of a set of words in Spanish [24]. The opacity of the mapping was varied such that half the items were transparent (a for feminine and o for masculine) and half the items were opaque (ending in d,e.n,l,r,s,t,z). The results revealed increased activity for the opaque items in the anterior insula, BA 44/45, and BA 44/6. BA 44/45 has been found to be active for studies which have looked at syntactic processing [21 , 25, 26 , 27] as well as in studies which have compared gender monitoring to semantic monitoring [28]. The anterior insula is known to be involved in articulation [29, 30] and BA 44/6 is known to be involved in phonological processing [22]. Furthermore, Heim et al. [31] found increased activity in BA 44/6 when German monolinguals were asked to generate the determiner (*der, die or das* gender marked *the* in German) for a picture compared to simply naming the picture. These results are consistent with the view that monolinguals generate the determiner in order to determine the gender of opaque items. BA 44/6 indicates the need for increased phonological retrieval demands, the anterior insula indicates the need for increased articulatory demands, and BA 44/45 shows increased activity because of the syntactic computation that occurs when checking determiner-noun agreement. In short, the neural data are consistent with the view that monolinguals covertly form a small syntactic phrase when retrieving the gender of opaque items. This strategy was confirmed in post-experimental interviews.

A subsequent unpublished study compared early Spanish-English bilinguals with late English-Spanish bilinguals using the gender decision described above. Early Spanish-English bilinguals are of interest because they are dominant in English but learn Spanish first [for further work with this population see 9, 11, 12, 32]. Participants were matched on proficiency in Spanish using tests of vocabulary, reading and syntax. Furthermore, participants were matched on performance in the gender decision task. Although both groups showed increased activity for the opaque items, each group showed a different pattern of activity. The late English-Spanish bilinguals showed a large area of increased activity which extended from the anterior insula into BA 47. The early Spanish-English bilinguals showed increased activity just superior to this in BA 44/45. Direct comparisons between the groups revealed increased activity in BA 47 for the late bilinguals. The results confirm that AoA modulates activity on grammatical tasks. Furthermore, it reveals that these differences are graded in nature. That is, for transparent items the group differences are very small. However, for opaque items the results reveal much larger differences. In short, not all grammatical functions show large AoA effects.

1.2 AoA effects during semantic processing

Work conducted both in my laboratory and in collaboration with others has begun to shed light on the central questions that will be addressed in the current proposal. First of all, work in my laboratory has confirmed the presence of AoA for non-grammatical processing. In a first study, I asked monolinguals to name a set of pictures in which AoA and word frequency were orthogonally manipulated [33]. A main effect of AoA remained even when controlling for word frequency and even when naming was delayed. More recent work in collaboration with colleagues at the Max Planck Institute of Cognitive Neuroscience has sought to uncover the neural correlates of word AoA in monolinguals performing an auditory and visual lexical decision task (Press right button if it is a word, press the left button if it is a pseudoword such as "mave") while being scanned with fMRI. Results revealed that the precuneus which is known to play a role in automatic retrieval from memory was activated for early learned words across auditory and visual presentation modalities. Additional activity in the auditory cortex was observed specifically for the reading of early acquired words. Late learned words revealed increased activity in BA 45/47 indicating more complex semantic retrieval. These results confirm that reaction time AoA effects are robust. Second, it appears that early lexical memories may be more automatic or auditory in nature whereas late learned lexical memory most likely involves complex retrieval. This latter result is consistent with findings from monolingual simulations of AoA [34]. This finding confirms that AoA effects appear in monolinguals for lexical tasks which do NOT involve grammatical processing. Furthermore, a number of studies have found that AoA effects appear in monolingual semantic tasks [35]. Hence, AoA effects are quite pervasive.

1.3 Overlap across languages

As noted earlier, work in the bilingual imaging literature has found that semantic effects were more sensitive to proficiency. A number of studies have confirmed this basic finding [7, 36]. Recent work in our laboratory Using fMRI late high proficient German-English second language learners were tested in L1 (German) and L2 (English) with concrete and abstract words that showed maximal overlap (cognates) in orthography or not (noncognates). All words were translation equivalents. Participants decided whether a visually presented word was abstract or concrete. Results revealed a graded language difference in neural activity with abstract non-cognates showing the most activation differences across languages and concrete cognates showing the fewest differences. Specifically, non-cognates showed more activity than cognates in superior BA 44 (near BA 6), BA 44/45 and in the insula extending into BA 47 in L2. There were no significant differences observed in L2 or for comparisons which looked for increased activity for cognates relative to noncognates. A second study using lexical decision yielded results which are consistent with those found in the first study. Taken together our results show that the amount of differential neural activity across languages depends on orthographic and semantic overlap. In short, a less proficient second language reveals a difference in items which overlap the least across languages (noncognates).

Recent work by Tokowicz and MacWhinney [37] sheds light on the nature of transfer in late second language learners. In that study, participants were asked to make grammaticality judgments to sentences which varied in the extent to which syntactic functions overlapped across languages. Participants brain activity was measured using ERP's. The first type of functions involved tense marking which is similar across languages. The second type of function involved determiner-noun agreement (las casas vs. la casas). Like Spanish number in English is marked on the noun (houses). However, unlike Spanish there is no need for the noun to agree with the determiner (the houses). Participants were also asked to make decisions about sentences which manipulated gender agreement, a function which is unique to Spanish (la casa vs. el casa). The results revealed increased activity for noun-verb agreement, a function which is similar across languages. However, participants did not show ERP differences for gender or number agreement in Spanish. Finally, results for determiner-noun gender agreement revealed ERP differences for this function. However, the distribution of the signal was diffuse. Taken together these results suggest that the nature of L1 influences brain responses to L2 during early learning. Furthermore, it confirms that functions which overlap across languages are easier to track in L2 than those which are not. These results suggest that in both semantic and syntactic tasks there is an effect of overlap. However, they leave open the question of whether AoA effects may interact with overlap. If syntactic function tend to rely on overlapping information, then AoA effects may be more dramatic in this domain. This would predict that AoA effects should be larger for semantic tasks for items with less conceptual overlap across languages.

1.4 FMRI studies of grammatical processing

Taken together the results reviewed are consistent with the view that both semantic and grammatical processing are graded in nature and that this continuity modulates differences at the neural level in early and late bilinguals. However, these results leave some questions unanswered. Unpublished work conducted by Hernandez et al. has looked at semantic processing in late German-English bilinguals. Cross-language differences could be due both to English being learned late and being the less dominant language. Follow up studies comparing early bilinguals and late bilinguals would help to elucidate whether there is indeed an effect of cognate status and concreteness in both groups and whether the pattern of activity differs across groups. This would also help to clarify whether AoA effects appear in a task which is more effortful and whether less overlap (i.e. noncognates) yields larger cross-group differences. The notion of overlap is also be important when considering AoA differences in the neural activity associated with syntactic processing. Previous studies have found that overlap between languages affects the speed with which a syntactic function is learned in late L2 learners. Significant differences in the pattern of neural activity between early and late learners of Spanish have been found during gender decision for opaque items. That is, when grammatical functions have little overlap across languages, groups which differ on AoA show differences in neural activity. However, when these functions are easier then there are much smaller differences. Finally, future studies will test whether the effects of difficulty and overlap extend to other syntactic functions and conditions of syntactic violation Of particular interest, will be the comparison of violation conditions across groups. Violations are known to lead to increased activity in L2 relative to L1 [38].

1.5 Theoretical Accounts of AoA

Despite consistently finding maturational effects on syntactic processing, very few accounts have stipulated the underlying mechanism (aside from emphasizing a general maturational constraint) for this effect. More recently, Ullman and colleagues [39-41] have proposed that second language acquisition can be viewed as being constrained by declarative and procedural memory. In this model, lexical learning is reliant on memorized facts whereas grammatical processing is dependent on rules and routines. Work in the literature has firmly established a frontal-basal ganglia circuit which is involved in procedural learning and a medial temporal lobe system which is involved in declarative memory, i.e. learning new facts. Ullman and colleagues provide considerable evidence to bolster their claim that grammatical and lexical processing rely on different neural systems. This includes evidence from grammatical processing in aphasia as well as Alzheimer's and Parkinson's disease [39]. They also present evidence that procedural learning ability decreases with age whereas declarative learning ability may actually improve with age . This framework sheds light on AoA syntactic effects that the neural correlates of syntactic processing are more sensitive to AoA [42, 43] because of their reliance on procedural memory which is affected to a greater extent by maturational constraints. Within this model L2 learners must rely on declarative memory for grammatical processing. This, in turn, predicts that late L2 learners will show increased activity of areas involved in declarative memory during grammatical processing relative to L1 learners.

The procedural/declarative account for AoA effects, however, cannot account for results in the monolingual literature. For a long time, theorists have suggested that AoA was due to differences in phonological completeness [44]. In this view, early learned words are represented in a phonologically complete manner whereas late learned words have to be assembled around these phonological primitives. Whereas this hypothesis is consistent with some aspects of the data such as slower picture naming times [45-47] they are less compatible with other effects [48]. More recent work using connectionist simulations suggest that early learned items are favored because the network is biased to "recognize" these items. Recognizing late learned words, however, requires more effortful retrieval because a network is using connection weights that are optimized for early learned words. The most interesting aspect of this model is that it suggests that AoA effects are very general effects. As such second language acquisition, bilingualism and language processing in general serve as methods to investigate the general mechanisms that are involved in learning.

References

1. Gomez-Tortosa, E., et al., *Selective deficit of one language in a bilingual patient following surgery in the left perisylvian area.* Brain and Language, 1995. **48**: p. 320-325.
2. Paradis, M., *Bilingualism and aphasia*, in *Studies in Neurolinguistics*, H. Whitaker and H.A. Whitaker, Editors. 1977, Academic Press: New York. p. 65-121.
3. Paradis, M., ed. *Bilingual aphasia 100 years later: Consensus and controversies*. 1995, Pergamon: Oxford.
4. Perani, D., et al., *Brain processing of native and foreign languages.* Neuroreport, 1996. **7**: p. 2439-2444.
5. Perani, D., et al., *The bilingual brain: Proficiency and age of acquisition of the second language.* Brain, 1998. **121**(10): p. 1841-1852.
6. Kohnert, K., *Elderly bilingual performance on the Boston Naming Test: Preliminary norms*. 1998.
7. Chee, M.W., et al., *Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. Blood oxygen level dependent.* Neuroimage, 2001. **13**(6 Pt 1): p. 1155-63.
8. Pallier, C., et al., *Brain imaging of language plasticity in adopted adults: can a second language replace the first?* Cereb Cortex, 2003. **13**(2): p. 155-61.
9. Hernandez, A.E., E. Bates, and L.X. Avila, *Processing across the language boundary: A cross modal priming study of Spanish-English bilinguals.* Journal of Experimental Psychology: Learning, Memory, &Cognition, 1996. **22**: p. 846-864.

10. Hernandez, A.E. and K. Kohnert, *Aging and language switching in bilinguals.* Aging, Neuropsychology and Cognition, 1999. **6**: p. 69-83.
11. Hernandez, A.E. and I. Reyes, *Within- and between-language priming differ: evidence from repetition of pictures in Spanish-English bilinguals.* J Exp Psychol Learn Mem Cogn, 2002. **28**(4): p. 726-34.
12. Kohnert, K.J., A.E. Hernandez, and E. Bates, *Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish & English.* Brain & Language, 1998. **65**(3): p. 422-440.
13. Iverson, P., et al., *A perceptual interference account of acquisition difficulties for non-native phonemes.* Cognition, 2003. **87**(1): p. B47-57.
14. Flege, J.E., M.J. Munro, and I.R.A. MacKay, *Effects of age of second-language learning on the production of English consonants.* Speech Communication, 1995. **16**(1): p. 1-26.
15. Munro, M.J., J.E. Flege, and I.R.A. MacKay, *The effects of age of second language learning on the production of English vowels.* Applied Psycholinguistics, 1996. **17**(3): p. 313-334.
16. Weber-Fox, C. and H.J. Neville, *Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers.* Journal of Cognitive Neuroscience, 1996. **8**: p. 231-256.
17. Wartenburger, I., et al., *Early setting of grammatical processing in the bilingual brain.* Neuron, 2003. **37**(1): p. 159-70.
18. Kutas, M. and S.A. Hillyard, *Event-related brain potentials to semantically inappropriate and surprisingly large words.* Biological Psychology, 1980. **11**: p. 99-116.
19. Kutas, M. and C. Van Petten, *Event-related brain potential studies of language*, in *Advances in Psychophysiology*, P.K. Ackles, J.R. Jennings, and M.G.H. Coles, Editors. 1988, JAI Press: Greenwich, Connecticut.
20. Friederici, A.D., A. Hahne, and A. Mecklinger, *Temporal structure of syntactic parsing: early and late event-related brain potential effects.* J Exp Psychol Learn Mem Cogn, 1996. **22**(5): p. 1219-48.
21. Dapretto, M., S. Bookheimer, and J. Mazziotta, *Form and Content: Dissociating syntax and semantics in sentence comprehension.* Neuron, 1999. **24**(2): p. 427-32.
22. Poldrack, R.A., et al., *Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex.* Neuroimage, 1999. **10**(1): p. 15-35.
23. D'Esposito, M., et al., *The neural basis of the central executive system of working memory.* 1995, 1995. **378**(6554): p. 279-281.
24. Hernandez, A.E., et al., *The neural correlates of grammatical gender decisions in Spanish.* Neuroreport, 2004. **15**(5): p. 863-6.
25. Moro, A., et al., *Syntax and the brain: disentangling grammar by selective anomalies.* Neuroimage, 2001. **13**(1): p. 110-8.
26. Kang, A.M., et al., *An event-related fMRI study of implicit phrase-level syntactic and semantic processing.* Neuroimage, 1999. **10**(5): p. 555-61.
27. Friederici, A.D., B. Opitz, and D.Y.v. Cramon, *Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types.* Cerebral Cortex, 2000. **10**(7): p. 698-705.
28. Miceli, G., et al., *The neural correlates of grammatical gender: an fMRI investigation.* J Cogn Neurosci, 2002. **14**(4): p. 618-28.
29. Bates, E., et al., *Voxel-based lesion-symptom mapping.* Nat Neurosci, 2003. **6**(5): p. 448-50.
30. Dronkers, N.F., *A new brain region for coordinating speech articulation.* Nature, 1996. **384**(6605): p. 159-61.
31. Heim, S., B. Opitz, and A.D. Friederici, *Broca's area in the human brain is involved in the selection of grammatical gender for language production: evidence from event-related functional magnetic resonance imaging.* Neurosci Lett, 2002. **328**(2): p. 101-4.
32. Kohnert, K., E. Bates, and A.E. Hernandez, *Balancing bilinguals: Lexical-semantic production and cognititve processing in children learning Spanish and English.* Journal of Speech, Language and Hearing Research, 1999. **42**: p. 1400-1413.
33. Meschyan, G. and A. Hernandez, *Age of acquisition and word frequency: determinants of object-naming speed and accuracy.* Mem Cognit, 2002. **30**(2): p. 262-9.
34. Ellis, A.W. and M.A. Lambon Ralph, *Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks.* Journal of Experimental Psychology: Learning, Memory, & Cognition, 2000. **26**(5): p. 1103-1123.
35. Brysbaert, M., I. Van Wijnendaele, and S. De Deyne, *Age-of-acquisition effects in semantic processing tasks.* Acta Psychol (Amst), 2000. **104**(2): p. 215-26.
36. Chee, M.W., et al., *Processing of visually presented sentences in Mandarin and English studied with fMRI.* Neuron, 1999. **23**(1): p. 127-37.
37. Tokowicz, N. and B. MacWhinney, *Implicit vs. Explicit Measures of Sensitivity to Violations in L2 Grammar: An Event-Related Potential Investigation.* Studies in Second Language Acquisition, in press.
38. Rueschemeyer, S.A., et al., *Cerebral representation of spoken language processing in bilinguals.*
39. Ullman, M., Corkin, S., Coppola, M., Hickok, G., Growden, J., & Koroshetz, W., *A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory and that grammatical rules are processed by the procedural system.* Journal of Cognitive Neuroscience, 1997. **9**: p. 289-299.
40. Ullman, M.T., *A neurocognitive perspective on language: the declarative/procedural model.* Nat Rev Neurosci, 2001. **2**(10): p. 717-26.

41. Ullman, M.T., *Contributions of memory circuits to language: the declarative/procedural model.* Cognition, 2004. **92**(1-2): p. 231-70.
42. Squire, L.R. and B.J. Knowlton, *The medial temporal lobe, the hippocampus, and the memory systems of the brain*, in *The new cognitive neurosciences*, M.S. Gazzaniga, Editor. 2000, MIT
Press: Cambridge, MA.
43. Eichenbaum, H. and N.J. Cohen, *From conditioning to conscious recollection: memory systems of the brain*. 2001, New York: Oxford University Press.
44. Brown, G.D. and F.L. Watson, *First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency.* Memory & Cognition, 1987. **15**(3): p. 208-216.
45. Ellis, A.W. and C.M. Morrison, *Real age-of-acquisition effects in lexical retrieval.* Journal of Experimental Psychology: Learning, Memory, & Cognition, 1998. **24**(2): p. 515-523.
46. Barry, C., C.M. Morrison, and A.W. Ellis, *Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement.* Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 1997. **50A**(3): p. 560-585.
47. Morrison, C.M., A.W. Ellis, and P.T. Quinlan, *Age of acquisition, not word frequency, affects object naming, not object recognition.* Memory & Cognition, 1992. **20**(6): p. 705-714.
48. Lewis, M.B., A.J. Chadwick, and H.D. Ellis, *Exploring a neural-network account of age-of-acquisition effects using repetition priming of faces.* Mem Cognit, 2002. **30**(8): p. 1228-37.

Developmental Stages of Perception and Language Acquisition in a Physically Grounded Robot

Peter Ford Dominey, Jean-David Boucher
Institut des Sciences Cognitives, CNRS 67 Blvd. Pinel, 69675 Bron Cedex, France,
dominey@isc.cnrs.fr
http://www.isc.cnrs.fr/dom/dommenu-en.htm

Abstract

The objective of this research is to develop a system for language learning based on a minimum of pre-wired language-specific functionality, that is compatible with observations of perceptual and language capabilities in the human developmental trajectory. In the proposed system, meaning (in terms of descriptions of events and spatial relations) is extracted from video images based on detection of position, motion, physical contact and their parameters. Mapping of sentence form to meaning is performed by learning grammatical constructions that are retrieved from a construction inventory based on the constellation of closed class items uniquely identifying the target sentence structure. The resulting system displays robust acquisition behavior that reproduces certain observations from developmental studies, with very modest "innate" language specificity. Most importantly, the demonstrates a certain degree of autonomy in adapting to the structural regularities of the environment.

1. Introduction

A challenge of developmental robotics is to demonstrate the successive emergence of behaviors in a developmental progression of increasing processing power and complexity. A particularly interesting avenue for this methodology is in language processing and the binding of language to developing perceptual relations. Generative linguists have posed the significant challenge to such approaches via the claim that the learning problem is too underconstrained and must thus be addressed by a highly pre-specified Universal Grammar (Chomsky 1995). The current research proposes an alternative, identifying a restricted set of functional requirements for language acquisition, and then demonstrating a possible framework for the successive emergence of these behaviors in developmentally plausible systems, culminating in a grounded robotic system that can learn a small language about visual scenes that it observes.

1.1. Functional Requirements:

We adopt a construction based approach to language in which acquisition is based on learning mappings between grammatical structure and meaning structure (Goldberg 1995). In this context, the system should be capable of: (1) extracting meaning from the environment, (2) learning mappings between grammatical structure and meaning, and (3) identifying-discriminating between different grammatical structures of input sentences. In the following sections we outline how these requirements can be satisfied in a biologically and developmentally plausible manner.

In this developmental context, Mandler (1999) suggested that the infant begins to construct meaning from the scene based on the extraction of perceptual primitives. From simple representations such as contact, support, attachment (Talmy 1988) the infant could construct progressively more elaborate representations of visuospatial meaning. Thus, the physical event "collision" is a form of the perceptual primitive "contact". Kotovsky & Baillargeon (1998) observed that at 6 months, infants demonstrate sensitivity to the parameters of objects involved in a collision, and the resulting effect on the collision, suggesting indeed that infants can represent contact as an event predicate with agent and patient arguments. Similarly, Quinn et al. (2002) have demonstrated that at 6-7 months, infants are sensitive to binary spatial relations such as above and below.

Bringing this type of perception into the robotic domain, Siskind (2001) has demonstrated that force dynamic primitives of contact, support, attachment can be extracted from video event sequences and used to recognize events including pick-up, put-down, and stack based on their characterization in an event logic. Related results have been achieved by Steels and Baillie (2003). The use of these intermediate representations renders the system robust to variability in motion and view parameters. Most importantly, this research demonstrated that the lexical semantics for a number of verbs could be established by automatic image processing.

Once meaning is extracted from the scene, the significant problem of mapping sentences to meanings remains. The functionalist perspective holds that learning plays a central role in language acquisition. The infant develops an inventory of grammatical constructions as mappings from form to meaning (Goldberg 1995). These constructions are initially rather fixed and specific, and later become generalized into a more abstract compositional form employed by the adult (Tomasello 1999). In this context, construction of the relation between perceptual and cognitive representations and grammatical form plays a central role in learning language (e.g. Feldman et al. 1990, 1996; Langacker 1991; Mandler 1999; Talmy 1998).

These issues of learnability and innateness have provided a rich motivation for simulation studies that have taken a number of different forms. Elman (1990) demonstrated that recurrent networks are sensitive to predictable structure in grammatical sequences. Subsequent studies of grammar induction demonstrate how syntactic structure can be recovered from sentences (e.g. Stolcke & Omohundro 1994). From the "grounding of language in meaning" perspective (e.g. Feldman et al. 1990, 1996; Langacker 1991; Goldberg 1995), Chang & Maia (2001) exploited the relations between action representation and simple verb frames in a construction grammar approach, and Cottrell et al. (1990) associated sequences of words with simple image sequences. In effort to consider more complex grammatical forms, Miikkulainen (1996) demonstrated a system that learned the mapping between relative phrase constructions and multiple event representations, based on the use of a stack for maintaining state information during the processing of the next embedded clause in a recursive manner.

In a more generalized approach, Dominey (2000) exploited the regularity that sentence to meaning mapping is encoded in all languages by word order and grammatical marking (bound or free) (Bates et al. 1982). That model was based on the functional neurophysiology of cognitive sequence and language processing and an associated neural network model that has been demonstrated to simulate interesting aspects of infant (Dominey & Ramus 2000) and adult language processing (Dominey et al. 2003).

1.2. Objectives

The goals of the current study are fourfold: First to test the hypothesis that meaning can be extracted from visual scenes based on the detection of contact and its parameters in an approach similar to but significantly simplified from Siskind (2001); Second to determine whether the model of Dominey (2000) can be extended to handle embedded relative clauses; Third to demonstrate that these two systems can be combined to perform miniature language acquisition; and finally to demonstrate that the combined system can provide insight into the developmental progression in human language acquisition without the necessity of a pre-wired parameterized grammar system (Chomsky 1995).

Figure 1. Structure-Mapping Architecture

2. The Behavioral Context

As illustrated in Figure 1, the human experimenter enacts and simultaneously narrates visual scenes made up of events that occur between a red cylinder, a green block and a blue semicircle or "moon" on a black matte table surface. A video camera above the surface provides a video image that is processed by a color-based recognition and tracking system (Smart – Panlab, Barcelona Spain) that generates a time ordered sequence of the contacts that occur between objects that is subsequently processed for event analysis (below). The simultaneous narration of the ongoing events is processed by a commercial speech-to-text system (IBM ViaVoice^{TM}). Speech and vision data were acquired and then processed off-line yielding a data set of matched sentence – scene pairs that were provided as input to the structure mapping model. A total of ~300 <sentence, scene> pairs were tested in the following experiments.

3. Requirement 1: Extracting Meaning

For a given video sequence (see snapshot in Figure 2) the visual scene analysis generates the corresponding event description in the format *event(agent, object, recipient).*

Single Event Labeling

Events are defined in terms of contacts between elements. A contact is defined in terms of the time at which it occurred, the agent, object, and duration of the contact. The agent is determined as the element that had a larger relative velocity towards the other element involved in the contact. Based on these parameters of contact, scene events are recognized as follows:

Touch(agent, object): A single contact, in which (a) the duration of the contact is inferior to *touch_duration* (1.5 seconds), and (b) the *object* is not displaced during the duration of the contact.

Push(agent, object): Similar to touch, with a greater contact duration, superior or equal to *touch_duration* and inferior to *take_duration* (5 sec), and object displacement.

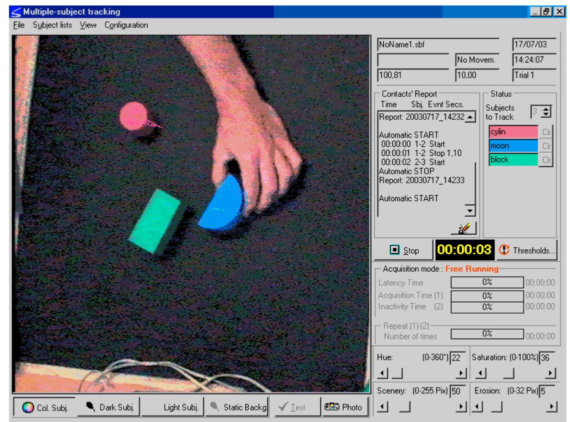
Take(agent, object): A single contact in which (a) the duration of contact is superior or equal to *take_duration*, (b) the object is displaced during the contact, and (c) the agent and object remain in contact.

Take(agent, object, source): Multiple contacts, as the agent takes the object from the source. Same as Take(a,o), and for the optional second contact between agent and source (a) the duration of the contact is inferior to *take_duration*, and (b) the agent and source do not remain in contact. Finally, contact between the object and source is broken during the event.

Give(agent, object, recipient): Multiple contacts as agent takes object, then initiates contact between object and recipient.

These event labeling templates form the basis for a template matching algorithm that labels events based on the contact list, similar to the spanning interval and event logic of Siskind (2001).

Complex "Hierarchical" Events: The events described above are simple in the sense that there have no hierarchical structure. This imposes serious limitations on the syntactic complexity of the corresponding sentences (Feldman et al. 1996, Miikkulainen 1996). The sentence "The block that pushed the moon was touched by the triangle" illustrates a complex event that exemplifies this issue. The corresponding compound event will be recognized and represented as a pair of temporally successive



simple event descriptions, in this case: *push(block, moon)*, and *touch(triangle, block).* The "block" serves as the link that connects these two simple events in order to form a complex hierarchical event.

Figure 2. Snapshot of scene event processing.

4. Requirement 2: Mapping Sentences to Meaning

Our approach is based on the cross-linguistic observation that open class words (e.g. nouns, verbs, adjectives and adverbs) are assigned to their thematic roles based on word order and/or grammatical function words or morphemes (Bates et al. 1982). The mapping of sentence form onto meaning (Goldberg 1995) takes place at two distinct levels: Words are associated with individual components of event descriptions, and grammatical structure is associated with functional roles within scene events (Fig 3). The first level has been addressed by Siskind (1996), Roy & Pentland (2000) and Steels (2001) and we treat it here in a relatively simple but effective manner. Our principle interest lies more in the second level of mapping between scene and sentence structure.

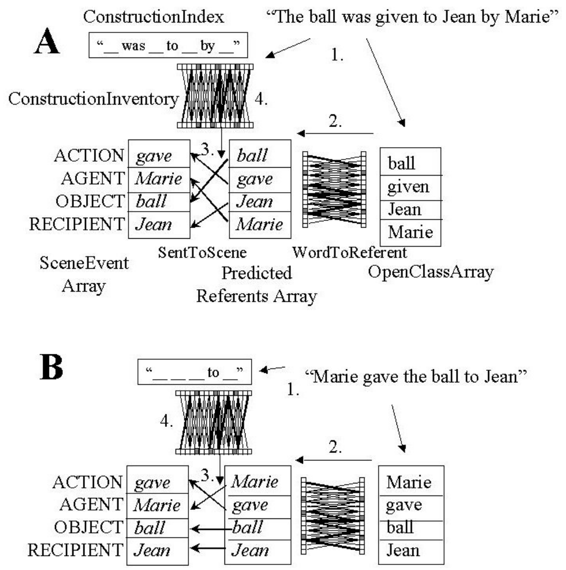


Figure 3. Model Overview: Processing of active and passive sentence types in A, B, respectively. On input, Open class words populate the Open Class Array (OCA), and closed class words populate the Construction index. Visual Scene Analysis populates the Scene Event Array (SEA) with the extracted meaning as scene elements. Words in OCA are translated to Predicted Referents via the WordToReferent mapping to populate the Predicted Referents Array (PRA). PRA elements are mapped onto their roles in the Scene Event Array (SEA) by the SentenceToScene mapping, specific to each sentence type. This mapping is retrieved from Construction Inventory, via the ConstructionIndex that encodes the closed class words that characterize each sentence type. Words in sentences, and elements in the scene are coded as single ON bits in respective 25-element vectors.

Word Meaning

In the initial learning phases there is no influence of syntactic knowledge and the word-referent associations are stored in the WordToReferent matrix (Eqn 1) by associating every word with every referent in the current scene ($\alpha=1$), exploiting the cross-situational regularity (Siskind 1996) that a given word will have a higher coincidence with referent to which it refers than with other referents. This initial word learning contributes to learning the mapping between sentence and scene structure (Eqn. 4, 5 & 6 below). Then, knowledge of the syntactic structure, encoded in SentenceToScene can be used to identify the appropriate referent (in the SEA) for a given word (in the OCA), corresponding to a zero value of α in Eqn. 1. In this "syntactic bootstrapping" for the new word "gugle," for example, syntactic knowledge of Agent-Event-Object structure of the sentence "John pushed the gugle" can be used to assign "gugle" to the object of push.

$$\text{WordToReferent}(i,j) = \text{WordToReferent}(i,j) + \text{OCA}(k,i) * \text{SEA}(m,j) * \text{Max}(\alpha, \text{SentenceToScene}(m,k)) \quad (1)$$

Indices: k(1:6) - words; m(1:6) – scene elements; i(1:25), j(1:25) – elements in word and scene item vectors, respectively.

Mapping Sentence to Meaning

In terms of the architecture in Figure 3, this mapping can be characterized in the following successive steps. First, words in the Open Class Array are decoded into their corresponding scene referents (via the WordToReferent mapping) to yield the Predicted Referents Array that contains the translated words while preserving their original order from the OCA (Eqn 2).

$$\text{PRA}(m,j) = \sum_{i=1}^{n} \text{OCA}(m,i) * \text{WordToReferent}(i,j) \quad (2)$$

Next, each sentence type will correspond to a specific *form to meaning* mapping between the PRA and the SEA. encoded in the SentenceToScene array. The problem will be to retrieve for each sentence type, the appropriate corresponding SentenceToScene mapping.

5. Requirement 3: Discriminating Between Grammatical Forms

The first step in discriminating between grammatical structures is to discriminate between open class (e.g. nouns, verbs) and closed class (e.g. determiners, prepositions) words. Newborn infants are sensitive to the perceptual properties that distinguish these two categories (Shi et al. 1999), and in adults, these categories are processed by dissociable neurophysiological systems (Brown et al. 1999). Similarly, artificial neural networks can also learn to make this function/content distinction (Morgan et al. 1996, Blanc et al. 2003). Thus, for the speech input that is provided to the learning model, open and closed class words are directed to separate processing streams that preserve their order and identity, as indicated in Figure 3.

Given this capability to discriminate between open and closed class words, we are still faced with the problem of using this information to discriminate between different sentence types. To solve this problem, we recall that each sentence type will have a unique constellation of closed class words and/or bound morphemes (Bates et al. 1982) that can be coded in a ConstructionIndex (Eqn.3) that forms a unique identifier for each sentence type, shifting the current contents by the index of the ON bit in FunctionWord, then ANDing the FunctionWord vector. The appropriate SentenceToScene mapping for each sentence type can be indexed in ConstructionInventory by its corresponding ConstructionIndex.

ConstructionIndex = $f_{circularShift}$(ConstructionIndex, FunctionWord) (3)

The link between the ConstructionIndex and the corresponding SentenceToScene mapping is established as follows. As each new sentence is processed, we first reconstruct the specific SentenceToScene mapping for that sentence (Eqn 4), by mapping words to referents (in PRA) and referents to scene elements (in SEA). The resulting, SentenceToSceneCurrent encodes the correspondence between word order (that is preserved in the PRA Eqn 2) and thematic roles in the SEA. Note that the quality of SentenceToSceneCurrent will depend on the quality of acquired word meanings in WordToReferent. Thus, syntactic learning requires a minimum baseline of semantic knowledge. Given the SentenceToSceneCurrent mapping for the current sentence, we can now associate it in the ConstructionInventory with the corresponding function word configuration or ConstructionIndex for that sentence, expressed in (Eqn 5). In Eqns 5, 6 SentenceToScene is linearized for simplification.

$$\text{SentenceToSceneCurrent}(m,k) = \sum_{i=1}^{n} \text{PRA}(k,i) * \text{SEA}(m,i) \tag{4}$$

$$\text{ConstructionInventory}(i,j) = \text{ConstructionInventory}(i,j) + \text{ConstructionIndex}(i) * \text{SentenceToSceneCurrent}(j) \tag{5}$$

Finally, once this learning has occurred, for new sentences we can now extract the SentenceToScene mapping from the learned ConstructionInventory by using the ConstructionIndex as an index into this associative memory, illustrated in Eqn. 6.

$$\text{SentenceToScene}(i) = \sum_{i=1}^{n} \text{ConstructionInventory}(i,j) * \text{ConstructionIndex}(j) \tag{6}$$

To accommodate the dual scenes for complex events Eqns. 4-7 are instantiated twice each, to represent the two components of the dual scene. In the case of simple scenes, the second component of the dual scene representation is null.

We evaluate performance by using the WordToReferent and SentenceToScene knowledge to construct for a given input sentence the "predicted scene". That is, the model will construct an internal representation of the scene that should correspond to the input sentence. This is achieved by first converting the Open-Class-Array into its corresponding scene items in the Predicted-Referents-Array as specified in Eqn. 2. The referents are then reordered into the proper scene representation via application of the SentenceToScene transformation as described in Eqn. 7.

$$\text{PSA}(m,i) = \text{PRA}(k,i) * \text{SentenceToScene}(m,k) \tag{7}$$

When learning has proceeded correctly, the predicted scene array (PSA) contents should match those of the scene event array (SEA) that is directly derived from input to the model. We then quantify performance error in terms of the number of mismatches between PSA and SEA.

6. Experimental results

Hirsh-Pasek & Golinkoff (1996) indicate that children use knowledge of word meaning to acquire a fixed SVO template around 18 months, then expand this to non-canonical sentence forms around 24+ months. Tomasello (1999) indicates that fixed grammatical constructions will be used initially, and that these will then provide the basis for the development of more generalized constructions (Goldberg 1995). The following experiments attempt to follow this type of developmental progression. Training results in changes in the associative WordToReferent mappings encoding the lexicon, and changes in the ConstructionInventory encoding the form to meaning mappings, indexed by the ConstructionIndex.

A. Learning of Active Forms for Simple Events

1. Active: The block pushed the triangle.
2. Dative: The block gave the triangle to the moon.

For this experiment, 17 scene/sentence pairs were generated that employed the 5 different events, and narrations in the active voice, corresponding to the grammatical forms 1 and 2. The model was trained for 32 passes through the 17 scene/sentence pairs for a total of 544 scene/sentence pairs. During the first 200 scene/sentence pair trials, α in Eqn. 1 was 1 (i.e. no syntactic bootstrapping before syntax is acquired), and thereafter it was 0. This was necessary in order to avoid the random effect of syntactic knowledge on semantic learning in the initial learning stages. The trained system displayed error free performance for all 17 sentences, and generalization to new sentences that had not previously been tested.

B. Passive forms

This experiment examined learning active and passive grammatical forms, employing grammatical forms 1-4. Word meanings were used from Experiment A, so only the structural SentenceToScene mappings were learned.

3. Passive: The triangle was pushed by the block.
4. Dative Passive: The moon was given to the triangle by the block.

Seventeen new scene/sentence pairs were generated with active and passive grammatical forms for the narration. Within 3 training passes through the 17 sentences (51 scene/sentence pairs), error free performance was achieved, with confirmation of error free generalization to new untrained sentences of these types. The rapid learning indicates the importance of lexicon in establishing the form to meaning mapping for the grammatical constructions.

C. Relative forms for Complex Events

Here we consider complex scenes narrated by relative clause sentences. Eleven complex scene/sentence pairs were generated with narration corresponding to the grammatical forms indicated in 5 – 10:

5. The block that pushed the triangle touched the moon.
6. The block pushed the triangle that touched the moon.
7. The block that pushed the triangle was touched by the moon.
8. The block pushed the triangle that was touched the moon.
9. The block that was pushed by the triangle touched the moon.
10. The block was pushed by the triangle that touched the moon.

After presentation of 88 scene/sentence pairs, the model performed without error for these 6 grammatical forms, and displayed error-free generalization to new sentences that had not been used during the training for all six grammatical forms

D. Generalization to Extended Construction Set

As illustrated above the model can accommodate 10 distinct form-meaning mappings or grammatical constructions, including constructions involving "dual" events in the meaning representation that correspond to relative clauses. Still, this is a relatively limited size for the construction inventory. We have subsequently demonstrated that the model can accommodate 38 different grammatical constructions that combine verbs with two or three arguments, active and passive forms and relativization, along with additional sentence types including: conjoined (John took the key and opened the door), reflexive (The boy said that the dog was chased by the cat), and reflexive pronoun (The block said that it pushed the cylinder) sentence types. The consideration of these sentence types requires us to address how their meanings are represented. Conjoined sentences are represented by the two corresponding events, e.g. *took(John, key), open(John, door)* for the conjoined example above. Reflexives are represented, for example, as *said(boy), chased(cat, dog)*. This assumes indeed, for reflexive verbs (e.g. said, saw), that the meaning representation includes the second event as an argument to the first. Finally, for the reflexive pronoun types, in the meaning representation the pronoun's referent is explicit, as in *said(block), push(block, cylinder)* for "The block said that it pushed the cylinder."

For this testing, the ConstructionInventory is implemented as a lookup table in which the ConstructionIndex is paired with the corresponding SentenceToScene mapping during a single learning trial. Based on the tenets of the construction grammar framework (Goldberg 1995), if a sentence is encountered that has a form (i.e. ConstructionIndex) that does not have a corresponding entry in the ConstructionInventory, then a new construction is defined. Thus, one exposure to a sentence of a new construction type allows the model to generalize to any new sentence of that type. In this sense, developing the capacity to handle a simple initial set of constructions leads to a highly extensible system. Using the training procedures as described above, with a pre-learned lexicon (WordToReferent), the model successfully learned all of a total of 38 distinct grammatical constructions, and demonstrated generalization to new sentences that it was not trained on.

That the model can accommodate these 38 different grammatical constructions with no modifications indicates its capability to generalize. This translates to a (partial) validation of the hypothesis that across languages, thematic role assignment is encoded by a limited set of parameters including word order and grammatical marking, and that distinct grammatical constructions will have distinct and identifying ensembles of these parameters.

E. Extension of the Construction Framework to Spatial Relations

Part of the developmental framework holds that existing processes can provide the basis for the emergence of new behavioral functionality. We have seen how the construction framework provides a basis for encoding the structural mappings between sentences and meaning in an organized and generalized manner. In theory this construction framework should extend to analogous cognitive domains. Here, we will investigate how this framework can be extended to the domain of spatial relations. Quinn et al (2002) have demonstrated that by the age of 6-7 months, infants can learn binary spatial relations such as left, right, above, below in a generalized manner, as revealed by their ability to discriminate in familiarization-test experiments. That is, they can apply this relational knowledge to scenes with new objects in these spatial relations.

In theory, the predicate-argument representation for event structure that we have described above can provide the basis for representing spatial relations in the form Left(X,Y), Above(X,Y) etc. where X is the object that holds the spatial relation with the referent Y. That is, Left(X,Y) corresponds to "X is left of Y".

In order to extract spatial relations from vision we return to the visual processing system described above. Based on the observations of Quinn et al. (2002) we can consider that by 6-7 months, the perceptual primitives of Relation(X,Y) are available, where Relation corresponds to Left, Right, Above and Below. The mapping of sentence structure onto the predicate argument then can proceed as described above for event meaning. One interesting problem presents itself however.

A

B

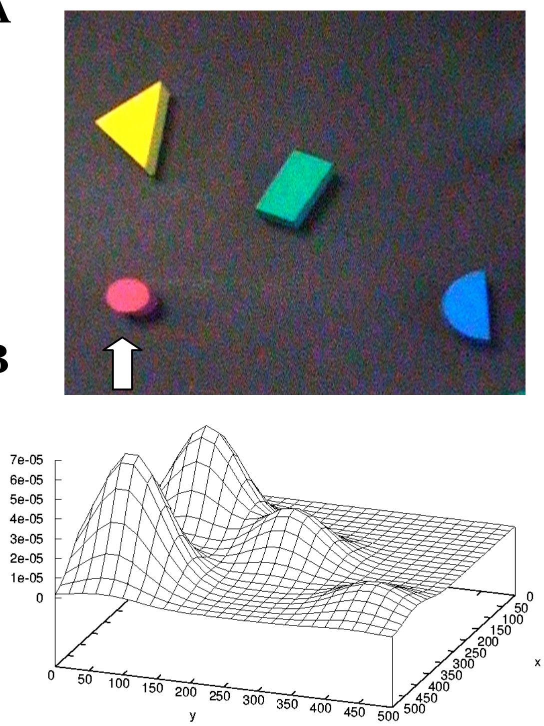


Figure 4. Spatial Attention for Relation Selection. The human user shows the robot a spatial relation and describes it. How does the robot know which of the multiple relations is the relevant one? A. The cylinder (lower left) has been moved into its current position, and now holds spatial relations with the three other objects. B. Based on parameters of (1) minimal distance from the target object and (2) minimal angular distance from the four principal directions (above, below, left, right).. In this case, the most relevant relation (indicated by the height of the two highest peaks) is Below(Cylinder, Triangle).

Figure 4 illustrates the spatial configuration after a human user has placed the cylinder in its current position and said "The cylinder is below the triangle". A simple attention mechanism based on motion is used to select the cylinder as the target object, but the intended referent for the "below" relation could be any one of the multiple other objets, and so the problem of referential ambiguity must be resolved. We hypothesize that this redundancy is resolved based on two perceptual parameters. First, spatial proximity will be used. That is, the observer will give more attentional preference to relations involving the target object and other objects that are closest to it. The second parameter is the angular "relevance" of the relations, quantified in terms of the angular distance from the cardinal positions above, below, left and right. Figure 4B represents the application of this perceptual attention mechanism that selects the relation Below(Cylinder, Triangle) as the most relevant, revealed by the height of the peak for the triangle in 4B.

We collected data training data in which a human observer demonstrated and narrated spatial relations with the four objects. The spatial attention mechanism extracted for each case the most relevant spatial relation, and the resulting <sentence, relation-meaning> pairs were used for training in the same procedure as in condition A for active sentences and simple events. The model demonstrated successful learning of the four object names and the four spatial relation terms, and could generalize this knowledge to a new <sentence, relation-meaning> generalization data set.

7. Discussion and Conclusion

The current study demonstrates (1) that the perceptual primitive of contact (available to infants at 5 months), can be used to perform event description in a manner that is similar to but significantly simpler than Siskind (2001), and can be extended to accommodate spatial relation encoding (2) that a novel implementation of principles from construction grammar can be used to map sentence form to these meanings together in an integrated system, (3) that relative clauses can be processed in a manner that is similar to, but requires less specific machinery (e.g. no stack) than that in Miikkulainen (1996), and finally (4) that the resulting system displays robust acquisition behavior that reproduces certain observations from developmental studies with very modest "innate" language specificity.

The goal was to identify minimal event recognition and form-to-meaning mapping capabilities that could be integrated into a coherent system that performs at the level of a human infant in the first years of development when the construction inventory is being built up. Rather than prewiring the language grammar, we demonstrate that the system can autonomously adapt to the regularities in the sentence form to meaning mappings in a systematic generalized manner.

8. Acknowledgments

Supported by the OHLL, EuroCores OMLL, French ACI Integrative and Computational Neuroscience, and HFSP MCILA Projects.

9. References

Bates E, McNew S, MacWhinney B, Devescovi A, Smith S (1982) Functional constraints on sentence processing: A cross linguistic study, *Cognition* (11) 245-299.

Blanc JM, Dodane C, Dominey PF (2003) Temporal processing for syntax acquisition: A Simulation Study, In Press, *Proceedings of the 25th Ann Conf. Cog. Sci. Soc.*, MIT, Cambridge MA

Brown CM, Hagoort P, ter Keurs M (1999) Electrophysiological signatures of visual lexical processing : Open- and closed-class words. *Journal of Cognitive Neuroscience*. 11 :3, 261-281

Chomsky N. (1995) The Minimalist Program. MIT

Chang NC, Maia TV (2001) Grounded learning of grammatical constructions, *AAAI Spring Symp. On Learning Grounded Representations*, Stanford CA.

Cottrel GW, Bartell B, Haupt C. (1990) Grounding Meaning in Perception. In Proc. GWAI90, 14th German Workshop on Artificial Intelligence, pages 307--321, Berlin, New York,. Springer Verlag.

Dominey PF, Ramus F (2000) Neural network processing of natural lanugage: I. Sensitivity to serial, temporal and abstract structure of language in the infant. *Lang. and Cognitive Processes*, 15(1) 87-127

Dominey PF (2000) Conceptual Grounding in Simulation Studies of Language Acquisition, *Evolution of Communication*, 4(1), 57-85.

Dominey PF, Hoen M, Lelekov T, Blanc JM (2003) Neurological basis of language in sequential cognition: Evidence from simulation, aphasia and ERP studies, *Brain and Language,* 86(2):207-25

Elman J (1990) Finding structure in time. *Cognitive Science*, 14:179-211.

Feldman JA, Lakoff G, Stolcke A, Weber SH (1990) Miniature language acquisition: A touchstone for cognitive science. In *Proceedings of the 12th Ann Conf. Cog. Sci. Soc.* 686-693, MIT, Cambridge MA

Feldman J., G. Lakoff, D. Bailey, S. Narayanan, T. Regier, A. Stolcke (1996). L0: The First Five Years. *Artificial Intelligence Review*, v10 103-129.

Goldberg A (1995) *Constructions*. U Chicago Press, Chicago and London.

Hirsh-Pasek K, Golinkof RM (1996) *The origins of grammar: evidence from early language comprehension.* MIT Press, Boston.

Kotovsky L, Baillargeon R, The development of calibration-based reasoning about collision events in young infants. 1998, *Cognition*, 67, 311-351

Langacker, R. (1991). *Foundations of Cognitive Grammar. Practical Applications, Volume 2*. Stanford University Press, Stanford.

Mandler J (1999) Preverbal representations and language, in P. Bloom, MA Peterson, L Nadel and MF Garrett (Eds) *Language and Space*, MIT Press, 365-384

Miikkulainen R (1996) Subsymbolic case-role analysis of sentences with embedded clauses. *Cognitive Science*, 20:47-73.

Morgan JL, Shi R, Allopenna P (1996) Perceptual bases of rudimentary grammatical categories, pp 263-286, in Morgan JL, Demuth K (Eds) *Signal to syntax*, Lawrence Erlbaum, Mahwah NJ, USA.

Quinn PC, Polly JL, Furer MJ, Dobson V, Nanter DB (2002) Young infants' performance in the object-variation version of the above-below categorization task. Infancy, 3, 323-347

Roy D, Pentland A (2002). Learning Words from Sights and Sounds: A Computational Model. *Cognitive Science*, 26(1), 113-146.

Shi R., Werker J.F., Morgan J.L. (1999) Newborn infants' sensitivity to perceptual cues to lexical and grammatical words, *Cognition*, Volume 72, Issue 2, B11-B21.

Siskind JM (1996) A computational study of cross-situational techniques for learning word-to-meaning mappings, *Cognition* (61) 39-91.

Siskind JM (2001) Grounding the lexical semantics of verbs in visual perception using force dynamics and event logic. *Journal of AI Research* (15) 31-90

Steels, L. (2001) Language Games for Autonomous Robots. *IEEE Intelligent Systems*, vol. 16, nr. 5, pp. 16-22, New York: IEEE Press.

Steels, L. and Baillie, JC. (2003). Shared Grounding of Event Descriptions by Autonomous Robots. Robotics and Autonomous Systems, 43(2-3):163--173. 2002

Stolcke A, Omohundro SM (1994) Inducing probablistic grammars by Bayseian model merging/ In *Grammatical Inference and Applications: Proc. 2nd Intl. Colloq. On Grammatical Inference*, Springer Verlag.

Talmy L (1988) Force dynamics in language and cognition. *Cognitive Science*, 10(2) 117-149.

Tomasello M (1999) The item-based nature of children's early syntactic development, *Trends in Cognitive Science*, 4(4):156-163

Texture Segmentation in 2D vs. 3D: Did 3D Developmentally Precede 2D?

Sejong Oh and Yoonsuck Choe
Department of Computer Science
Texas A&M University
3112 TAMU, College Station, TX 77843-3112
{sejong,choe}@tamu.edu

Abstract

Texture boundary detection (or segmentation) is an important capability in human vision. Usually, texture segmentation is viewed as a 2D problem, as the definition of the problem itself assumes a 2D substrate. However, an interesting hypothesis emerges when we ask a question regarding the nature *of textures:* What are textures, and why did the ability to discriminate texture evolve or develop? *A possible answer to this question is that textures naturally define physically distinct surfaces, thus, we can hypothesize that 2D texture segmentation may be an outgrowth of the ability to discriminate surfaces in 3D. In this paper, we investigated the relative difficulty of learning to segment textures in 2D vs. 3D configurations. It turns out that learning is faster and more accurate in 3D, very much in line with our expectation. Furthermore, we have shown that the learned ability to segment texture in 3D transfers well into 2D texture segmentation, bolstering our initial hypothesis, and providing a possible explanation for the developmental origin of 2D texture segmentation function in human vision.*

1. Introduction

Detection of a tiger in the shrub is a perceptual task that carries a life or death consequence for preys trying to survive in the jungle [1]. Here, figure-ground separation becomes an important perceptual capability. Figure-ground separation is based on many different cues such as luminance, color, texture, etc. In case of the tiger in the jungle, texture plays a critical role. What are the visual processes that enable perceptual agents to separate figure from ground using texture cues? This intriguing question led many researchers in vision to investigate the mechanisms of texture perception.

Beck [2][3] and Julesz [4] conducted psychological experiments investigating the features that enable humans to discriminate one texture from another. These studies suggested that texture segmentation occurs based on the distribution of simple properties of "texture elements", such as brightness, color, size, and the orientation of contours, or other elemental descriptors [5]. Julesz also proposed the texton theory, in which textures are discriminated if they differ in the density of simple, local textural features, called textons [6]. Most models based on these observations lead to a feature-based theory, in which segmentation occurs when feature differences (such as difference in orientation) exist. Furthermore, psychophysical and physiological studies have shown that human texture processing may be based on the detection of texture boundaries between heterogeneous textures using contextual influences via intra-cortical interactions in the primary visual cortex [7][8].

In the current studies of texture segmentation and boundary detection, texture is usually defined in 2D. However, an interesting hypothesis arises when we ask an important question regarding the nature of textures: *What are textures, and why did the ability to discriminate textures evolve or develop?* One possible answer to the question is that texture is that which defines physically distinct surfaces, belonging to different objects, and that texture segmentation function may have evolved out of the necessity to distinguish different surfaces. Human visual experience with textures can be, therefore, in most cases to use them as cues for surface perception, depth perception, and 3D structure perception. In fact, psychological experiments by Nakayama and He [9][10] showed that the visual system cannot ignore information regarding surface layout in texture discrimination and proposed that surface representation must actually precede perceptual functions such as texture perception (see the discussion section for more on this point).

From the discussion above, we can reasonably infer that texture processing may be closely related to surface discrimination. Surface discrimination is fundamentally a 3D task, and 3D cues such as stereopsis and motion parallax may provide unambiguous information about the surface. Thus, we can hypothesize that 3D surface perception could have contributed in the formation of early texture segmentation processing capabilities in human vision. In this paper, through computational experiments using artificial neural networks, we investigate the relative difficulty of learning to discriminate texture boundaries in 2D vs. 3D arrangements of texture. We will also study whether the learned ability to segment texture in 3D can transfer into 2D. In the following, we will first describe in detail the methods we used to prepare the 2D and 3D texture inputs (Section 2.1), and the procedure we followed to train multilayer perceptrons to discriminate texture boundaries (Section 2.2). Next, we will present our main results and interpretations (Section 3), followed by discussion (Section 4) and conclusion (Section 5).

2. Methods

To test our hypothesis proposed in the introduction, we need to conduct texture discrimination experiments with 2D and 3D arrangements of texture. In this section, we will describe in detail how we prepared the two different arrangements (Section 2.1), and explain how we trained two standard multi-layer perceptrons to discriminate these texture arrangements (Section 2.2). We trained two separate networks that are identical in architecture, one with input prepared in a 2D arrangement (we will refer to this network as the *2D-net*), and the other in a 3D arrangement (the *3D-net*).

2.1. Input preparation

We prepared three sets of texture stimuli S_1, S_2, and S_3. Textures in S_1 were simple artificial texture images (oriented bars of orientation $0, \frac{\pi}{4}, \frac{\pi}{2}$, or $\frac{3\pi}{4}$ at two different spatial frequencies); those in S_2 were more complex texture images such as crosses and circles, adapted from Krose [11] and Julesz [12]; and those in S_3 were real texture images from the widely used Brodatz texture collection [13] (Figure 1). For the training of the 2D-net and the 3D-net, the eight simple texture stimuli in S_1 were used. For testing the performance of the 2D-net and the 3D-net, all sets of texture stimuli (S_1, S_2 and S_3) were used.

To extract the primitive features in a given texture, we used Gabor filters. Previous results have shown that Gabor filters closely resemble experimentally measured receptive fields in the visual cortex [14] and they have been widely used to model the response of visual cortical neurons. A number of texture analysis studies also used oriented Gabor filters or difference of Gaussian (DOG) filters to extract local image features [15][16][17].

We used a bank of oriented Gabor filters to approximate the responses of simple cells in the primary visual cortex. The Gabor filter is defined as follows [18]:

$$G_{\theta,\phi,\sigma,\omega}(x,y) = \exp^{-\frac{x'^2+y'^2}{2\sigma^2}} \cos(2\pi\omega x' + \phi), \quad (1)$$

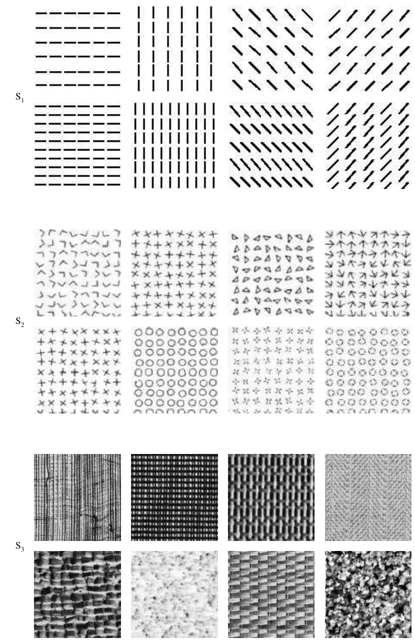


Figure 1. Texture stimuli. Three texture sets S_1, S_2, and S_3 are shown from the top to the bottom row.

where θ is the orientation, ϕ is the phase, σ is the standard deviation (width) of the envelope, ω is the spatial frequency, (x, y) represents the pixel location, and x' and y' are defined as:

$$x' = x\cos(\theta) + y\sin(\theta) \quad (2)$$

$$y' = -x\sin(\theta) + y\cos(\theta). \quad (3)$$

For simplicity, only four different orientations ($0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$) were used for θ. (Below, we will refer to $G_{\theta,\phi,\sigma,\omega}$ as simply G.) To adequately sample the spatial-frequency features of input stimuli, three frequency ranges (1 to 3 cycles/degree) were used for ω. The size of filter was 16×16, $\sigma = 16/3$, and $\phi = \pi/2$. This resulted in 12 filters G_i (for $i = 1..12$) for the computation of simple cell responses as shown in Figure 2. To get the Gabor response matrix C_i, a gray-level intensity matrix I was obtained from the images randomly selected from S_1 and convolved with the filter bank G_i:

$$C_i = I * G_i, \quad (4)$$

where $i = 1..12$ denotes the index of a filter in the filter bank, and $*$ represents the convolution operator. The Gabor filtering stage is linear, but models purely based on linear mechanisms are not able to reproduce experimental data [19]. Thus, half-wave rectification is commonly used to provide a nonlinear response characteristic following linear filtering. However, in our experiments, full-wave rectification was used as in [20], which is similar to half-wave rec-

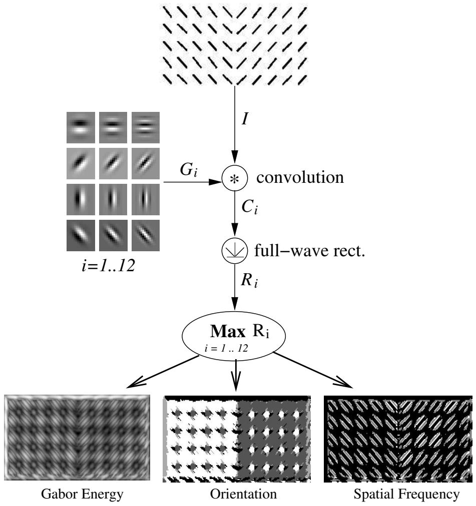
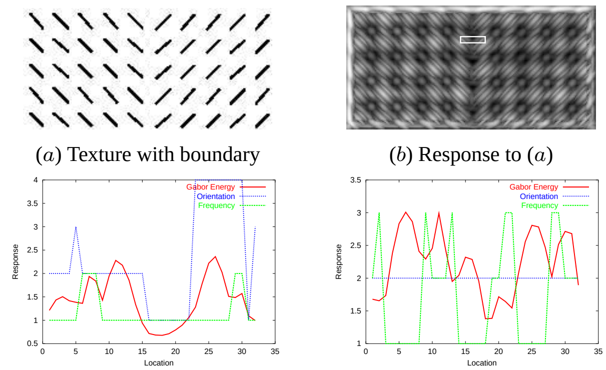


Figure 2. Gabor filter bank. The process used to generate two orientation response matrices is shown. The texture I is first convolved with the Gabor filters G_i (for $i = 1..12$), and the resulting responses are passed through a full-wave rectifier resulting in R_i. Finally, we get the Gabor energy matrix $E(x, y)$, Orientation response matrix $O(x, y)$, and Frequency response matrix $F(x, y)$.

tification, but is simpler to implement. Full-wave rectification is equivalent to summing the outputs of the two corresponding half-wave rectification channels (see, e.g. Bergen and Adelson [21] [19]). The final full-wave rectified Gabor feature response matrix is calculated as

$$R_i = |C_i|, \tag{5}$$

for $i = 1..12$. For each sample texture pair, we acquired three Gabor response matrices (or maps), which were the Gabor energy response matrix E, the orientation response matrix O, and the frequency response matrix F. First, to get the Gabor energy response matrix E, only one maximally responding values at location (x, y) from the twelve response matrices R_i were selected. In addition to the Gabor energy matrix, orientation response matrix and frequency response matrix were computed to avoid the loss of orientation and frequency properties at the given location. The orientation response matrix O had orientation index $(1 \leq O(x, y) \leq 4)$ of the filter that had maximum response at location (x, y) out of 12 filters. The frequency response matrix F had frequency index $(1 \leq F(x, y) \leq 3)$ of the filter that had maximum response at location (x, y) out of 12 filters. The same filtering procedure was used for both the 2D and the 3D arrangement of textures, which will be described below. Figure 2 shows the Gabor filter bank and the three response matrices E, O, and F of the given texture pair.



(a) Texture with boundary (b) Response to (a)

(c) Response profile of (b) (d) Response profile (no-boundary)

Figure 3. Generating the 2D input set (2D preprocessing). The procedure used to generate the 2D training data is shown. (a) Input with a texture boundary. (b) Orientation response calculated from (a). Only the E matrix is shown. (c) The response profile from the 32-pixel wide area marked with a white rectangle in (b). The three curves represent the profiles in the E, O, and F matrices. (d) A similarly calculated response profile in a different input texture, for an area without a texture boundary (note the periodic peaks).

To get the 2D training samples for the 2D-net, two randomly selected textures from S_1 were paired and convolved with the Gabor filter bank (figure 2). Gabor energy response matrix was acquired first, and orientation response matrix and frequency response matrix were computed from the 12 different response matrices that were used to get the Gabor energy response matrix. Each training input in the 2D training set consisted of three 32-element vectors (say, ξ_k^{2D}, where k is the training sample index) taken from a short horizontal strip (response profile) of the Gabor response matrix, the orientation response matrix, and the frequency response matrix, which resulted in a 96-element vector. A single scalar value (say, ζ_k^{2D}) indicating the existence $(= 1)$ or nonexistence $(= 0)$ of a texture boundary within that strip was paired with ξ_k^{2D}. The vector ξ_k^{2D} was taken from a horizontal strip centered at (x_c, y_c) within the Gabor energy matrix, the orientation response matrix, and the frequency response matrix respectively (e.g., the white rectangle in figure 3b), where x_c is the horizontal center where the two textures meet, and y_c is randomly chosen within the full height of the matrix. The Gabor energy matrix was normalized so that each value in the matrix has the range $0 \leq E(x, y) \leq 5$. When the two selected textures were the same, a texture boundary will not occur at the center; and if they were different, a texture boundary will occur. The number of input-target pair (ξ_k^{2D},ζ_k^{2D}) in each class, either boundary or no boundary, was balanced so that each class is equally represented. Figure 3c shows an example vector ξ_k^{2D} when there was a texture boundary, and figure 3d a case without a boundary.

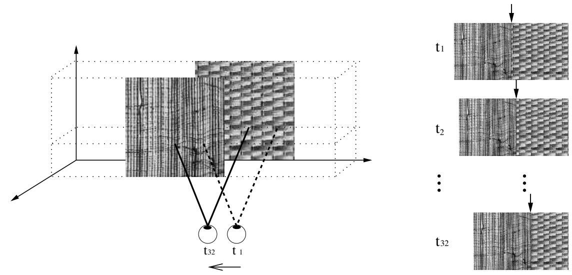
For the training samples for the 3D-net, motion cue was applied to simulate self-motion of an observer as shown in figure 4. One texture from a pair of textures was overlayed on top of the other and the texture above was allowed to slide over the one below, which resulted in successive further occlusion of the texture below. The texture above was moved by one pixel 32 times and each time the resulting 2D image (I'_j, for $j = t_1 ... t_{32}$; figure 5a) was convolved with the oriented Gabor filter bank followed by full-wave rectification as in the 2D preprocessing case (figure 5b). To generate a single training input-target pair (ξ_k^{3D}, ζ_k^{3D}) for the 3D-net, at each time step the Gabor energy response value $E(x_c, y_c)$, orientation response value $O(x_c, y_c)$ and frequency response value $F(x_c, y_c)$ were collected into a 92-element vector, where x_c was 16 pixels away to the right from the initial texture boundary in the middle, and y_c was selected randomly for each new input-target pair but remained the same within the same pair (the white square in figure 5b shows an example). Figure 5c shows an example of such a vector ξ_k^{3D} (note that the x-axis represents time, unlike in the 2D case where it is space) for a case containing a texture boundary, and figure 5d for a case without a boundary. The target value ζ_k^{3D} of the input-target pair (ξ_k^{3D}, ζ_k^{3D}) was set in a similar manner as in the 2D case, either to 0 (no boundary) or to 1 (boundary). When collecting the training samples for the 3D-net, the above procedure was performed with two different 3D configurations. In the first 3D configuration, the texture on the left side was on top of the texture on the right side with self-motion of observer from right to left. In the second configuration, the texture on the right was on top of the texture on the left side with self-motion of observer from left to right. For an unbiased training set, the same number of samples were collected for each 3D configuration.

For a fair comparison between the 2D and the 3D arrangements, 400 training samples were collected for each combination of two different textures to make 2,400 samples with target value of 1, and the same number of samples with target value of 0. This resulted in 4,800 input-target samples for each case ($1 \leq k \leq 4{,}800$). These 4,800 input-target samples from each training set were then randomly ordered during training.

2.2. Training the texture segmentation networks

We used standard multilayered perceptrons (MLPs) to perform texture boundary detection. The networks (2D-net and 3D-net), which consisted of two layers including 96 input units, 16 hidden units and 1 output unit, were trained for 2,000 epochs each using standard backpropagation[1]. The goal of this study was to compare the relative learnability of the 2D vs. the 3D texture arrangements, thus a backpropagation network was good enough for our purpose. The hyperbolic tangent function was used for the activation function of the hidden layer, which is defined as $f(v) = a \tanh(bv)$, where $a = 1.7159$ and $b = \frac{2}{3}$ respectively (following [22]). For the activation function of the output layer that consisted of one unit, radial basis function (RBF) was used. The use of the radial basis function in standard MLP is not common: It is usually used as an activation function of the hidden layer in radial basis networks, which has additional data-independent input to the output layer. In the experiment, as shown in the previous section, an input vector to MLP is symmetric about the center when there is no boundary. On the other hand, an input vector to MLP is quite asymmetric when there is a boundary, but the mirror image of that vector should result in the same class. This observation led us to use the radial basis function, which has a Gaussian profile. Several preliminary training trials showed that the use of the RBF as the activation function enabled both the 2D-net and the 3D-net to converge faster (data not shown here). For the training, the input vectors were drawn from the texture set S_1. Backpropagation with momentum and adaptive learning rate was applied to train the weights.

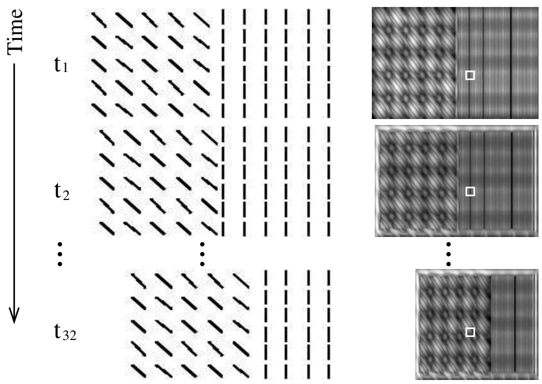
To determine the best learning parameters, several preliminary training runs were done with combinations of learning rate parameter $\eta \in \{0.01, 0.1, 0.5\}$ and momentum constant $\alpha \in \{0.0, 0.5, 0.9\}$. MLP with each combination was trained with the same set of inputs so that the results of



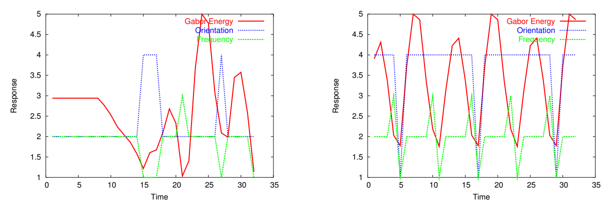
(a) Texture in 3D (b) Resulting 2D view

Figure 4. Generating the 3D input set (3D preprocessing). (a) A 3D configuration of textures and (b) the resulting 2D views before, during, and after the movement are shown. As the viewpoint is moved from the right to the left (t_1 to t_{32}) in 32 steps, the 2D texture boundaries in (b) (marked by black arrows) show a subtle variation.

[1]Matlab neural networks toolbox was used for the simulations.



(a) Input over time (b) Response to (a)



(c) Temporal profile of (b) (d) Temporal profile (no boundary)

Figure 5. Generating 3D input set through motion (3D preprocessing). (a) Texture pair images resulting from simulated motion: I'_j for each $j = t_1..t_{32}$. (b) The response matrix of the texture pair: R^{3D}_{ij}. (c) Response profile obtained over time near the boundary of two different texture images (marked by the small squares in c). (d) A similarly measured response profile collected over time, using a different input texture, near a location without a texture boundary (note the periodic peaks).

the experiment can be directly compared. Each training set consisted of 280 examples, drawn from S_1 and processed by the preprocessing procedure. The training process continued for 1,000 epochs. The MLPs with other combination of parameters failed to converge. Based on these preliminary training tests, we chose the learning parameters as follows: learning rate $\eta = 0.01$, and momentum constant $\alpha = 0.5$.

We also applied standard heuristics to speed up and stabilize the convergence of the networks. First, each input variable was preprocessed so that its mean value, averaged over the entire training set, is close to zero. Secondly, adaptive learning rate was applied. For each epoch, if the mean squared error (MSE) decreased toward the goal (10^{-4}), then the learning rate (η) was increased by the factor of η_{inc}:

$$\eta_n = \eta_{n-1}\eta_{\text{inc}}, \tag{6}$$

where n is the epoch. If MSE increased by more than 1.04, the learning rate was adjusted by the factor of η_{dec}:

$$\eta_n = \eta_{n-1}\eta_{\text{dec}}. \tag{7}$$

The constants selected above ($\eta = 0.01, \alpha = 0.5$) were used for the second test training to choose the optimal adaptive learning rate factors (η_{inc} and η_{dec}). Combinations of the factors $\eta_{\text{inc}} \in \{1.01, 1.05, 1.09\}$ and $\eta_{\text{dec}} \in \{0.5, 0.7, 0.9\}$ were used during the test training to observe their effects on convergence. The combination of factors $\eta_{\text{inc}} = 1.01$ and $\eta_{\text{dec}} = 0.5$ were chosen based on these results.

The 2D-net and the 3D-net were trained 10 times each with parameters chosen from the preliminary training above ($\eta = 0.01$, $\alpha = 0.5$, $\eta_{\text{inc}} = 1.01$, and $\eta_{\text{dec}} = 0.5$). After the training of the two networks, the speed of convergence and the classification accuracy were compared. To test generalization and transfer potentials, test stimuli drawn from the texture sets S_1, S_2, and S_3 were preprocessed using both 2D- and 3D-preprocessing to obtain six sample input sets. These input samples were then presented to the 2D-net and the 3D-net to compare the performances of the two networks on these six sample input sets. The results from these experiments will be presented in the following section.

3. Experiments and Results

We compared the performance of the two trained networks (2D-net and 3D-net), and also compared the performance of the two networks over novel texture images that were not used in training the networks.

3.1. Speed of convergence and accuracy on the training set

Figure 6 shows the 3 best learning curves of each network out of 10 trials during the training. The learning processes continued for 2,000 epochs. After 2,000 epochs, the average mean squared error (MSE) of the 2D-net was 0.0742 and that of the 3D-net was 0.0073. For the 10 trials, the results were comparable each time (data not presented here). The fact that the final MSEs of the three curves for each network did not vary significantly as shown in figure 6 suggests that the number of epochs was adequate. A noticeable difference in the two learning curves is that there are significant fluctuations in the learning curves of the 2D-net, which often prevented convergence of the network. These results indicate that the 3D-net is easier to train than the 2D-net. In other words, texture arrangements represented in 3D may be easier to segment than those in 2D. The misclassification rate, which was computed by using a threshold of 0.5 on the output response, in the 2D-net for the 2D training set was 11.2% and that of the 3D-net for the 3D training set was 0.2%, thus, accuracy was also higher in the 3D-net for the training data.

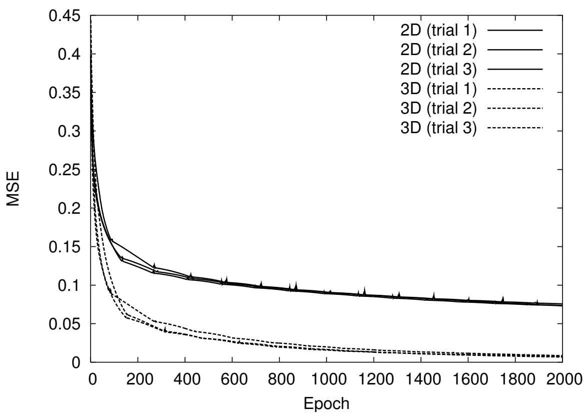


Figure 6. Learning curves of the networks. The learning curves of the 2D-net and the 3D-net up to 2,000 epochs of training on texture set S_1 are shown. The 3D-net is more accurate and converges faster than the 2D-net (near 100 epochs), suggesting that the 3D preprocessed training set may be easier to learn than the 2D set.

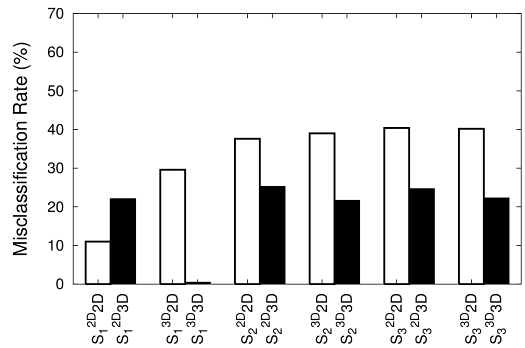


Figure 7. Comparison of misclassification rates. The misclassification rates of the different test conditions are shown (white bars represent the 2D-net, and the black bars the 3D-net). The x-axis label $\mathrm{S}_i^{n\mathrm{D}}m\mathrm{D}$ indicates that input set i preprocessed in n-D was used as the test input, and the m-D network was used to measure the performance. In all cases, the 3D-net shows a lower misclassification rate compared to that of the 2D-net, except for $\mathrm{S}_1^{2\mathrm{D}}2\mathrm{D}$.

3.2. Generalization and transfer

The 2D-net and the 3D-net trained with the texture set S_1 were tested on texture pairs from S_1, S_2 and S_3. (Note that for the texture set S_1, input vectors different from those in the training set were used.) For this testing, 500-sample sets of 2D and 500-sample of 3D per each texture set, which were prepared in the same manner as the training samples sets, were used. All six sample sets were presented to the 2D-net and the 3D-net. Two methods to compare the performance of the networks were used. First, we compared the misclassification rate, which is the percentage of misclassification. Misclassification rates were calculated for all 12 cases (= 6 sample sets $\times$ 2 networks): Figure 7 shows the result. The 3D-net outperformed the 2D-net in all cases, except for the sample set from S_1 with 2D preprocessing, which was similar to those used for training the 2D-net. It is also notable that the 3D-net outperformed the 2D-net on the sample sets from S_2 and S_3 prepared with 2D preprocessing (third and the fifth pair in figure 7; these are basically a 2D texture segmentation problem), where one would normally expect the 2D-net to perform better because of the manner in which the input was prepared.

As another measure of performance, we compared the absolute error (= $|target - output|$) for each test case for the two networks. The results are shown in figure 8. The plot shows the mean absolute errors and their 99% confidence intervals. The results are comparable to the misclassification rate results reported above. The 3D-net consistently outperformed the 2D-net for the sample sets from S_2 and S_3, and the differences were found to be statistically significant (t-test: $n = 500, p << 0.001$). However, the 2D-net outperformed the 3D-net for sample set from S_1 (figure 8 first pair from the left). Again, since S_1 preprocessed in 2D was used for training the 2D-net, this was expected from the beginning.

4. Discussion

Since the early works of Julesz [4] and Beck [2] on texture perception, many studies have been conducted to understand the mechanisms of the human visual system underlying texture segmentation and boundary detection in both psychophysical research and in pattern recognition research. In most cases their main concerns have been about the texture perception ability of human in 2D. The work presented in this paper suggests an alternative approach to the problem of texture perception, with a focus on boundary detection. First, we demonstrated that texture boundary detection in 3D is easier than in 2D. We also showed that the learned ability to find texture boundary in 3D can easily be transferred to texture boundary detection in 2D. Based on these results, our careful observation is that the outstanding ability of 2D texture boundary detection of the human visual system may have been derived from an analogous ability in 3D.

Our preliminary results allow us to challenge one common belief that many other texture boundary detection stud-

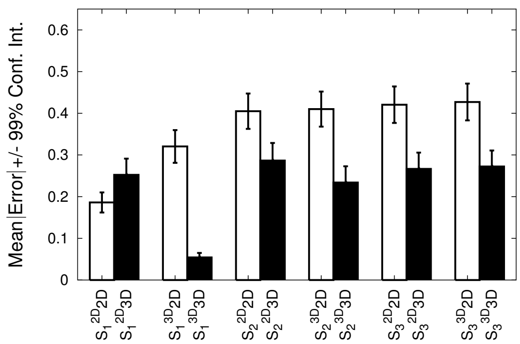
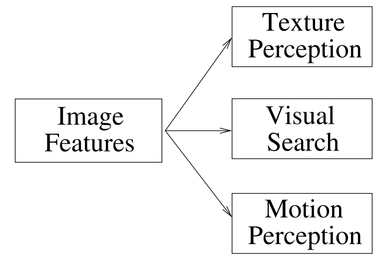
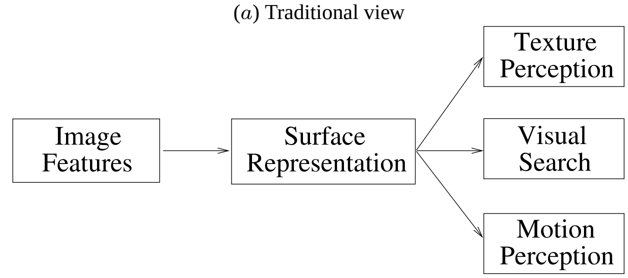


Figure 8. Comparison of output errors. The mean error in the output vs. the target value in each trial and its 99% confidence interval (error bars) are shown for all test cases (white bars represent the 2D-net, and the black bars the 3D-net). In all cases, the differences between the 3D-net and the 2D-net are significant (t-test: $n = 500, p << 0.001$). Note that for $S_1^{2D}, 2D < 3D$.



(a) Traditional view



(b) An alternative view

Figure 9. Two views of intermediate visual processing. (a) In the traditional view, texture perception, visual search, motion perception depend on feature processing in early cortical areas. (b) In an alternative view, surface representation must precede intermediate visual tasks [10]. (Adapted from [10].)

ies share. In this view, intermediate visual processing such as texture perception, visual search and motion process do not require object (in our context, "3D") knowledge, and thus perform rapidly; and texture perception is understood in terms of features and filtering, so the performance is determined by differences in the response profiles of receptive fields in low-level visual processing. A similar point as ours was advanced by Nakayama and his colleagues [9][10]. In Nakayama's alternative view on intermediate visual processing, visual surface representation is necessary before other visual tasks such as texture perception, visual search, and motion perception can be accomplished (figure 9). Such an observation is in line with our results indicating that 3D performance can easily transfer into a 2D task. (Note that there is yet another possibility, where all of these visual tasks are processed concurrently at the same stage, but we do not have enough evidence to either accept or reject such a proposal.)

The main goal of our work was to understand the nature of textures, and from that emerged the importance of 3D cues in understanding the texture detection mechanism in human visual processing. To emulate 3D depth, we employed motion cues to provide depth. This imposes potential limitations on our work, which is that additional information in 3D input may have become available to the 3D-net–some form of temporal information that that 2D inputs do not have. This can be seen as an unfair advantage for the 3D-net, but on the other hand, the 2D-net had additional spatial information which the 3D-net did not have, so eventually these two relative advantages may have canceled out. One way of addressing this issue may be to normalize (or equalize) the information content in the 2D vs. the 3D input preparation, which may allow us to more fairly assess the differences between the two modes of texture processing. Finally, another potential criticism may be that we only used S_1 for training. Would a contradictory result emerge if S_2 or S_3 was used to train the networks? We are currently investigating this issue as well, but we believe our main conclusion in this paper will hold even in different training scenarios.

5. Conclusion

We began with the simple question regarding the nature of textures. The tentative answer was that textures naturally define distinct physical surfaces, and thus the ability to segment texture in 2D may have grown out of the ability to distinguish surfaces in 3D. To test our insight, we compared texture boundary detection performance of two neural networks trained on textures arranged in 2D and in 3D. Our results revealed that texture boundary detection in 3D is easier to learn than in 2D, and that the network trained in 3D easily solved the 2D problem as well, but not the other way around. Based on these results, we carefully conclude that the human ability to segment texture in 2D may have originated from a module evolved to handle 3D tasks. One immediate future direction is to extend our current approach to utilize stereo cues as well as monocular cues used in this paper.

Acknowledgments

The authors wish to thank Ricardo Gutierrez-Osuna, Takashi Yamauchi, and three anonymous reviewers for their valuable comments; and Jyh-Charn Liu for his support. This research was supported in part by the Texas Higher Education Coordinating Board grant ATP#000512-0217-2001 and the National Institute of Mental Health Human Brain Project grant #1R01-MH66991.

References

[1] M. Tuceryan, *The Handbook of Pattern Recognition and Computer Vision*, 2nd ed. World Scientific Publishing Co., 1998, ch. 2.1 Texture Analysis, pp. 207–248.

[2] J. Beck, “Effect of orientation and of shape similarity on grouping,” *Perception and Psychophysics*, pp. 300–302, 1966.

[3] ——, “Textural segmentation, second-order statistics and textural elements,” *Biological Cybernetics*, vol. 48, pp. 125–130, 1983.

[4] B. Julesz, “Texture and visual perception,” *Scientific American*, vol. 212, pp. 38–48, Feb 1965.

[5] M. S. Landy and N. Graham, *The Visual Neurosciences*. MIT Press, 2004, ch. Visual perception of texture, pp. 1106–1118.

[6] B. Julesz and J. Bergen, “Texton theory of preattentive vision and texture perception,” *Journal of the Optical Society of America*, vol. 72, 1982.

[7] Thielscher and Neumann, “Neural mechanisms of cortico-cortical interaction in texture boundary detection: a modeling approach,” *Neuroscience.*, vol. 122, no. 4, pp. 921–939, 2003.

[8] Z. Li, “Pre-attentive segmentation in the primary visual cortex,” *Spatial Vision*, vol. 13, no. 1, pp. 25–50, 2000.

[9] Z. J. He and K. Nakayama, “Perceiving textures: beyond filtering,” *Vision Research*, vol. 34(2), pp. 151–62, 1994.

[10] K. Nakayama, Z. J. He, and S. Shimojo, *Visual Cognition*, 2nd ed. MIT Press, 1995, ch. Visual Surface Representation: A Critical Link between Lower-level and Higher-level vision, pp. 1–70.

[11] B. J. Krose, “A description of visual structure,” Ph.D. dissertation, Delft University of Technology, Delft, Netherlands, 1986.

[12] B. Julesz, “Textons, the elements of texture perception, and their interactions,” *Nature*, vol. 290, pp. 91–97, 1981.

[13] P. Brodatz, *Textures: A Photographic Album for Artists and Designer*. New York: Dover Publication, 1966.

[14] J. P. Jones and L. A. Palmer, “An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex,” *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1223–1258, 1987.

[15] I. Fogel and D. Sagi, “Gabor filters as texture discriminator,” *Biological Cybernetics*, vol. 61, pp. 102–113, 1989.

[16] A. Bovik, M. Clark, and W. Geisler, “Multichannel texture analysis using localized spatial filters,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55–73, January 1990.

[17] H.-C. Lee and Y. Choe, “Detecting salient contours using orientation energy distribution,” in *Proceedings of the International Joint Conference on Neural Networks*. IEEE, 2003, pp. 206–211.

[18] J. Daugman, “Two-dimensional spectral analysis of cortical receptive field profiles,” *Vision Research*, vol. 20, pp. 847–856, 1980.

[19] J. Malik and P. Perona, “Preattentive texture discrimination with early vision mechanisms,” *Journal of Optical Society of America A*, pp. 923–932, 1990.

[20] J. R. Bergen and E. H. Adelson, “Early vision and texture perception,” *Nature*, vol. 333, pp. 363–364, May 1988.

[21] J. Bergen and M. Landy, “Computational modeling of visual texture segregation,” 1991.

[22] Y. LeCun, “Efficient learning and second-order methods,” in *A tutorial in Neural Information Processing Systems*, Denver, 1993, pp. 76–77.

Development of emotional facial processing: Event-related brain potentials to happy and angry facial expressions in 7-month-old infants and adults

Tobias Grossmann, Tricia Striano

Facial expressions are an important way to communicate emotions. Over the last decades, a variety of behavioral studies have shown that human infants are adept at discriminating a range of facial expressions. By 3 months of age, infants discriminate happy from angry expressions. Furthermore, there is evidence showing that infants are sensitive to the affective tone these expressions convey. By 4 to 6 months of age, infants display more positive expressions in response to happy faces and more negative expressions in response to angry faces. While behavioral evidence points to an early ability to discriminate and recognize happy and angry facial expressions, little is known about neural processing of facial expressions as revealed by event-related brain potentials and how this processing might develop over the course of ontogeny.

Event-related brain potentials were measured in 20 adults and 20 7-month-old infants to assess neural processing of angry and happy facial expressions. ERPs were recorded while participants viewed photographs of a woman wearing either a happy or an angry facial expression.

Adults' ERPs when shown angry faces were more negative than ERPs elicited by happy faces as early as 250 ms after stimulus onset. This difference between the emotions reached its peak amplitude around 350 ms, and was statistically significant in the 300-400 ms latency interval ($F(1,19) = 20.92$, $p < 0.0002$). In contrast to the findings with adults, infants' ERPs for happy faces were more negative than ERPs elicited by angry faces as early as 350 ms after stimulus onset. This difference between the emotions reached its peak amplitude around 450 ms, and was statistically significant in the 400-500 ms latency interval ($F(1,19) = 5.59$, $p < 0.03$).

In adults, angry faces elicited a more negative amplitude than happy faces. This replicates previous findings in adults and in children. Interestingly we found a negativity in infants' ERPs distinguishing happy from angry faces. This indicates that, even by 7 months of age, there are measurable differences in the electric brain responses corroborating with the ability to discriminate happy and angry emotions behaviorally. The negative component invoked by both expressions is similar to a negative component observed in previous infant studies, thought to reflect allocation of attention. In contrast to adults, the amplitude of the negative component was larger for happy than for the angry face. This suggests that 7-month-olds allocated more attentional resources to the happy than to the angry face. This interpretation is consistent with previous findings showing that infants 4 and 6 months of age look longer at happy than angry expressions. The finding that, in contrast to adults, infants at 7 months of age do not allocate more attentional resources to the angry face, suggests that experience with facial expressions may play a key role in the development of specialised cortical mechanisms responsible for processing emotional information conveyed by the face. However, it remains to be specified when during development the human brain begins to allocate more attention to angry as compared to happy facial expressions.

An explanation of complex cell development by information separation

Akira Date
National Institute of Information and Communications Technology
Kyoto 619-0289 Japan
date@nict.go.jp

Koji Kurata
University of the Ryukyus
Faculty of Engineering
Okinawa 903-0213 Japan
kurata@mibai.tec.u-ryukyu.ac.jp

Complex cells in the primary visual cortex exhibit approximate invariance to position within a limited range. Hubel and Wiesel (1962) assume that complex cells receive their major inputs from simple cells or simple-cell-like subunits selective for the same orientation in different positions[1]. Nagano & Kurata (1981) and Földiák (1991) explain the shift invariance property of complex cells by using a modified Hebbian learning in which the modification of the synaptic strength is proportional not to the pre- and post-synaptic activity, but to the presynaptic activity and a temporal average of the postsynaptic activity [2, 3].

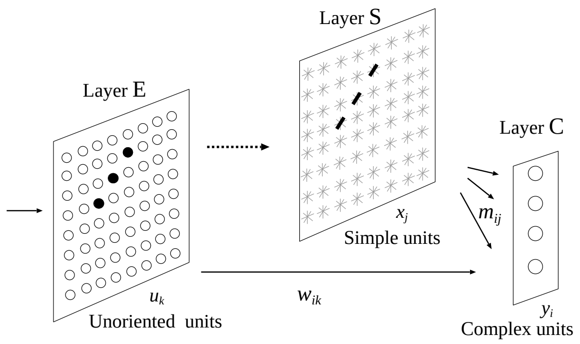


Figure 1. Architecture of our model. Units in the layer E have position selective. Our model is considered to be one in which layer E is added to the model of Földiák[3].

Although postsynaptic activity seems to be sustained for a period of time, there might be another mechanism by which the shift invariance property is obtained. Here, we propose a new computational model of complex cell development based on another possible mechanism. The model network(Fig.1) consists of three, E, S, and C, layers by which we model excitatory cells in LGN and/or V1, and simple cells, and complex cells in V1 respectively. Units in the layer E are assumed to be position-selective, and units in the layer S are line detectors for each specific location. During the learning phase, the network is exposed to randomly located short oriented bars, and units in the layer C develop its selectivity. The units in the layer C receive and learn inputs from the layer S through Hebbian or SOM (Self-organizing Map[4]) type connections, while anti-Hebbian connections from the layer E to C are assumed to force the layer C to represent aspect of the inputs uncorrelated to that represented on the layer E. We demonstrate that units in the layer C learns invariance to shift in input position.

Complex cell development might be explained in terms of information separation. The input signal to complex units carries essentially three-dimensional information, that is the location(two-dimensional) and the orientation(one-dimensional) of the bar. In our model, the “E” units are position-selective, and the “S” units are both position- and orientation-selective. When an “E” unit and a “C” unit are activated simultaneously, anti-Hebbian connections between them becomes more inhibitory to discourage the simultaneous activation of these two units in the future, and their correlation is decreased.

This works was partly supported by Grant-in-Aid (14750329, 16700232) from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT).

References

[1] Hubel, D. H. and Wiesel, T. N. (1962) Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *Journal of Physiology* **60**: 106–154.

[2] Nagano, T. and Kurata, K. (1981) A self-organizing neural network model for the development of complex cells. *Biol. Cybern.* **40**(3):195–200.

[3] Földiák, P. (1991) Learning invariance from transformation sequences. *Neural Computation*, **3** (2):194–200.

[4] Kohonen, T (2001) Self-Organizing Maps, Third Ed., Springer Verlag, Berlin.

Information maximization in face processing

Marian Stewart Bartlett an Javier R. Movellan
Machine Perception Lab Institute for Neural Computation
University of California, San Diego
marni@salk.edu; movellan@mplab.ucsd.edu

Abstract

This presentation explores principles of unsupervised learning and how they relate to face recognition. Dependency coding and information maximization appear to be central principles in neural coding early in the visual system. We argue that these principles may be relevant to how we think about higher visual processes such as face recognition as well. We first review examples of dependency learning in biological vision, along with principle s of optimal information transfer and information maximization. Next, we examine algorithms for face recognition by computer from a perspective of information maximization. The eigenface approach learns first and second-order dependencies among face image pixels, and maximizes information transfer only in the case where the input distributions are Gaussian. Independent component analysis (ICA) learns high-order dependencies in addition to first and second-order relations, and maximizes information transfer for a more general set of input distributions. Face representations based on ICA gave better recognition performance than eigenfaces, supporting the theory that dependency learning is a good strategy for high level visual functions such as face recognition. Finally, we review perceptual studies suggesting that dependency learning is relevant to human face perception as well, and present an information maximization account of perceptual effects such as the atypicality bias, and face adaptation aftereffects.

Finding People by Contingency: An Infomax Controller Approach

Javier R. Movellan

Abstract

We frame social interaction as a problem in real-time system's identification and control. System's identification refers to the task of identifying the properties of a system we are trying to control. Control refers to the problem of sending an input sequence to a system in order to maximize expected returns with respect to desired goals. The brain faces a control problem when sending motor commands to the limbs, where it has to account for the inertial properties of the arm and the delays and levels of uncertainty in the efferent and afferent transmission lines. Riding a bicycle, shooting baskets, and playing a musical instrument are also control problems.

Social interaction can be approached from the point of view of stochastic control theory under conditions of random delays and uncertainty much larger than those found when controlling non-social instruments. We illustrate this framework on the problem of finding people via contingency. There is strong evidence that infants use contingency analysis to identify other humans (Watson 1972, 1979) and that in fact contingency information may be more powerful than morphological properties of human faces (like the presence of eyes). Indeed we found that by ten moths of age infants used contingency information in a very active manner, ascertaining in a matter of seconds, whether a robot was or was not responsive to them (Movellan and Watson, 1987, 2002).

We formalize the control problem in terms of a generative model in which observed behaviors can be produced by three different causes: (1) Self feedback (e.g., when we hear our own vocalization). (2) Responses from other humans that are related to our activity. (3) Background reponses unrelated to our activities. There are two possible control conditions: (1) A human is responding to us; and (2) Human are not responding to us.

Given the random delays and noise typical of social interaction, the goal of the controller is to generate a sequence of behaviors that gather as much information as possible in a minimum amount of time about which of the two conditions we are operating under. We call this an infomax controller.
Interestingly the controller exhibits patterns of behavior very similar to those found in 10 month infants when interacting with a robot system for the first time.

The controller is currently being implemented on RUBI, a social robot under development at our laboratory, and will be ready for demonstration at ICDL2004.

DEVELOPMENT OF FACE PROCESSING IN AUTISM: A LOOK INTO SPATIAL FREQUENCIES AND THE INVERSION EFFECT.

C. Deruelle & C. Rondan Institute of Physiologic and Cognitive Neurosciences, CNRS, Marseille, France & Laboratory of Psychology and Neurocognition, UMR 5105, CNRS, Grenoble, France

This research was aimed at exploring the development of abnormal face processing strategies in children and adults with autism (ASD). Subjects were asked to match faces (upright or upside down) or chairs filtered in high (HSF, local processing) and low spatial frequencies (LSF, holistic processing) to non-filtered images. Results show an evolution of face processing strategies with age. Indeed, adults with autism rather use LSF than HSF information in faces, just as controls, whereas children with autism exhibit the opposite pattern of preference, using rather HSF. Our findings demonstrate that the processing of social objects such as faces possibly meet the typical pattern of performance in adults and this may have implications on the development of social deficits observed in this pathology.

A Toy-like Robot in the Playroom for Children with Developmental Disorder

Hideki Kozima[1], Cocoro Nakagawa[1], Yuriko Yasuda[2], Daisuke Kosugi[3]

Fifteen children, with PDD, autism, or other developmental disorder, interacted with a toy-like robot, Keepon. This report describes our preliminary findings of how the children changed their ways of interaction over 5-month longitudinal observations.

1. Keepon, the Robot

Keepon is a small (12cm tall), soft (made of silicone rubber), simple (yellow snowman-like) robot (Figure 1) [1]. It has two video cameras and a microphone at the nose. Keepon can perform two types of motion: (1) expressing its *attention* by orienting its face to a certain target in the environment, and (2) expressing its *emotional states*, such as pleasure and excitement, by rocking its body from left to right and by bobbing up and down. Keepon is connected by wireless links to a remote PC, from which a human operator or a computer program controls the motion.

Beforehand we observed 23 normal infants interact with Keepon. Each infant together with his/her mother interacted with Keepon (Figure 2, left). We found from the observations that (1) one-year-olds recognized Keepon as an autonomous agent, responding to its attention and emotion by following its gaze and mimicking its gesture, and (2) two-year-olds, as a social being to play with by showing toys, asking questions, and stroking its head.

2. Interactions in the Playroom

We placed Keepon in the playroom, with a lot of toys to play with, at a day-care center for preschoolers (2 to 4 years old) with developmental disorders (Figure 2, right). Fifteen children, often with their parents and nursing staffs, could interact with Keepon spontaneously anytime during the group remedial session (about 3 hours). Through a series of longitudinal observations for about five months (12 to 15 sessions for each child), we observed various types of change in the children's ways to interact with Keepon. Here we exemplify two cases:

- A four-year-old boy with PDD: He showed strong interest from Session 1 (hereafter, S1). His touch to Keepon got stronger through S1 to S3. After S4, the interaction became gentle, as if Keepon was his exclusive pet, and became social gradually over the subsequent sessions — putting a cap on Keepon's head (caregiving), giving a toy cookie (pretense), mimicking Keepon's gesture, talking jargons, etc.
- A four-year-old girl with autism: She also showed interest from S1, but did not get close to Keepon. Through S1 to S7, she avoided being looked straight at by Keepon (gaze aversion); however she often looked into Keepon's profile; the distance to Keepon gradually got shorter. Her first touch was in S11, since then she started social interaction including eye-contact, putting a cap, talking jargons, etc.

3. Summary

We longitudinally observed the children's behavior from the view point of Keepon as their playmate. Their style of interaction suggests how they recognize Keepon — as a moving toy, a living creature, or a social partner. Although it is difficult to dissociate qualitative change in their social capability from quantitative increase in their familiarity with Keepon, this longitudinal observation gives us a lot of information beneficial to the remedial service (e.g. facilitating interpersonal communication) as well as the psychological research (e.g. modeling cognitive development).

References

[1] Hideki Kozima, Cocoro Nakagawa, and Hiroyuki Yano (2004). "Can a robot empathize with people?", *Artificial Life and Robotics*, Vol. 8 (to appear).

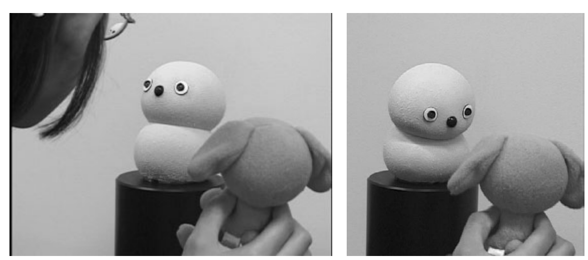


Figure 1. Keepon gazing at a human face and then at a toy.

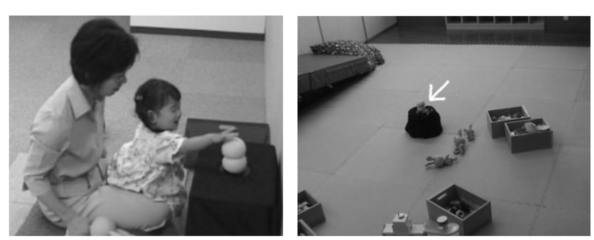


Figure 2. Keepon interacting with a 27 m.o. girl (left), waiting for children in the playroom (right).

[1]National Institute of Information and Communications Technology (NICT), Kyoto, Japan, {xkozima, cocoron}@nict.go.jp

[2]Omihachiman City Day-Care Center for Children with Special Needs (*Hikari-no Ko*), Shiga, Japan, 010810@city.omihachiman.lg.jp

[3]Kyoto University *and* JSPS Fellow, Kyoto, Japan, dkosugi@bun.kyoto-u.ac.jp

Comparing emotional expressions using eyes or mouths: a perceptual advantage in autism?

A W Hendriks[1], P J Benson[2], M Jonkers[1], S Rietberg[1] ([1] Department of Special Education, University of Nijmegen, Spinoza Building, PB 9104, NL 6500 HE Nijmegen, The Netherlands; [2] School of Psychology, College of Life Sciences and Medicine, King's College, University of Aberdeen, Aberdeen AB24 2UB, Scotland, UK)

While autistic individuals are able to process many types of visual information at a level commensurate with their age, studies have shown that they have difficulty interpreting facial expressions. One reason could be that autistics suffer from weak central coherence, i.e. a failure to integrate parts of information into globally coherent wholes [Frith, 1989 *Autism: Explaining the Enigma* (Oxford: Blackwell); Frith and Happé, 1994 *Cognition* 50 115 - 132]. To test this hypothesis we presented autistic and age-, sex- and IQ-matched normal children with pairs of facial images of the same sex but different identities. Their task was to decide whether the faces showed a similar expression (experiment 1), or whether either solely the eyes (experiment 2) or the mouths (experiment 3) displayed the same emotion while ignoring the rest of the face. The second stimulus in each pair was digitally altered in half of the trials so that the expression of the target feature was incongruent with the expression of the rest of the face, e.g. happy eyes within the context of an angry face. Although autistics were expected to show relatively greater difficulty comparing whole facial expressions, we proposed that they should be better than normal children when judging the similarity of single expressive features and that they should be less adversely affected by incongruent face contexts.

Overall accuracy did not differ between the autistic and control children, either when full facial displays were compared (exp 1) or the eyes (exp 2) or mouths (exp 3) alone were judged. However, reaction times in experiments 2 and 3 differed significantly: autistics were significantly faster than controls in judging the similarity of the emotional expression of eyes as well as mouths. This result indicates that autistics were better able to concentrate on a single feature within the faces; the finding also suggests that our autistic group was not perturbed by having to compare facial expressions (albeit an emotional expression of just single feature). Contrary to expectations, incongruent facial contexts were equally problematic for both the autistics and controls, causing increased error rates and response times.

Our results suggest we may have to reconsider the remit of weak central-coherence theory as an explanation of impoverished facial expression perception. We discuss this new finding in terms of theories of emotion-deficit disorders and current evidence on atypical visual-information processing in autism.

Cognitive Development in Context: Learning to Pay Attention

Petra Björne
petra.bjorne@lucs.lu.se

Christian Balkenius
christian.balkenius@lucs.lu.se

Lund University Cognitive Science
Kungshuset, Lundagård, SE-222 22 Sweden

Abstract

A developing system must be able to learn new things without forgetting what it has learned before. It should be capable of reacting in different ways to the same stimuli in different contexts. Context sensitive reinforcement learning, which parallels some of the functions of the basal ganglia, is a learning algorithm that fulfills this requirement when the context is explicitly given. Here, we extend the algorithm with the ability to identify the relevant features of the environment that defines the context. It is suggested that this is a critical component of an architecture for cognitive development and we present simulation results that illustrate the operation of the system.

1. Introduction

A developing system must be able to learn new things without forgetting what it has learned before. It should also be capable of reacting in new ways in new contexts. Ideally, what has already been learned should be generalized to new situations, while new learning should not interfere in a negative way with previously learned behavior that is still appropriate.

For this to be possible, it is essential that new and old contexts can be distinguished in an efficient way. Otherwise, it would not be possible for the system to know when behaviors should be modified to fit the new situation or simply forgotten and relearned.

However, in most cases, there is no individual stimulus in the environment that signals that the situation is new or different. This can usually only be determined by examining several cues in combination.

and even when a single stimulus indicates a new context, this stimulus has to be attended to influence learning.

Here, we want to develop a computational system that can automatically learn to attend to relevant aspects of the environment and use these to determine when they should be used as indications of a changed context. The system takes its starting point in behavioral data on the role of context in learning and relearning.

It is yet not possible to build a large scale model of the brain at a detailed physiological level. Too many details are simply unknown. Instead, we strive for a model that parallels the brain at a systems level where the different components of the model functionally correspond to different brain regions, but we do not attempt to model how these regions work at a neuronal level. Below, we focus on the basal ganglia and its role in the production of context dependent action and in the selection of contextual cues.

A task such as reaching for an object will involve several serially connected chains of specialized motor structures. At the same time, information will be processed in a parallel organization of multiple cortical, basal ganglia, thalamic and cerebellar structures (Salinas et al., 2000). Thus, a model of the functional role of the basal ganglia in a motor task must be consistent with a model of the functional role of other specialized cortical areas which the basal ganglia interact. This is not to say that at every stage of modeling we need to have a complete model of the brain, but it is necessary to be aware of the fact that no brain structure works in isolation from other structures. A model of any single structure must aim at integration with models of the other structures.

In the case of modeling the functions of the basal ganglia, besides accounting for their involvement in motor tasks, such a model should strive to be consistent with their involvement in non-motor tasks, such as sensory decision making, motivation, attentional and volitional modulation of other neural structures.

Further, a model of the basal ganglia would need to be consistent with data on impairments caused by degenerative illnesses afflicting the basal ganglia, such as Parkinson's disease and Huntington's chorea. A model of this kind would be able to provide valuable information on the interaction of cognitive and motor impairments of these patients.

There is always the possibility that we will find a mismatch between a neurocognitive model and experimental data, human or animal. This, however, need not signal a drawback. Instead, a mismatch can guide further empirical research and help improve the model (Hyland, 2000). A system model of neurocognitive functioning, if based on neurophysiological and behavioral data, will provide a powerful instrument for analyzing experimental data and develop hypotheses for further research. Both the results from simulations and new experimental data will, in turn, improve the model.

In the following subsections we describe data and models that have something to say about how learning and attention interacts with context. A more detailed review can be found in Balkenius (2000).

1.1. Learning in context

In most models of learning, context does not play a part at all. In those models where context have a role, it is often in the form of a dedicated input. Although such a solution is a step forward, it neglects that in the real world stimuli and contexts are not labeled to indicate their role in the learning experiment.

Another problem with many learning models where context appears is that it acts in the same was as any additional stimuli. While this acknowledges that a distinction between stimulus and context is not always easy to make, it ignores that in many learning experiments, the roles of stimulus and context are very different. While initial learning appears to be mainly context insensitive, relearning makes behavior context dependent. This has been shown in an important experiment by Bouton (1991) where learning in one context generalizes to other contexts, but extinguished behavior reappears outside the extinction context (see also Balkenius and Morén, 2000). This learning strategy is very powerful as it maximizes the generalization of learned behavior between contexts, while still being able to differentiate behavior in different contexts when needed. These results imply that there is a need to distinguish between stimulus and context.

1.2. Context and attention

We have earlier proposed that a context code can be constructed from a sequence of attentional fixations (Balkenius, 2000). Balkenius and Morén (2000) describe a computational model that can automatically generate context codes from sequences of attentional fixations of features in the environment. The model binds each environmental feature to its location before all features are combined into a context code. The selection of features was controlled by a fixed mechanism that would scan the features of the environment in a sequential manner.

In many cases, it would be useful if the selection of the features that make up the context could be put under reinforcement control. This would potentially allow the system to select the critical features that define the context or task to be accomplished. The initial steps towards such a mechanism were described by Balkenius (2000), where it was suggested that attentional shifts should be considered as any other action and learned in the same way. This principle was called *attention-as-action*.

An important consequence of this principle is that learned attention shifts will become context dependent in the same way as other actions. Since attention controls which stimuli are treated as parts of the context, this will make the contexts codes themselves context dependent.

1.3. The basal ganglia

Traditionally, the basal ganglia have been considered to be important for voluntary control and planning of body movements (Middleton and Strick 1994; Hikosaka, Takikawa and Kawagoe, 2000). However, through studies of persons with impairments of the basal ganglia, such as Parkinson's disease and Huntington's chorea, increasing insight into the cognitive functions of the basal ganglia has emerged. Along with the above mentioned neurodegenerative disorders, research into neurodevelopmental disorders, such as ADHD, autism and obsessive compulsive disorders (Bradshaw 2001) has further highlighted the importance of the basal ganglia in higher cognitive functions.

The basal ganglia operate by exerting tonic inhibition with phasic disinhibition (Kimura 1995; Hikosaka, Takikawa and Kawagoe, 2000), i.e., they select appropriate behaviors rather than controlling their detailed execution. This is probably true for both motor and non-motor functions controlled by the basal ganglia. An example would be the orienting response, which requires integration of information from several sensory modalities. From this integrated information an appropriate signal is selected, probably by processes in the basal ganglia (Redgrave, Prescott and Gurney 1999; Hikosaka, Takikawa and Kawagoe, 2000). The actual motor response is controlled by the superior colliculus (SC), which receives input from the frontal eye fields (FEF) and areas of the parietal cortex constituting the neural correlates for selection of saccades or attention. The role of the basal ganglia is to inhibit the SC (Hikosaka, Takikawa and Kawagoe, 2000). This is done though the substantia nigra (SN), which projects to the intermediate layer of the SC. The SN, in turn, is inhibited by the caudate nucleus. Occasionally the SN releases the inhibition of the SC, which results in a saccade to the contralateral side. Here, the basal ganglia select to produce the response, but the specific target of the orientation is controlled by the cortical input to the superior colliculus and not by the basal ganglia.

The output neurons of the SN or GPi show very high spontaneous activities. In contrast, the projection neurons of the striatum become active only when the animal performs an appropriate task, and whereas the neurons of the putamen can be activated by simple motor tasks, complex behavioral tasks are needed to activate the caudate. This suggests that the neurons in this region are sensitive to the behavioral context in which an action should or should not be selected.

The responses of the neurons in the caudate resemble the responses of those in the SN but with opposite signs, changing their activity when the location of the stimuli must be remembered or attended, or when the saccade uses the working memory.

The function of the basal ganglia has been linked to behavioral learning that is sensitive to reward. The responses of the dopamine cells appear to code for the temporal difference error between the expected and actual reward received. Recently, the responses of dopaminergic cells in the basal ganglia have been shown to react in a context dependent way (Nakahara, et al., 2004).

1.4. Working memory

The basal ganglia are also involved in the manipulation of working memory. In patients with Parkinson's disease, the degeneration of dopaminergic neurons projecting to the basal ganglia leads to a difficulty in manipulating information that is stored in working memory (Lewis, Cools et al. 2003; Lewis, Dove et al. 2004). These patients seem to be able to maintain information over a short time span in a verbal memory task and then retrieve it in an unmodified version. However, they seem to have difficulties in manipulating the same information. According to the authors, this would correspond to the visuospatial tasks of executive functioning that are also considered to be particularly difficult for the PD patients, namely tasks that involve manipulation of spatial information. One way to investigate working memory is through continuous performance tests.

Continuous performance tests are frequently used in assessments of sustained attention (Lin 1999; Oades 2000). The CPT-AX is a continuous performance test that puts high demands on working memory. Frank et al. (2001) have developed an even more demanding version of the task (Fig. 1). In the original version, the subject is presented with a sequence of letters and is expected to respond to the letter X if the previous stimulus was an A. In the extended version, the subject has to respond to the X preceded by an A within a context defined by the number 1. If the number 2 instead defines the context, the subject has to respond to the letter Y if preceded by a B.

This calls for rapid updating of working memory, i.e., an incoming stimulus has to be encoded. Furthermore, the context 1 or 2 has to be maintained stably while interference occurs from processing of targets and distracters. Finally, the task calls for selective updating of working memory, where the context 1 or 2 remains stable, while the sequence of letters is continuously updated.

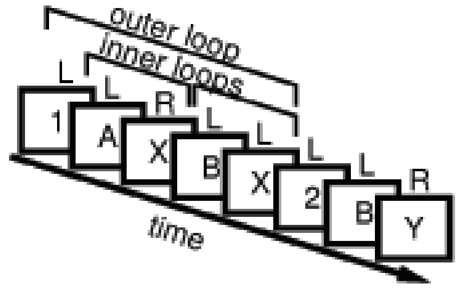


Figure 1. The 1-2-AX task.

The thalamus is tonically inhibited by the GPi/SNr and phasically disinhibited by the firing of striatal neurons. This functions as a gating mechanism, enabling but not causing other functions to occur, though as mentioned earlier, the context of the action is not defined by the disinhibition by the striatum. Frontal neurons react momentarily to irrelevant stimuli, returning to the task-relevant stimuli and maintaining these after the irrelevant stimuli have disappeared. This intrinsic maintenance is important for working memory and robust maintenance of task-relevant stimuli.

Disinhibition by striatal firing will modulate the intracellular switch of the frontal neurons, leading to an update of current and maintained information. Thus, according to Frank et al. (2001) stimuli will activate corresponding frontal representations and they will be maintained if they trigger the intrinsic maintenance switch. Those stimuli that do not have this intracellular switch activated will decay quickly, but will be maintained by recurrent excitation until other stimuli are presented. This latter function is important for learning what will be relevant to maintain.

Striatal neurons fire for a specific conjunction of environmental stimuli and internal context representations through descending projections from the cortex. Thus, striatal neurons would fire in response to the encoding of a frontal representation of the task 1-2-AX together with the incoming of some stimulus (1 or 2) and the encoding of the sequence of letters, enabling the response to the letters X and Y when appropriate.

2. Toward a Model

Taken together, the data presented above suggest that the basal ganglia is a central structure in the learning of context sensitive behavior dependent on reward contingencies whether the actions be external, such as orienting movements, or internal, such as manipulations of working memory. We now turn to a computational model that attempts to cover the central ideas described above. These include the need for a context code that can be adapted to the task and the ability to put attention under reinforcement control. The model is based on a context sensitive reinforcement learning algorithm and rests on the assumption that the basal ganglia perform a form of reinforcement learning (Schultz, Dayan and Montague, 1997, Doya,1999, 2000).

2.1. Context sensitive reinforcement learning

Like most on-line learning algorithms, the standard reinforcement learning algorithms are sensitive to catastrophic forgetting (cf. French, 1999). If it first learns one task and then another, the second learning experience is likely to interfere with the first. This is especially the case when a look-up table is used to store the value of each stimulus-response association.

For a developing system, it is essential that new tasks can be learned without erasing older ones. Balkenius and Winberg (2004) developed a novel context sensitive reinforcement learning algorithm, ContextQ, that overcomes this problem by using an additional input that codes for the context in addition to the input that codes for the current stimulus or state. The algorithm is an extension of the popular Q-learning algorithm (Watkins and Dayan, 1992) and uses a function approximator to estimate the function,

$$Q(c, s, a),$$

which assigns a value to each action a in state s and context c. The algorithm starts out with a zero value for all actions and as long as the received reinforcement is larger than predicted by the Q-function, learning increases the weight of a linear mapping between the stimulus and response. This learning is not influenced by the context and will allow the system to automatically generalize all learning to new contexts. This part of the algorithm is identical to what Sutton and Barto (1998) called LAQ. The difference compared the LAQ algorithm occurs during extinction, i. e. when the received reinforcement is lower than expected. In this case, the linear associator is unaffacted. Instead, a shunting inhibition from the context to the active stimulus-response association increases. This will make behavior learned during extinction context sensitive. Optionally, the two modes of learning can be mixed such that both acquisition and extinction involve both the stimulus and the context but to a different extent.

A detailed description of ContextQ can be found in Balkenius and Winberg (2004), where it was shown how the algorithm could learn a number of cognitive experiments including task-switching, a version of the Wisconsin Card Sorting Test, and context sensitive categorization. In Balkenius and Björne (2004), it was applied to an attention switch task to model impairments of attention in autism.

ContextQ, like ordinary Q-learning, is an off-policy algorithm, which means that it does not need to follow its own policy during learning (Sutton and Barto, 1998). This is a very important property when the reinforcement learning system is used as a component in a larger architecture. Since many different subsystems can suggest actions, it is not always the output of the reinforcement learning system that is used, but it should still learn the consequences of such actions. For example, reflex actions can be triggered directly by external stimuli and the reinforcement learning system could listen-in to these associations and learn to produce them voluntarily.

Fig. 2 illustrates three ways to add context information in ContextQ module. The context can either come from a context buffer that stores a sequence of attended stimuli or it can come from a working memory where stimuli decays over time. Finally, it is possible that the ContextQ module controls the context itself through an action that explicitly makes a stimulus part of the context.

2.2. Context buffer

The simplest way to construct a context is to select every stimulus and store them in a context buffer in such a way that the last few stimuli are always stored. Assume that the system encounters a stimulus sequence A, B, B. The task is to learn that a B preceded by an A is to be ignored, while when preceded by another B, the B should evoke a response. Using only A and B as inputs, a reinforcement system would behave as if responses to B were rewarded half of the time. With ContextQ, the correct behavior would be learned if the foregoing stimulus is used as context. In this case the reaction to B would be extinguished in the context of A. This would correspond to a context buffer of length 1.

This type of context processing results when the output of the context system is used directly as context input to the ContextQ module as shown in figure Fig. 2. The context buffer can be implemented as a tapped delay-line (Balkenius and Morén, 1999), i. e. a sequence of storage units that remembers the stimulus at the previous time steps.

2.3. Working memory decay

It is also possible to use a strategy where all stimuli gradually decay in a working memory. In this case, all stimuli are stored to some extent depending on how long ago they were encountered. To allow the ContextQ module to control working memory, it could influence the decay rate to make stimuli that contribute to the selection of the correct response last longer in working memory.

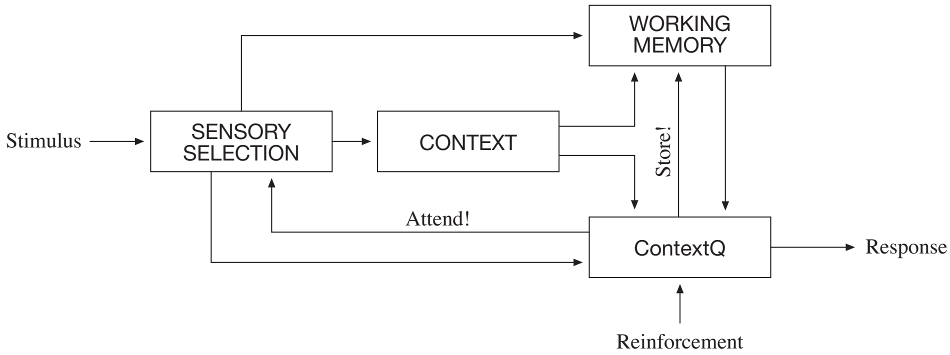


Figure 2. Three possible ways to put attention under reinforcement control. An action "Attend!" can be added that makes the system attend to its current input. Alternatively, storage in working memory can be controlled by temporal difference error to store stimuli that correlates with an error.

Let the input stimuli or context be coded in $I = \langle I_0, I_1, \ldots, I_i \rangle$. The working memory is coded by i memory nodes $m = \langle m_0, m_1, \ldots, m_i \rangle$. To each memory node is assigned a bias b_i which codes how much each stimulus or context input should be attended. The memory nodes are updated according to the following equation,

$$m_i(t+1) = (\alpha + b_i)m_i(t) + I_i$$

and the bias is updated according to the temporal difference error in the ContextQ algorithm as,

$$b_i(t+1) = b_i(t) - \beta \Delta Q(t) m_i(t),$$

where b_i is constrained to stay within the range $[0 \ldots 1]$, and

$$\Delta Q_t = \left[r_{t+1} + \gamma \max_a Q(c_{t+1}, s_{t+1}, a_{t+1}) - Q(c_t, s_t, a_t) \right],$$

i. e., the temporal difference error. This corresponds to the signal ”Store!” in Fig. 2.

The bias b_i is a measure of how useful stimulus I_i is and will control how long it is stored in working memory. The result of these equations is that the bias of any stimulus that is positively correlated with a negative reward prediction error will increase, thus allowing such stimuli to be stored in working memory. In the example above where A predicts that a response to B will not be rewarded, this would lead to the storage of A in working memory.

The main difference between this method and the use of a fixed context buffer is that the time interval that can be bridged is not set by the buffer length. Fig. 2 shows two possible inputs to a working memory module of this kind. The input to working memory either comes directly from the stimulus selection system or from the context module. In the first case, it is individual stimuli that are stored in working memory. In the second case, whole contexts are stored.

2.4. Attention as action

Another possibility is to allow the ContextQ module to select stimuli as part of the context by explicitly attending to them. In this case, an additional action is included in the system called ”Attend!” that makes a stimulus a part of the context. This action could be learned in the same way as other actions (Balkenius, 2000).

3. Simulations

To test the ability of the model to learn to pay attention to stimuli that would change the reward contingencies of actions, we ran three simulations with different conditions. In the first simulation, the system was required to learn to respond to an B except when directly preceded by an A. As described above, the previous stimulus was used as context. As expected, the model quickly learned this task and began to distinguish between the two contexts. This simulation was a version of one described in Winberg (2004).

In the second simulation, different distractor stimuli where placed between the A and B. In this case, it is necessary to keep the A in memory even when the distractor appeared in the input. We first tested this task with a tapped delay-line as context. Although the model could easily learn this task, it is sensitive to the exact timing of the stimulus A and B. If the inter-stimulus interval changes, it has to relearn the task again. Another drawback of using a tapped delay-line is that in a more realistic situation, the amount of stimulus data to store becomes intractable.

Fig. 3 shows the result of the final simulation, which included the working memory mechanism. The task was as above to respond to B except when proceeded by A. In this simulation, however, there were two presentation of the distractor stimulus X in between each A or B. As can be seen in Fig. 3A, the bias for the three stimuli changes over time. The bias for A increases while the biases for B and X decreases. As a result, A till be stored in working memory while B and X will not. This is shown in Fig. 3C, where the activation of the different stimuli in working memory are shown during a few presentation after the system has learned the task. The activation of A when the system should respond to B is lower than the activation when the response should be inhibited. This makes it possible for the ContextQ module to learn to discriminate between the two presentations of B. The graph also illustrates that stimulus B and X immediately drops out of working memory and are essentially not stored at all.

4. Discussion

We have extended an earlier model of context sensitive reinforcement learning with the ability to control attention and working memory. Preliminary simulations shows that the extended architecture is able to used its control of attention to explicitly store the appropriate stimuli in working memory and use them as contextual cues.

The main component of the system was inspired by the function of the basal ganglia in working memory and one future goal is to bring the model closer to the actual physiology of the basal ganglia. Another goal is to apply the model to developmental disorders (cf. Balkenius and Björne, 2001, 2004).

References

Balkenius, C. (2000). Attention, habituation and conditioning: Toward a computational model, *Cognitive Science Quarterly*, 1, 2, 171-214.

Balkenius, C., and Björne, P. (2001). Toward a robot model of attention-deficit hyperactivity disorder (ADHD). In Balkenius, C., Zlatev, J., Kozima, H., Dautenhahn, K., and Breazeal, C. (Eds.), *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Studies, 85.

Balkenius, C., and Björne, P. (2004). First steps toward a computational theory of autism. In Berthouze, L. et al. (Eds.), *Proceedings of the Fourth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Studies.

Balkenius, C. and Hulth, N. (1999). Attention as selection-for-action: a scheme for active perception. In Schweitzer, G., Burgard, W., Nehmzow, U., and Vestli,

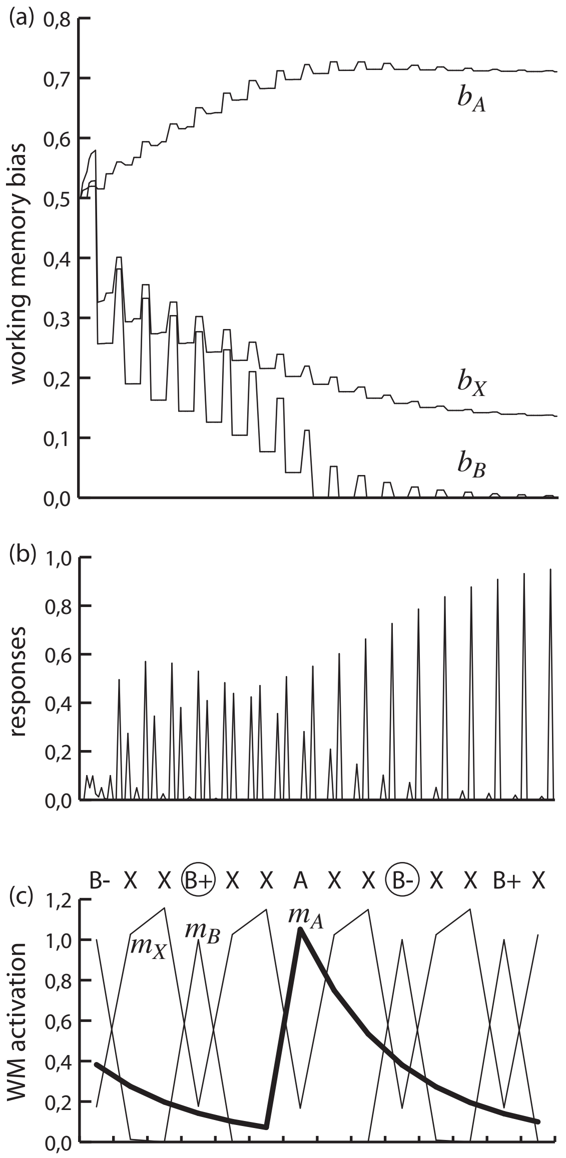


Figure 3. (a) Development of the level of attention (b_A, b_B, b_X) assigned to stimulus A, B and X. (b) Development of the response to AXXBXXBXX sequences during the simulation. Initially the model responds to every B. Later in the simulation, the response to B after an A is suppressed. (c) The working-memory activation (m_A, m_B, m_X) when the model has learned to remember A. The activation of A (thick-line) when B is presented is smaller when the response should be produced (circled B+) than when it should be inhibited (circled B-).

S. J. (Eds.), *Proceedings of EUROBOT '99* (pp. 113-119). IEEE Press.

Balkenius, C., and Morén, J. (1999). Dynamics of a classical conditioning model. *Autonomous Robots, 7*, 41–56.

Balkenius, C., and Morén, J. (2000). A computational model of context processing, In Meyer, J-A., Berthoz, A., Floreano, D., Roitblat, H. L., Wilson, S. W. (Eds.) *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior*, Cambridge, MA: MIT Press.

Balkenius, C., and Winberg, S. (2004). Cognitive modeling with context sensitive reinforcement learning, In Malec, J. et al. (Eds.), *Proceedings of AILS 2004*.

Bouton, M. E. (1991). Context and retrieval in extinction and in other examples of interference in simple associative learning. In L. W. Dachowski and C. F. Flaherty (Eds.), *Current topics in animal learning: Brain, emotion, and cognition* (pp. 25–53). Hillsdale, NJ: Erlbaum.

Bradshaw, J. L. (2001). Developmental disorders of the frontostriatal system: neuropsychological, neuropsychiatric and evolutionary perspectives. Hove: Psychology Press.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks, 12*, 961–974.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology, 10*, 732–739.

de Haan, M., Pascalis, O., and Johnson, M.H. (2002). Specialization of neural mechanisms underlying face recognition in human infants, *Journal of Cognitive Neuroscience, 14*, 199-209.

Frank, M. J., Loughry, B., and O'Reilly, R. C. (2001). Interacitons between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, & Behavioral Neuroscience, 1*, 137–160.

French, R. M. (1999). Catastrophic Forgetting in Connectionist Networks. *Trends in Cognitive Sciences, 3*, 128–135.

Gilbert, P. F. C. (2001). An outline of brain function. *Cognitive Brain Research, 12*, 61-74.

Hikosaka, O., Miyashita, K., Miyachi, S., Sakai, K., and Lu, X. (1998). Differential roles of the frontal cortex, basal ganglia, and cerebellum in visuomotor sequence learning. *Neurobiology of learning and memory, 70*, 137–149.

Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Reviews, 80*, 953–978.

Hyland, B. I. (2000). Relationships of substantia nigra dopamine neurone activity to behaviour. In Miller, R. and Wickens, J.R. (Eds.), *Brain dynamics and the striatal complex* (pp. 1–27). Amsterdam: Harwood Academic Publishers.

Johnson, M.H. (1998). Developing an attentive brain, In Parasuraman, R. (Ed.), *The attentive brain* (pp. 428-443), Cambridge, MA: MIT Press.

Kimura, M. (1995). Role of basal ganglia in behavioral learning. *Neuroscience Research, 22*, 353–358.

Lewis, S. J., Cools, R., Robbins, T. W., Dove, A., Barker, R. A., and Owen, A. M. (2003). Using executive heterogeneity to explore the nature of working memory deficits in Parkinson's disease. *Neuropsychologia, 41*, 645–654.

Lewis, S. J., Dove, A., Robbins, T. W., Barker, R. A., and Owen, A. M. (2004). Striatal contributions to working memory: a functional magnetic resonance imaging study in humans. *European Journal of Neuroscience, 19*, 755–760.

Lin, C. C. H., Hsiao, Chuhsing K., Chen, Wei J. (1999). Development of Sustained Attention Assessed Using the Continuous Performance Test among Children 6-15 Years of Age. *Journal of Abnormal Child Psychology, 27*, 403–412.

Middleton, F. A., and Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science, 266*, 458–461.

Miller, R. and Wickens, J.R. (Eds.) (2000). *Brain dynamics and the striatal complex* Amsterdam : Harwood Academic Publishers.

Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., Hikosaka, O. (2004). Dopamine Neurons Can Represent Context-Dependent Prediction Error. *Neuron 41*, 269–280.

Niebur, E. aand Koch, C. (1998). Computational architectures for attention, In Parasuraman, R. (Ed.), *The attentive brain* (pp. 163–186), Cambridge, MA: MIT Press.

Oades, R. D. (2000). Differential measures of "sustained attention" in children with attention-deficit/hyperactivity or tic disorders: realtions to monoamine metabolism.*Psychiatry Research, 93*, 165–178.

Parasuraman, R (Ed.). (1998). *The attentive brain*, MIT, Cambridge.

Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience*, *89*, 1009–1023.

Ruff, H.A. and Capozzoli, M.C. (2003). Development of attention and distractibility in the first 4 years of life, *Developmental Psychology, 39*, 5, 877–890.

Schultz, W., Dayan, P., Montague, P. R. (1997). A neural substrate of prediction and reward. *Nature*, *275*, 1593–1599.

Sutton, R. S., and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Watkins, C. J. C. H., and Dayan, P. (1992). Q-learning. *Machine Learning*, 9, 279–292.

Winberg, S. (2004). MSc Thesis. Lund University Cognitive Science.

Modeling Cognitive Development in the Human Brain

L. Andrew Coward
Department of Computer Science
Australian National University, ACT 0200, Australia
landrewcoward@shaw.ca

Abstract

Any architecture for modeling cognitive development must have several general characteristics. It must be possible to learn complex combinations of interacting cognitive capabilities using information derived from the same experience stream. Learning must be bootstrapped from experience with minimal a priori guidance and limited external guidance during learning, but in such a way that later learning does not interfere with earlier learning. Learning must be possible from single experiences. The architecture must provide an account for the observed dissociations between the various types of memory including semantic, episodic and procedural memory. A connectionist architecture with these characteristics is described.

1. Introduction

Much recent computational modeling of human cognition concentrates on accurate modeling of the phenomenology without regard to plausible matching between the detailed information processes employed and neural capabilities. An example is ACT [1], which provides quantitative modeling of experimental psychological data in terms of cognitive categories and processes which are defined by psychological observation but implemented by computer without regard to physiology. Another approach which focuses on practical applications is expert systems [e.g. 12] which capture the skills of human experts in a computer. These approaches make no effort to realistically model development of cognitive skills. Connectionist modeling is the only approach which makes claims to plausible modeling both of cognitive learning and of information processes at the neuron level [2]. However, the types of high level learning which have been successfully modeled tend to use restricted input and behavioural domains. For example, Roy et al [17] model the learning of words from visual and auditory inputs, but comment that "…the model is limited in its ability to deal with complex scenes…". Although Roy et al argue that their model has the potential for modeling learning of real cognitive processes, an issue often encountered with connectionist learning of more complex domains is the catastrophic forgetting problem [13] in which later learning sometimes overwrites and obliterates prior learning.

This paper argues that to be a plausible approach to modeling cognitive development, a model must demonstrate the potential to achieve a number of general characteristics exhibited by human development. A connectionist architecture which appears capable of these characteristics is then described. This architecture employs neural device algorithms with qualitative differences from conventional connectionist algorithms. The ability of the architecture to meet the required general characteristics and the strong dependence of this ability on the different device algorithms is described. In particular, the general ability to bootstrap memory and behaviour from experience and to use the same information recorded during experience to support episodic, semantic and procedural memory are described.

2. Criteria for Effective Modeling of Cognitive Development

There are a number of sometimes overlooked characteristics of human cognition which must be effectively addressed by any cognitive architecture which aims to model human development. Firstly, human beings learn a complex combination of different types of behaviour making use of the same experiences. For example, experimental psychology distinguishes between episodic, semantic, procedural and working memory and priming [19]. However, perceptual processing on the same stream of experience must generate information to support all these memory types, and information initially available to one memory type must over time become available in suitable form to others, while still remaining available to the original type. Thus episodic memories can result in semantic and procedural memories while still being accessible to episodic memory. Secondly, humans can learn new behaviour types with minimal interference with existing behaviours. This capability poses problems for conventional connectionist models, which tend to exhibit the catastrophic forgetting problem [13]. Thirdly, humans can bootstrap their cognitive capabilities from experience with minimal a priori guidance. For example, genetic guidance would not be able to specify categories of visual objects, but could perhaps provide preliminary and general associations between types of sensory input and types of behaviour which would need to be corrected and made much more specific by experience. The feedback available following behaviour can be reward or punishment, but not supervision in the sense of explicit indication of targets in terms of internal brain information structures as is required for connectionist supervised learning algorithms. However, a genetically defined tendency to imitate can make general reward and punishment feedback more efficient. Fourthly, humans are capable of significant, permanent learning from single experiences. For example, given a few seconds to examine each photograph in a set of 2500, subjects can later pick the familiar photograph from pairs in which only one came from the examined set at an accuracy level of 90% [20].

3. The Recommendation Architecture

Any system which must learn to perform a complex combination of interacting features with limited information handling resources in such a way that new learning does not interfere with prior learning tends to be constrained within a set of architectural bounds called the recommendation architecture [8]. These bounds define how the operations of the system are separated into modules, the ways in which modules interact, and the type of learning algorithms available to modules and devices. For a detailed description of the design of an electronic system implemented within these bounds, see [9].

In the recommendation architecture there is a primary architectural separation between a modular hierarchy called clustering and a subsystem called competition. Clustering defines a population of conditions within the available sensory information space and detects the occurrence of any defined condition. A subset of the conditions detected at any point in time is communicated to competition. Competition interprets each such condition as a set of recommendations in favour of a range of different behaviours, each with a different weight. Competition adds the weights of each recommended behaviour across all currently detected conditions, and implements the most strongly recommended behaviours. Consequence feedback following a behaviour can change the recommendation weights of recently active conditions into recent behaviours, but cannot change the definition of the conditions.

Devices in clustering learn and respond in radically different ways from conventional connectionist device algorithms. A clustering device permanently records a set of similar conditions. To be recorded, a condition must actually occur within the information available to the device, be similar in an information sense to conditions already recorded on the device, and at the time it occurs the device must also be receiving signals encouraging it to record conditions. These signals come from devices in other modules within clustering. The device is activated by the recording of a condition or by any subsequent repetition of the condition. On activation a device produces an output which is a series of activity spikes. The average rate of spike production indicates the number of its programmed conditions which are currently present, and a frequency modulation of the spike rate (i.e. bunching of spikes close to peaks in a regular modulation frequency) indicates the input population within which the condition was detected. In general, conditions will be detected by a device within a group of inputs much more strongly if a frequency modulation is present at the same phase on all the inputs, because otherwise fewer activity spikes will arrive within the time interval over which the device integrates its inputs. As an example of this frequency modulation in practice, if modulation is imposed on a subset of visual inputs corresponding with an area within a closed boundary (i.e. a visual object), only conditions within the object will be detected and recommended behaviours will be in response to this object. Frequency modulation also makes it possible to detect separate populations of conditions within two different objects in the same physical set of devices, if different phases of frequency modulation are imposed on inputs from the two objects. For a more detailed discussion of the frequency modulation mechanism, see [10].

A device in clustering thus records a set of similar conditions and indicates any repetition of a previously recorded condition. This device algorithm is in strong contrast with conventional connectionist device algorithms, in which devices have inputs with different weights which can be constantly adjusted, with no guarantee of response to an exact repetition of a condition that previously generated a response.

Devices in clustering are arranged in layers in which the condition defining inputs to one layer come from just one preceding layer. The first layer receives raw sensory inputs. This arrangement ensures that all the conditions detected within one layer are within the same range of complexity, where the complexity of a condition is the number of raw sensory inputs (including duplicates) that contribute to the condition either directly or via intermediate conditions. The layering also means that all the conditions detected at one time within a layer tend to be present within a system input state at one time (such as one visual object). Conditions within one range of complexity may be more appropriate for a particular behavioural function than conditions in other ranges.

The clustering device algorithm means that tight management is required over when and where additional conditions will be recorded. This management of change is a major role of the modular hierarchy in clustering. The first level of module above the device is a small area on one device layer. The next level is a column made up of a sequence of such areas across several layers. The next level is an array of such columns and the next level is a sequential block of such arrays. Each module detects a set of conditions made up of the sum of the sets detected by each device in the module. However, most of these conditions are only communicated within the module and used for change management within the module, only a small subset are communicated to other modules. A column module manages when conditions will be recorded within the column and within other columns in the same array. An array module ensures that some conditions are detected in every input state from a specific input domain. A block module ensures that the conditions detected within its constituent arrays are consistent with each other as indicated by a tendency for conditions in different arrays to have been active and recorded at the same time in the past. A block module then generates outputs in behaviourally useful ranges of complexity to competition.

Only some arrays target outputs on to competition, but any column within such an array can target competition. However, only devices within a column which detect conditions within a specific range of complexity (i.e. are located within a specific deviec layer) can target competition. These devices have sets of conditions which they detect, and the sum of these sets for a column is called the portfolio of the column. Portfolios are important in understanding the processes which lead to cognition.

4. Definition of Conditions and Portfolios

A vast range of raw inputs containing information about the external environment and the internal state of both the brain and the body are available to the brain from the senses. A somewhat oversimplified way of understanding the definition of information conditions is that one condition corresponds with a specific set of these inputs each being present to an individually specified degree. Because conditions cannot be specified a priori, there is a random element to the definition of conditions, and because conditions are not changed after being recorded, any one condition or portfolio is perceptually, cognitively and behaviourally ambiguous. Unambiguous meanings are only achieved in competition across populations of conditions. However, conditions with complexities of the same order of magnitude as visual object perceptions will tend to be less ambiguous with respect to categorization of visual objects than conditions on other levels of complexity, athough no conditions on any level correlate unambiguously with such categories.

This simple view of conditions is made considerably more complex because a column portfolio can be activated not only by the presence of its conditions within sensory inputs, but also indirectly by two other types of mechanism. One mechanism is that it can be activated if a number of other columns are already active which have often been active in the past at the same time as the column. The other mechanism is that columns can be activated if a number of other columns are already active which have recorded conditions at the same time in the past as the column. These indirect activations are behaviours which must be recommended by the already active columns into competition and accepted. When there is simultaneous activity or condition recording in two columns, there is a strong recommendation weight created in competition in favour of the activity of one column activating the other, but this recommendation weight declines fairly rapidly with time. However, if an indirect activation actually occurs in the course of generating a behaviour which is followed by positive consequences, the decline is reduced, and frequent such occurences stabilize or increase the weight.

These indirect activation mechanisms can be viewed as supplementing the conditions present in current sensory inputs with other conditions which have a significant probability of being relevant to determining the most appropriate current behaviour. For example, conditions which have been active in the past at the same time as currently present conditions may contain information about the current environment which cannot currently be observed [7, 11].

A newly recorded condition is made up of a set of currently active component conditions. Some of these component conditions may be combinations of currently present sensory inputs, and some could have been activated by one of the two indirect mechanisms. Both the definition of conditions in terms of sensory inputs and the relationship between sensory inputs and the resultant pattern of condition activation can therefore become very complex.

Learning occurs by permanent addition of conditions to modules on many levels including device, column, array and block. Conditions are defined heuristically, and there is no a priori knowledge of which higher level modules such as arrays will require many column portfolios, or which column portfolios will need many device level portfolios and which devices will require inputs from which other devices etc. Hence assignment of column and device resources must be performed heuristically on the basis of need. A resource management function must therefore assign provisional conditions to devices, devices to columns and columns to arrays on the basis of current need.

Resource management requires two components. One is a map of resources identifying which are unassigned, the other is a process for identifying appropriate connectivity [6]. Resource management is then a periodic process during which requirements for new resources are identified, resources are assigned, and appropriate provisional connectivity provided. Connectivity to support indirect portfolio activation on the basis of simultaneous condition recording could be efficiently provided via the resource map, at least initially. Connectivity to support indirect portfolio activation on the basis of prior or subsequent condition recording would for efficiency reasons tend to continue to be dependent on the map.

5. Behavioural Interpretation of Portfolios

Competition is made up of devices which total the excitatory and inhibitory weights of currently active inputs from a range of sources, and produce an output if the total exceeds a threshold. The devices adjust their input weights in response to consequence feedback. Unlike the device algorithms used in clustering, these algorithms are generally similar to the perceptron type algorithms used in conventional connectionist networks.

The competition system is made up of components corresponding on a one-for-one basis with all possible system behaviours. Each component is a device or group of devices. Some components correspond with behaviours which are "atomic" in the sense that the system can implement the behaviour or not implement it, can vary the speed and perhaps the degree with which the behaviour is performed, but cannot change the nature of the behaviour. An atomic behaviour could be the contraction of an individual muscle, or genetically programmed groups and sequences of such contractions. Other components correspond with higher level behaviours such as groups and sequences of atomic behaviours, and yet higher behaviours which are groups and sequences of such groups. At the highest cognitive levels, behaviour is achieved by outputs from clustering driving a sequence of competition components which in turn activate more specific competition components. Any very frequently occurring sequence or set of behaviours will tend to result in a new component in competition which receives most of the inputs from clustering and drives behaviour more directly into atomic behaviours.

The outputs of a component are the outputs of specific devices within the component. Because of the use of consequence feedback within competition, such outputs cannot have operationally complex meanings. Only two types of operational meaning are possible. One is a recommendation to perform the behaviour corresponding with the component. If such an output exits competition, it becomes a command to perform the corresponding behaviour. Otherwise it is directed at a range of components corresponding with more detailed or specific behaviours within the recommended type, and increases the probability of such behaviours being accepted. These more detailed components could also receive inputs directly from clustering.

The other type of operational meaning is a recommendation against performing any behaviour other than the component behaviour. Such outputs are directed at competition components corresponding with different behaviours. A high proportion of these outputs are directed at peer components, in other words components corresponding with different behaviours on roughly the same level of detail.

When a condition is recorded in clustering, it can immediately acquire a range of different behavioural meanings either directly through recommendation weights in competition or indirectly by incorporation in other conditions with such recommendation weights. Any subsequent change to the condition would therefore result in a wide range of uncontrolled behavioral side effects. The restrictions that conditions cannot change, and devices can only add similar conditions, limits these side effect much more effectively than perceptron type algorithms [9].

6. An electronic system with the recommendation architecture

An electronic system with the recommendation architecture has been implemented, and demonstrated the capability to define portfolios from experience with no a priori guidance, and to associate different combinations of portfolios with different behaviours using only reward and punishment feedback. The ability to learn with minimal interference with prior learning has also been demonstrated [8, 9]. Processes within the electronic implementation which strongly resemble cognitive processes including category learning, learning to activate appropriate visual information in response to words, and activation of mental images have been observed [9].

7. The Recommendation Architecture cognitive model

In the recommendation architecture, information can be accessed by four qualitatively different mechanisms. Firstly, the actual presence of a condition within current sensory inputs activates the substrate on which the portfolio containing the condition is recorded. Secondly, an activated portfolio can recommend activation of other portfolios which have often been active at the same time in the past, and the recommended portfolio will activate if adequate recommendation strength is present. A variant of this mechanism is activation of portfolios which were often active somewhat before or somewhat after activity in the active portfolio.

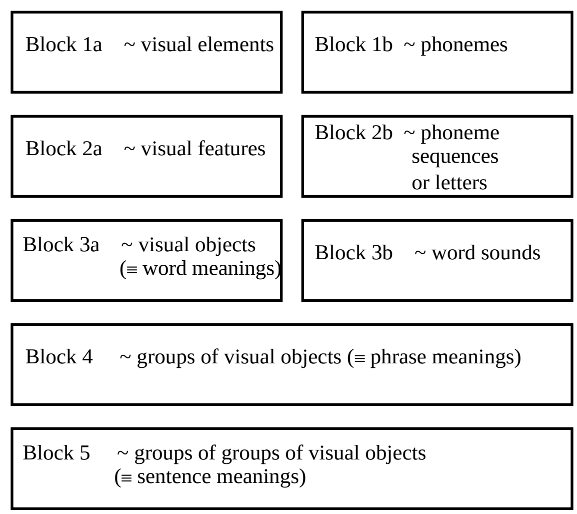


Figure 1. Architecture to support cognitive processes. Block modules detect conditions on five different levels of complexity, with condition defining information passing sequentially from top to bottom. The outputs of a block indicate the detection of conditions within the same range of complexity as the indicated cognitive category (features, objects, groups of objects etc.) but conditions do not correlate unambiguously with such categories. The subdivision of levels 1 through 3 reflects different input domains within which conditions are detected.

Thirdly, an activated portfolio can recommend activation of other portfolios which recorded conditions at the same time in the past. A variant of this mechanism is activation of portfolios which recorded conditions somewhat before or somewhat after an episode of condition recording in the active portfolio. Indirectly activated portfolios can in turn recommend activation of yet other portfolios. The fourth mechanism is comparison of recommendation weights. The weights of all active behaviours into each recommended behaviour are totaled, and the behaviours with the strongest weights are implemented.

The simplest arrangement of clustering blocks able to support complex cognitive behaviour is illustrated in figure 1, and examples of competition subsystems associated with one block are illustrated in figure 2. In figure 1, outputs from block 2a to competition indicate the detection of portfolios with a complexity comparable with visual features. Outputs from block 3a indicate portfolios comparable with visual objects, outputs from block 4 indicate portfolios comparable with groups of objects, and outputs from block 5 indicate portfolios comparable with groups of groups of objects. Block 4 will therefore detect portfolios in a sequence of perceptual objects, and block 5 will detects portfolios incorporating information derived from several area 4 outputs, in other words portfolios containing information derived from all members of the sequence.

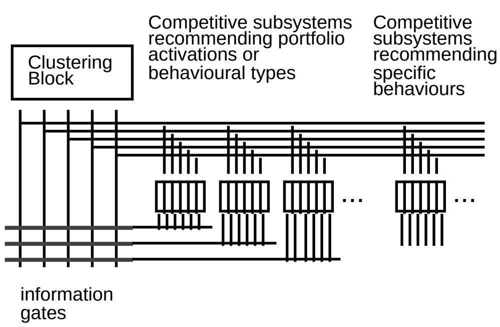


Figure 2 Competitive components receiving outputs from one sequence module. Different behavioural interpretations are placed upon the same clustering outputs by different components. There is competitive inhibition between and within competitive components to limit selected behaviours to a small, consistent set. In some cases the behaviour accepted by a competitive subsystem is release of the outputs from clustering which correspond with the behaviour to either the next clustering level or to a more detailed competition subsystem. This release behaviour is indicated by the information gates. In other cases competition outputs drive their corresponding individual behaviours, either external (e.g. eye movements) or internal (e.g. prolonging the activity of clustering neurons in specific modules).

Ten types of competition component corresponding with ten types of behaviour which could be recommended by a clustering area are as follows: prolong the activity of some currently active portfolios (for example of a group recommending a sequence of behaviours until the sequence is complete); activate portfolios active at the same time in the past as the currently active portfolio; activate portfolios containing conditions recorded at the same time in the past as some conditions in the currently active portfolio; activate portfolios containing conditions recorded just before or just after some conditions in the currently active portfolio; synchronize the activity (i.e. phase of modulation frequency) of several different groups of currently active portfolios; perform a general sequence of attention behaviours; perform a specific sequences of attention behaviours; perform an individual attention behaviour; speak a word; and say a phrase.

Competition receives outputs from portfolios currently being detected by clustering. If a portfolio has been present in the past at the same time as the performance of a number of different behaviours, it will have acquired recommendation weights in favour of or against those behaviours in the component corresponding with the behavior, depending on the consequence feedback from those behaviours. Competition adds the weights of all currently recommended behaviours and selects the behaviour with the largest weight. New portfolios are given an initial weight similar to the weights of the most similar previously existing portfolios, or genetically defined initial weights. If portfolios are new and no significant recommendation strength has been assigned in these ways, a behaviour can be selected randomly. Such random selection can be limited to behaviours within a behaviour type which has already been selected. Alternatively, a behaviour can be selected by imitation of an externally observed behaviour.

8. Bootstrapping of memory and behaviour

Behaviours can be defined heuristically with limited a priori guidance. Such definition will be illustrated by describing a possible process for acquisition of simple speech, with careful attention to the nature of the a priori (genetic) guidance needed. Learning goes through a series of steps generally consistent with observation of how humans learn to speak and understand words, but the purpose of this section is not to offer a formal model for speech acquisition but rather to demonstrate that speech can be acquired in the recommendation architecture model with only limited and plausible genetic guidance.

Genetic information specifies creation of a set of detailed competition components which drive muscle movements contributing to sound generation. Every possible such movement has a corresponding genetically specified competition component, and activation of the component results in the movement. Genetic information also specifies the existence of intermediate competition components which activate randomly selected sequences of detailed components and therefore generate sounds. Learning proceeds in a series of partially overlapping steps.

The first step is creation of an array of portfolios in clustering blocks 1b and 2b of figure 1 in response to hearing sounds, at different levels of condition complexity. Because speech is somewhat different from other sounds, there will be a tendency for speech related portfolios to be somewhat separate from the portfolios created in response to other sounds. This tendency could be reinforced by genetically determined connectivity biases within clustering.

The next step is generation of sounds using intermediate competition components. Initially a component is randomly selected after any activation of a portfolio population in the array of sound related portfolios in levels 1 and 2. A positive consequence feedback is genetically programmed to be generated if the portfolios activated in response to hearing an external sound (i.e. not self generated) are similar to the portfolios activated shortly afterwards in response to hearing a self generated sound. One effect of this consequence feedback is that the sequence of detailed components activated by the intermediate component is fixed long term. In other words, the presence of a sound in the environment results in production of that sound becoming instantiated in an intermediate component. If there is no feedback within some period of time, the intermediate component is reconfigured with a different randomly selected sequence of sound generating muscle movements, or deleted. The second effect of the consequence feedback is that the activated portfolios acquire recommendation strength in favour of activating the intermediate component. In other words, the behaviour of imitating sounds which are heard is acquired.

The next step in learning is that portfolios are created in area 3b of figure 1 in response to sequences of sequences of sounds which are heard. These portfolios will correlate partially (or ambiguously) with frequently heard sequences, and therefore with words which are heard. Higher level competition components are defined which activate randomly selected sequences of intermediate components. If a self generated sound sequence activates a portfolio population in area 3b similar to an immediately prior population activated by an external sound sequence, a genetically defined consequence feedback results in the higher level component being fixed and the active portfolios acquiring recommendation strength in favour of activating it. Thus the behaviour of imitating words which are heard is acquired.

The next step utilizes a genetically programmed tendency for the portfolios created in level 3b in response to sequences of sounds to have recommendation strength in favour of activation of portfolios created in level 3a in response to visual experiences, if the visual experience portfolios are often active at the same time as the sound sequence portfolios. The effect is that hearing the word will tend to activate a partial visual image of the type of object often seen when the word was present in the past. The portfolios making up a visual image will also recommend any other behaviours which have become associated with the object. In addition, the visual experience portfolios acquire recommendation strength in favour of activating the higher level component which tends to be activated by the sound sequence portfolios. The effect is that seeing the object will tend to result in speaking the corresponding word. Consequence feedback associated with the perceived behavior of adults in response to activating a higher level component (i.e. speaking a word) will affect the recommendation strengths of active portfolios in favour of the word just spoken.

Thus a set of genetically defined tendencies result in relatively efficient acquisition of simple speech behaviours. Learning does not require a priori internal definition of cognitive categories. Genetic information provided three types of information. Firstly, it indicates the available range of detailed muscle movements. Secondly, it biases initial connectivity in favour of the types of sensory inputs and portfolio condition complexity ranges which will most effectively drive those behaviours. Thirdly, it defines in general terms the circumstances in which consequence feedback will be generated and the effects of such feedback.

9. Different Types of Memory

It has been argued that there are a number of different memory systems in the brain, based on the observed dissociations between different memory phenomena [19]. These sytems include semantic, episodic, procedural and working memory. The following discussion will focus on semantic, episodic, and procedural memory, working memory is discussed in detail in [10].

There are two mechanisms by which information can be recorded in the recommendation architecture. One is permanent recording of conditions in clustering, the other is adjustment to recommendation weights in competition. The permanent recording of conditions means that the system has the capability to learn from single experiences. In the model there will be a level of condition recording in response to every experience, with higher levels for experiences with higher levels of novelty [6]. This higher level of recording in response to novelty accounts for the high human capability to detect the novelty or otherwise of an experience. For example, subjects exposed briefly to a set of several thousand photographs could a few days later distinguish between photographs in or not in the set with 90% accuracy[20].

As discussed earlier, there are four mechanisms by which information can be accessed in the recommendation architecture. The use of different combinations of these mechanisms can account for the phenomena and dissociations between semantic, episodic and procedural memory.

10. Episodic Memory

Episodic memory is memory of the past with a context of what else happened at the time, in contrast with semantic memory in which memories of facts are detached from memories of where those facts were learned [21]. Various types of recall experiments measure episodic memory.

In targeted recall subjects are asked to recall particular past events [14]. In cued recall, subjects are given a cue which may be a word [16] or a type such as "vivid memories" [18]. In involuntary recall, some specific environmental stimulus such as a smell or a taste brings a memory to mind unsought [3].

The starting point for targeted recall in the recommendation architecture model is hearing words which describe an event. Portfolios are activated which contain conditions within the sounds of the words. Secondary populations of portfolios containing conditions which occurred within visual and other sensory inputs are activated on the basis that the secondary portfolios have often been active in the past at the same time as the primary "auditory" portfolios. A significant proportion of the portfolios in these secondary populations were also active during the event, and a somewhat smaller proportion recorded conditions during that event. Because of the words used, the proportions are larger for the target event than for any other event.

All active portfolios have recommendation strengths in favour of activating other portfolios which recorded conditions at similar times in the past. Active populations at higher levels derived from the presence of words like "recall" have recommendation strength in favour of accepting these types of recommendations. Because the target event has the highest proportion of activated portfolios in the secondary population, acceptance of such recommendations will tend to result in an active tertiary population with an even higher proportion of portfolios which recorded conditions during the target event. This process is self reinforcing, especially if a large number of conditions were recorded during the target event. The resultant population will be experienced as a general re-perception of the original event, although in general the portfolios closest to input from the senses are not reactivated. The activated portfolios in this population have recommendation strength in favour of, for example, generating verbal descriptions of the event. Use of recommendation strength in favour of activating portfolios which recorded conditions somewhat before or somewhat after condition recording in currently active portfolios allows the re-perception to be set at the beginning of the event and to be moved through the event.

Recommendation strengths will always also be present in favour of activating portfolios on the basis of simultaneous past activity and simultaneous past recording during other events. The activated population is therefore unlikely to be an exact match for the original, although in general the higher the level of condition recording during the event, the greater the probability of a close match.

Cued recall operates in a very similar fashion. However, the initial secondary population may contain portfolios which recorded conditions during a number of past events. In the absence of specific indication of one event in the verbal cue, the tertiary population will evolve towards the event which happened to be represented by the highest degree of condition recording in the initial population. Events which resulted in a high degree of condition recording across many portfolios will tend to be the end points of this process.

Involuntary recall is the result of strong condition recording in portfolios activated in response to a sensory stimulus (for example, a novel smell or taste) at the same time as strong condition recording in other portfolios in response to some event. A later repetition of the sensory stimulus activates the portfolios which originally responded to that stimulus. These portfolios in turn activate portfolios which recorded conditions at the same time in the past, resulting in a re-perception of the event.

11. Semantic Memory

A typical way of measuring semantic memory in the laboratory is sentence and category verification. Sentence verification experiments measure the time for subjects to respond with the correctness of sentences like "Is a robin a bird" or "Is a penguin a bird". Category verification experiments are essentially equivalent and measure such times for simple category-exemplar pairs like bird-robin or bird-tree. It is found that for different members of the same category paired with the correct category, responses are faster for more typical category exemplars. For example, the response to bird-robin is faster than to bird-chicken. However, responses to clearly incorrect category-exemplar pairs like "Is canary an animal" is also fast [15].

The recommendation architecture model for category verification experiments can be understood by considering the portfolio populations activated in response to the words indicating category and exemplar. The portfolios activated in response to hearing the name of the category are portfolios which have often been active in the past when the category name was also present. Hence they will be portfolios also present when exemplars of the category were present, since this is the way the category is learned. Hence the active population is the set of portfolios which have most often been present when different category exemplars have been present. There will therefore be an overlap between the population activated in response to the name of the category and that activated in response to the name of the exemplar. This overlap will be greater for more typical exemplars, and very small if the exemplar is not a member of the category. The degree of overlap is itself a condition which can be detected and recommends for or against identifying the exemplar as a category member. If the exemplar is typical, overlap is substantial and its detection is rapid. Atypical exemplars have more moderate overlap, and more time may be required to expand the portfolio populations to include portfolios active at the same time in the past but slightly less often to achieve an overlap adequate to generate the appropriate verbal response. However, objects which are not in any way members of the category will have negligible overlap which again is detected rapidly. The model is therefore in agreement with the observations of [15].

In contrast with the spreading activation model of Collins & Loftus [4], there are no units which correspond with concepts like categories or the features of categories. Portfolios are groups of conditions in which there has been a degree of randomness in the definition of individual conditions, but conditions within one portfolio have some similarity with each other and have tended to occur at similar times in the past. A portfolio may therefore have a probabilistic correlation with many different features and categories, with the probabilities expressed, for example, as recommendation weights into naming the features or categories.

12. Procedural Memory

Procedural memory is defined as the ability to acquire skills. Observations of amnesics indicate that such memory is at least partially dissociated from semantic memory, since amnesics can acquire such skills at apparently normal rates. Thus amnesics can acquire motor skills such as mirror tracing tasks [5].

In the recommendation architecture model for learning a skill, portfolios activated within clustering in environments where the skill is relevant must acquire weights in competition associated with skilled behaviours. These portfolios will generally include both new information elements resulting from novelty in the environments and information elements recorded in prior experiences. Although the new elements may be particularly useful for recommending the new behaviours, some skill learning would be possible using only previously recorded elements which happen to occur in the new environments. Thus skill acquisition could proceed in the absence of condition recording.

13. Conclusions

The recommendation architecture cognitive model demonstrates the general capabilities to learn complex combinations of capabilities utilizing information drawn from a common experience stream, to bootstrap learning from experience with minimal and genetically plausible a priori guidance and limited and plausible external guidance during experience. The catastrophic interference between new and prior learning found in other connectionist architectures can be avoided. Significant, permanent learning is possible in response to single experiences. The information recording and access mechanisms of the recommendation architecture can provide an account for the phenomena of and disociations between semantic, episodic and procedural memory. The recommendation architecture thus has considerable advantages as a starting point for modeling cognitive development.

References

[1] Anderson, J.R. and Lebiere, C. (1998). *The Atomic Components of Thought*. NJ: Erlbaum.Coward, L. A. (1990). *Pattern Thinking*. New York: Praeger.

[2] Ballard, D. H. (1997). Introduction to Natural Computing. MIT Press.

[3] Berntsen, D. (1996). Involuntary autobiographical memories. *Applied Cognitive Psychology* 10, 435 - 454.

[4] Collins, A. & Loftus, E. (1975). A spreading-activation theory of semantic processing. *Psychological Review* 82(6), 407 – 428.

[5] Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia* 6, 225 - 264.

[6] Coward, L. A. (1990). *Pattern Thinking*. New York: Praeger.

[7] Coward, L. A. (1999). A physiologically based theory of consciousness, in S. Jordan (Ed.). *Modeling Consciousness Across the Discipline*. Maryland: UPA.

[8] Coward, L.A. (2000). A Functional Architecture Approach to Neural Systems. *International Journal of Systems Research and Information Systems*, 9, 69-120.

[9] Coward, L. A. (2001). The recommendation architecture: lessons from the design of large scale electronic systems for cognitive science. *Journal of Cognitive Systems Research*, 2, 2, 111-156.

[10] Coward, L. A. (2004). The Recommendation Architecture Model for Human Cognition.Proceedins of the 2004 Conference on Brain Inspired Cognitive Systems, Stirling, Scotland.

[11] Coward, L.A. and Sun, R. (2004). Some Criteria for an Effective Scientific Theory of Consciousness and Examples of Preliminary Attempts at Such a Theory. Consciousness and Cognition in press.

[12] Firebaugh, Morris W. (1993). Artificial Intelligence, A Knowledge-Based Approach. PWS-Kent Publishing Company.

[13] French, R. M. (1999). Catastrophic Forgetting in Connectionist Networks. *Trends in Cognitive Science* 3(4), 128-135.

[14] Neisser, U., & Libby, L. K. (2000). Remembering life experiences. In E. Tulving and F. I. M. Craik (Eds.).*The Oxford Handbook of Memory*. Oxford: Oxford University Press.

[15] Rips, L., Shoben, J. & Smith, E. Semantic Distance and Verification of Semantic Relations. *Journal of Verbal Learning and Verbal Behaviour* 12, 1 – 20.

[16] Robinson, J. A. (1976). Sampling autobiographical memory. *Cognitive Psychology* 8, 578 - 595.

[17] Roy, D. K. and Pentland, A. P. (2002). Learning words from sights and sounds: a computational model. Cognitive Science 26, 113 - 146.

[18] Rubin, D. C. & Kozin, M. (1984). Vivid memories. *Cognition* 16, 81 - 95.

[19] Schacter, J.C., Wagner, A. D., & Buckner, R. L. (2000). Memory systems of 1999. In E. Tulving & F. I. M. Craik (Eds.). *The Oxford Handbook of Memory*. New York: Oxford University Press.

[20] Standing, L., Conexio, J., & Haber, R.N. (1970). Perception and memory for pictures: single-trial learning of 2500 visual stimuli. *Psychonomic Science* 19(2), 73 - 74.

[21] Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Reviews of Psychology* 53, 1 – 25.

Solving Complex Problems Using Hierarchically Stacked Neural Networks Modeled on Behavioral Developmental Stages

Michael Lamport Commons, Ph.D.
Department of Psychiatry
Harvard Medical School
Commons@tiac.net

Myra Sturgeon White, Ph.D., J.D.
Department of Psychiatry
Harvard Medical School
mswhite@fas.harvard.edu

Abstract

Neural networks have greatly improved how well we model human behavior and solve complex problems. Their success lies in their schematic representation of model neuronal function and organization within the brain. Neural networks do not fully mimic the brain's capacity to combine behaviors in novel ways to solve complex problems so they cannot solve complex problems that humans solve easily. Hierarchically stacked neural networks model how humans acquire complex behavioral sequences. We present a blueprint for designing neural networks that incorporate Commons' Model of Hierarchical Complexity (1998) and thus, more closely parallel the behavioral learning process in humans with its capacities to flexibly solve and respond to complex problems. Commons' Model is based on research showing that cognitive development in humans proceeds through a series of ordered stages. Actions and tasks performed at increasingly higher stages are built on each proceeding stage. Hierarchical stacked neural networks in our design parallel this process by being ordered in the same way as the developmental learning sequence outlined in Commons' model. The mathematical models used within each network in a stack are based on its developmental stage and not the logic of a task.

Using our model, we have designed a system that directs incoming customers' calls to correct departments in a large organization based on customers' oral statements and responses to questions asked by the system.

In this system, hierarchical stacked computer neural networks are based on Commons' (1998) Model of Hierarchical Complexity which models human development and learning. The system allows computers to mimic higher level human cognitive processes and make sophisticated distinctions between stimuli; and allow computers to solve more complex problems.

Traditional neural networks are limited because they only model neuronal function and relatively simple physiological structures in the brain. By failing to model the manner in which human cognition develops, these networks are unable to reproduce the more complex behaviors of humans and have limited problem-solving ability. As a consequence, they cannot solve many problems that humans solve easily.

Theoretical Underpinnings of Commons' Model

Humans pass through a series of ordered stages of development. Behaviors performed at each higher stage of development are always more complex than those performed at the immediately preceding stage. Movement to a higher stage of development occurs by the brain combining, ordering and transforming the behavior which was at the preceding stage. This combining and ordering of behaviors must be non-arbitrary.

Commons' Model of Hierarchical Complexity

The model identifies 14 stages of hierarchical complexity in development. According to this model, individual tasks are classified by their highest stage of hierarchical complexity. The model is used to deconstruct tasks into the behaviors that must be learned at each stage in order to build the behavior needed to successfully complete a task

Hierarchical Stacked Computer Neural Networks Based on Commons' Model

Hierarchical stacked computer neural networks based on Commons' (1998) Model recapitulate the human developmental process. Thus, they learn the behaviors needed to perform increasingly complex tasks in the same sequence and manner as humans. This allows them to perform high-level human functions such as monitoring complex human activity and responding to simple language.

They can consist of up to 14 architecturally distinct neural networks ordered by stage of hierarchical complexity. The number of networks in a stack depends on the hierarchical complexity of the task to be performed. The type of processing that occurs in a network corresponds to its stage of hierarchical complexity in the developmental sequence. In solving a task, information moves through each network in ascending order by stage

Design of Neural Networks Based on Commons' Model

The task to be performed is first analyzed to determine the sequence of behaviors needed to perform the task and the stages of development of the various behaviors. The number of networks in the stack is determined by the highest stage behavior that must be performed to complete the task. Behaviors are assigned to networks based on their stage of hierarchical development.

Example: System to Answer Customer Calls and Transfer Them to a Department

Features

Answers calls and based on callers' oral statements and

directs them to a department
Queries callers for more information
Achieves the language proficiency of a three year-old
Asks simple questions

Design of Network

Uses 4 neural networks, N2, N3, N4 and N5.
N2: Circular Sensory Motor Stage Network: Forms open-ended classes
N3: Sensory Motor Stage Network: Recognizes classes
N4: Nominal Stage Network: Identifies relationships between simple concepts and labels them
N5: Sentential Stage Network: Forms simple sentences, constructs complex relationships and orders relationships

Processes Performed at Each Stage

Input: Front-end speech recognition system
N2: Uses inter-word intervals to group words
N3: Maps words to pretaught words central to organizational environment
N4: Identifies relationships between words and links to concepts
N5: Maps relationships between concepts and makes simple queries to caller
Output: Chooses department and checks with caller to see if it is the correct place to send call. Figure 1 illustrates a stacked neural network 10 in accordance with one embodiment of the present invention. Stacked neural network 10 comprises a plurality of up to 14 architecturally distinct, ordered neural networks 20, 22, 24, 26, ..., of which only four are shown. The number of neural networks in stacked neural network 10 is based on the number of consecutive stages needed to complete the task assigned. A sensory input 60 to stacked neural network 10 enters lowest stage neural network 20. The output of each of neural networks 20, 22, 24, 26, ..., is the input for the next neural network in the stack.

The highest-stage neural network 26 in the stack produces an output 62. Each of neural networks 20, 22, 24, 26, ..., except for the first in the stack, neural network 20, can provide feedback 30, 32, 34, 36, 38, 40 to a lower-stage neural network 20, 22, 24, ... Feedback adjusts weights in lower stage neural networks. Neural networks in the stack 20, 22, 24, 26 ... can send a request 50 for sensory input 60 to feed more information to neural network 20. A neural network can send this request when its input does not provide enough information for it to determine an output.

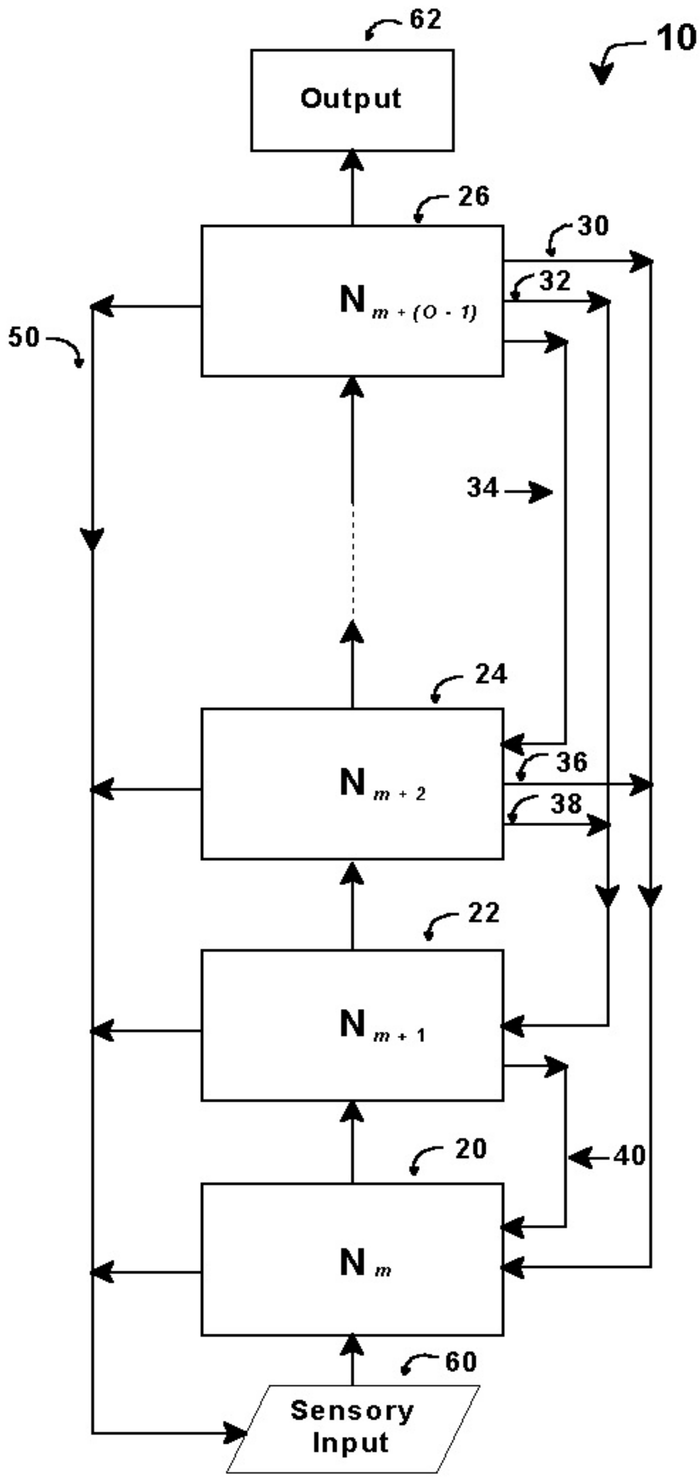
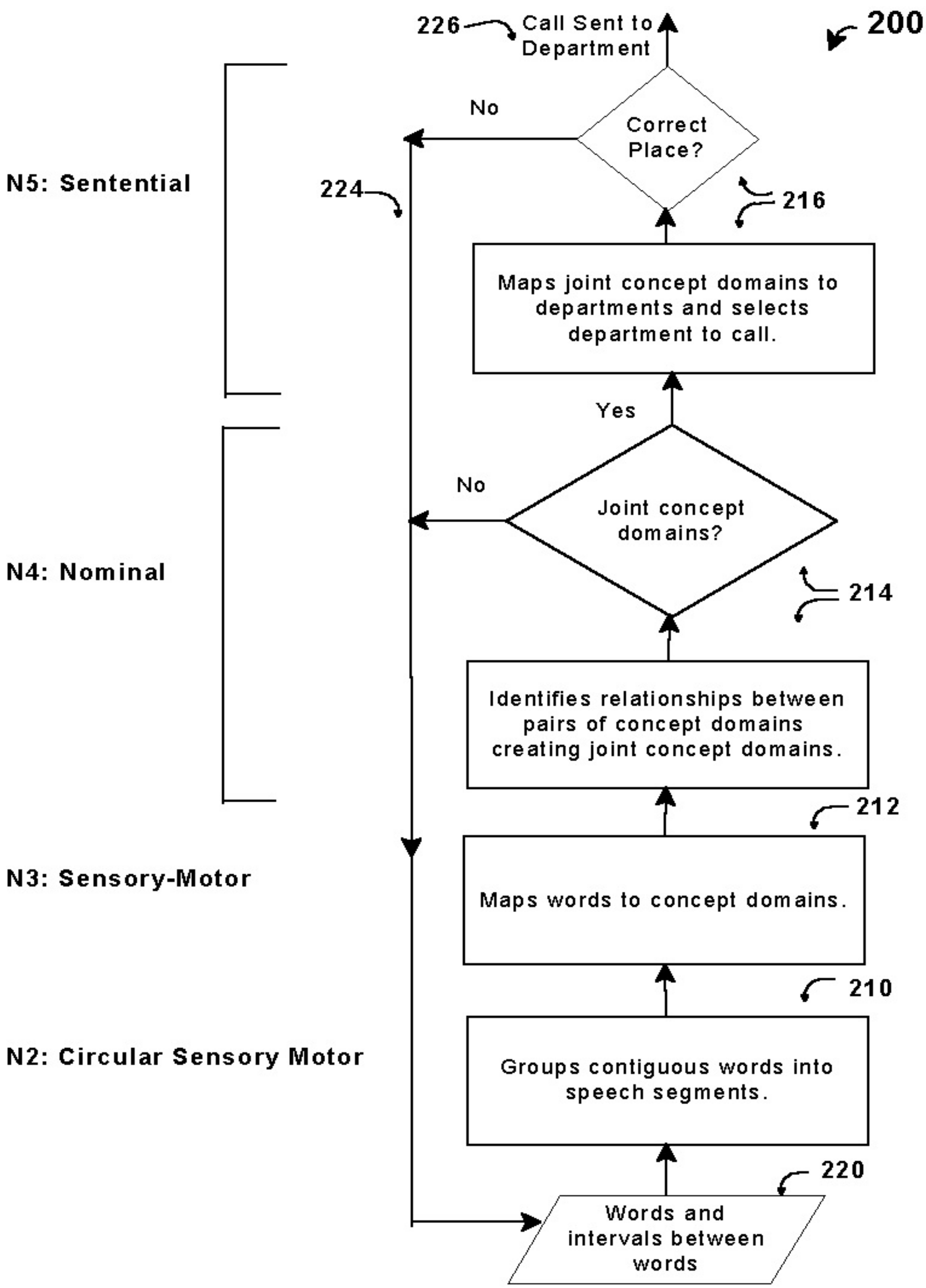

62
10
Output
26
30
32
$N_{m+(O-1)}$
50
34
24
36
N_{m+2}
38
22
N_{m+1}
20
40
N_m
60
Sensory Input

FIG. 1

Figure 7 is a high level flow chart 200 that illustrates a series of four major processing steps 210, 212, 214, and 216 for the second embodiment of the present invention: An Intelligent Control System that Directs Customer Calls to the Correct Department in a Large Organization. A front-end speech recognition software system 220 translates customers' utterances into words, measures the time intervals between each word and removes articles, prepositions and conjunctions from the utterances. Words and time intervals between words are processed at a step 210 which performs tasks at the Circular Sensory Motor stage/order. At this stage/order, open ended classes are formed. In this step, time intervals between the words are used to break the word stream into contiguous word groups that reflect natural speech segments.

A next processing step 212, maps words in each group produced at processing step 210 into clusters called concept domains which represent concepts that are central to the company's functions. Processing step 212 uses tasks at the Sensory-Motor stage/order. At this stage/order, words are identified as belonging to meaningful classes. A processing step 214, next identifies simple relationships between pairs of concept domains produced by process 212. Processing step 214 uses tasks at the Nominal stage/order. At this stage/order, simple relationships are formed between concepts. If processing step 214 is unable to identify any joint concept domains from the concept domains input from step 212, the customer is queried for more information. The customer's responses are sent to the front-end speech recognition system 220 and then processed at steps 210 and 212 before being processed at step 214.

Once step 214 identifies joint concept domains, then a processing step 216 maps the joint concept domains to clusters of neurons that represent relationships between company products and functions. This step operates at the Sentential stage/order. At this stage/order, simple sentences are formed, relationships between more than two concepts are understood and relationships are ordered. A department is competitively selected at this step based on the patterns of activation from the mapping of joint concept domains. The customer is queried to determine whether they would like their call sent to this department. If the customer answers affirmatively, a connection 226 is made to the department selected by the system. If they do not want this department, the customer is queried for more information. A response set of their utterances 224 is sent to the front-end speech recognition system 220. The words produced by the speech recognition system are input to processing step 210 and are processed in the same manner as the customer's initial utterances.



200
226
Call Sent to Department
No
Correct Place?
216
N5: Sentential
224
Maps joint concept domains to departments and selects department to call.
Yes
No
Joint concept domains?
214
N4: Nominal
Identifies relationships between pairs of concept domains creating joint concept domains.
212
N3: Sensory-Motor
Maps words to concept domains.
210
N2: Circular Sensory Motor
Groups contiguous words into speech segments.
220
Words and intervals between words

FIG. 7

Figure 8 illustrates a stacked neural network 230 for the second embodiment of the present invention: An Intelligent Control System that Directs Customer Calls to the Correct Department in a Large Organization. Stacked neural network 230 comprises a stack of 4 architecturally distinct, ordered neural networks 240, 242, 244, and 246. Words and the time intervals between words are input into a neural network 240 from a front-end speech recognition system 220 that translates customer utterances into words, computes time intervals between words, and removes articles, prepositions and conjunctions.

Neural network 240 performs Circular Sensory Motor stage/order tasks that group words into contiguous word groups based on the time intervals between words that naturally segment speech. Output from neural network 240 is input into a neural network 242. Neural network 242 performs Sensory-Motor stage/order tasks that map words into concept domains that represent company functions. Output from neural network 242 is input into a neural network 244 that performs Nominal stage/order tasks that identify simple relationships between pairs of concept domains, thereby creating joint concept domains. If no joint concept domains are identified by neural network 244, then a query 252 is output to the customer for more information. This new information from the customer is sent to the front-end speech recognition system 220 and then processed by neural networks 240 and 242 before neural network 244 continues processing the customer's speech. Once joint concept domains are identified in neural network 244, they are input into a neural network 246. It performs Sentential stage/order tasks that map the joint concept domains to clusters representing product/department relationships. Based on levels of department activation, a department is selected to be the department most likely to satisfy the customer's needs. A query 254 is then sent the customer to ask them if they would like to be sent to this department. If the customer responds "yes," then the call is sent to the department selected by neural network 246. If the customer responds "no," then the customer is further queried. A set of the customer's responses 254 are sent to the front-end speech recognition system 220 and then processed by neural networks 240, 242 and 244 before being processed by neural network 246. A group of feedback adjustments 256 are sent to neural networks 246 and 244 to adjust their weights based on the success or failure of the stacked neural network in selecting a department for the customer.

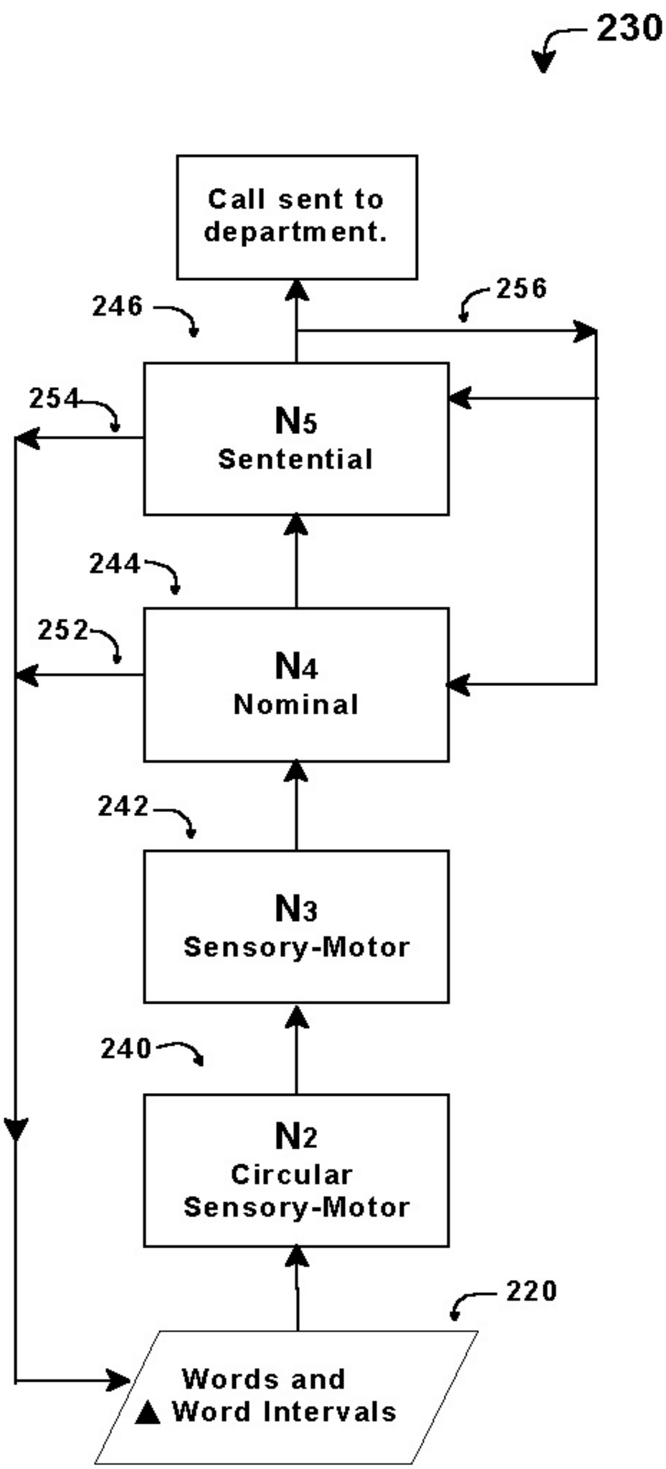

230
Call sent to department.
246
256
254
N5
Sentential
244
252
N4
Nominal
242
N3
Sensory-Motor
240
N2
Circular Sensory-Motor
220
Words and
▲ Word Intervals

FIG. 8

Simulating development in a real robot

Gabriel Gómez*, Max Lungarella**, Peter Eggenberger Hotz*, Kojiro Matsushita* and Rolf Pfeifer*
*Artificial Intelligence Laboratory
Department of Information Technology, University of Zurich, Switzerland
Andreasstrasse 15, CH-8050 Zurich, Switzerland
gomez@ifi.unizh.ch
**Neuroscience Research Institute
Tsukuba AIST Central 2, Japan
max.lungarella@aist.go.jp

Abstract

We present a quantitative investigation on the effects of a discrete developmental progression on the acquisition of a foveation behavior by a robotic hand-arm-eyes system. Development is simulated by increasing the resolution of the robot's visual system, by freezing and freeing mechanical degrees of freedom, and by adding neuronal units to its neural control architecture. Our experimental results show that a system starting with a low-resolution sensory system, a low precision motor system, and a low complexity neural structure, learns faster that a system which is more complex at the beginning.

1. Introduction

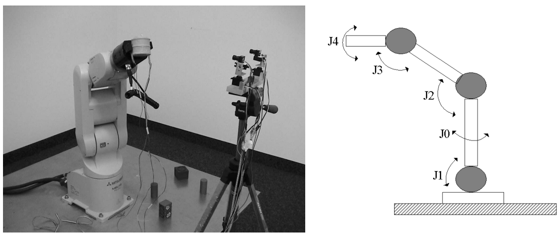
Development is an incremental process, in the sense that behaviors and skills acquired at a later point in time can be bootstrapped from earlier ones, and it is historical, in the sense that each individual acquires its own personal history [15]. It is well known that newborns and young infants have various morphological (bodily), neural, cognitive, and behavioral limitations, e.g., in neonates color perception and visual acuity are poor (implying a poor tracking behavior) [14]; working memory and attention are initially restricted (giving rise to reduced predictive abilities); motor immaturity is even more obvious, movements have a lack of control and coordination (producing inefficient and jerky movements).

The state of immaturity of sensory, motor, and cognitive systems, a salient characteristic of development, at first sight appears to be an inadequacy. But rather than being a problem, early morphological and cognitive limitations effectively decrease the amount of information that infants have to deal with, and may lead, according to a theoretical position pioneered by [16], to an increase of the adaptivity of the organism. A similar point of view was made with respect to neural information processing by [4]. For instance, it has been suggested that by initially limiting the number of the mechanical degrees of freedom that need to be controlled, the complexity of motor learning is reduced. Indeed, an initial freezing (i.e., not using) of degrees of freedom followed by a subsequent freeing (i.e., release) might be the strategy figured out by Nature to solve the degrees of freedom problem first pointed out by [1], that is, despite the highly complex nature of the human body, well-coordinated and precisely controlled movements emerge over time. In other words, it is possible to conceptualize initial sensory, motor, and cognitive limitations as an adaptive mechanism on its own right, which effectively helps speeding up the learning of tasks, and acquisition of new skills by simplifying the external world of the agent.

The aim of this paper is to provide support for the hypothesis that ”starting small” makes and agent more adaptive, and robust against environmental perturbations. Other attempts have shared explicitly or implicitly a similar research hypothesis. [11], for instance, applied a developmentally inspired approach to robotics in the context of joint attention. The authors showed that by having the visual capabilities of a robot mature over time, the robot could learn faster. The effect of phases of freezing and freeing of mechanical degrees of freedom for the acquisition of motor skills was examined by [8] and [2]. For a detailed review of the field of developmental robotics see [9]. Although based on the same research hypothesis, the present study makes at least two novel contributions: (a) it considers the concurrent ”developmental changes” in three different systems, i.e., sensory, motor, and neural; and (b) it quantitatively compares a ”developing” system to a ”nondeveloping” system.

Obviously, an understanding of development cannot be limited to investigate control architectures only, but must include considerations on physical growth, change of shape, and body composition, which are salient characteristics of maturation. Given the current state of technology, however, it is not easy to construct physically growing robots. We propose a method to "simulate" development in an embodied artifact at the levels of sensory, motor, and neural system. We use a high-resolution sensory system and a high-precision motor system with a large number of mechanical degrees of freedom, but we start out by simulating, in software, lower resolution sensors (i.e. by averaging over neighboring pixels in the camera image, or by using only a few pressure sensors) and an increased "controllability" (i.e., by freezing most degrees of freedom). Over time, we gradually increase the resolution of the sensors and the precision of the motors by successively freeing these "degrees of freedom" (i.e. by starting to use the "frozen" joints) and added neuronal units to the neural control architecture.
In the following, we present quantitative results demonstrating how a concurrent increase of sensory resolution, motor precision and neural capabilities can shape an agent's ability to learn a task in the real world, and speed up the learning process.

In the following section we introduce our experimental setup, we then proceed to specify the robot's task in section 3. The neural network and how it is embedded in the robot are described in section 4. The developmental approach is described in sections 5 and 6. The experiments performed are described in section 7, and the results are discussed in section 8. Finally, we point to some future research prospects in the last section.



(a) (b)

Figure 1. Experimental setup consisting of six degrees of freedom robot arm, four degrees of freedom color stereo active vision system, and a set of tactile sensors placed in the robots gripper.

2 Experimental setup

We performed our experiments by using the experimental setup shown in Figure 1. It consisted of the following components:

- *Robot arm*. An industrial robot manipulator (Mitsubishi MELFA RV-2AJ) with six degrees of freedom (DOF). As can be seen in the Figure 1b, joint $J0$ ("shoulder") was responsible for the rotation around the vertical axis, joint $J2$ ("elbow"), joint $J1$ ("shoulder") and joint $J3$ ("wrist") were responsible for the up and down movements; joint $J4$ ("wrist") rotated the gripper around the horizontal axis. The additional DOF came from the gripping manipulator.

- *Color stereo active vision system*. Two frame grabbers were used to digitalize images with a resolution of 128x128 pixels, down sampled at a rate of 20Hz.

- *Sensory-motor control board*. The communication between the computer and the motor control board that drives the active vision system and gets the tactile information was via a USB controller based on the Hitachi H8 chip.

- *System architecture*. The system architecture was composed of two computers Pentium III/600 MHz and the robot arm controller connected together in a private local area network based on the TCP/IP protocol, one computer controlled the robot arm and the other acquired the tactile input as well as the visual input from the active vision system.

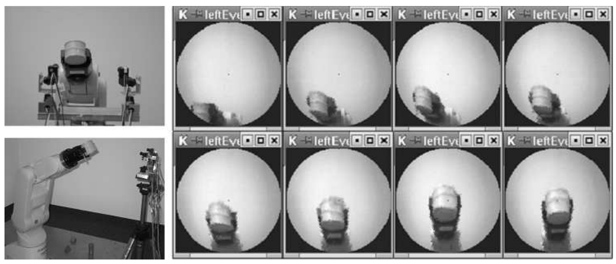


Figure 2. Robotic setup performing an experiment moving an object from the bottom-left corner of its visual field to the center of it. The observer's perspective can be seen on the left side, while the robot's perspective is shown on the right side.

3 Task specification

The task of the robot was to learn how to bring a colored object from the periphery of the visual field to the center of it by means of its robotic arm. It is important to note that although it would have been possible to program the robot directly to perform this task, our aim here is to quantify the effects of developmental changes on the learning performance. We are not seeking biological plausibility, but biologically inspired mechanisms of adaptive and autonomous behavior.

At the outset of each experiment the active vision system was initialized looking at the center of the visual scene *(x_c, y_c)* and the position of its motors were kept steady throughout the operation. The robot arm was placed at a random position at the periphery of the robot's visual field and a colored object was put on its gripper. Once the object was detected by the pressure sensors the robot started to learn how to move the arm in order to bring the object from the periphery of the visual field *(x_0, y_0)* to the center of it *(x_c, y_c)*. In other words, the eyes should teach the robot arm to solve the task, the object was the visual stimulus and the way to solve the task was the movement of the robot arm. A typical experiment is shown in Figure 2. For more details see [5, 6].

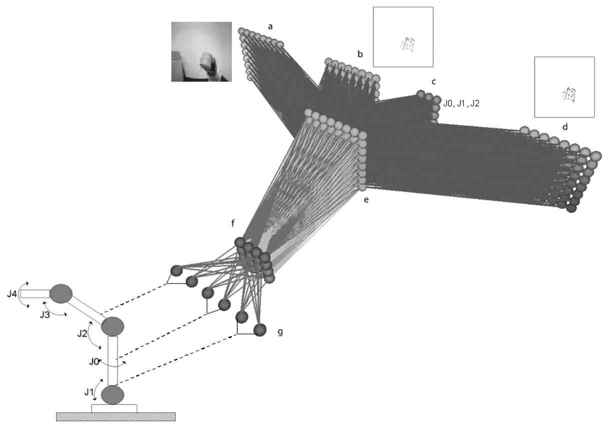


Figure 3. Neural structure and its connections to the robot's sensors and motors. Neuronal areas: (a) RedColorField. (b) RedMovementToRightField. (c) ProprioceptiveField. (d) RedMovementToLeftField. (e) NeuronalField. (f) MotorField. (g) MotorActivites.

4 Neural control architecture

The components of the neural structure and its connections to the robot arm are depicted in Figure 3.

4.1 Sensory field

- *Color information.* Three receptor types are considered: red (r), green (g), and blue (b). A "broadly" color-tuned channel was created for red:

$$R = r - (g + b)/2 \tag{1}$$

This channel yields maximum response for the fully saturated red color, and zero response for black and white inputs. The negative values were set to zero. Each pixel was then mapped directly onto the 8x8 neuronal units of area *RedColorField* (see Figure 3a). The activity S_i of the i-th neuron of this area was calculated as:

$$S_i = \begin{cases} 1.0 & : \ if \ \ R_i > \theta_1 \\ 0.0 & : \ otherwise \end{cases} \tag{2}$$

Where R_i is the value of the red color-tuned channel for the i-th pixel; and θ_1 is a threshold value.

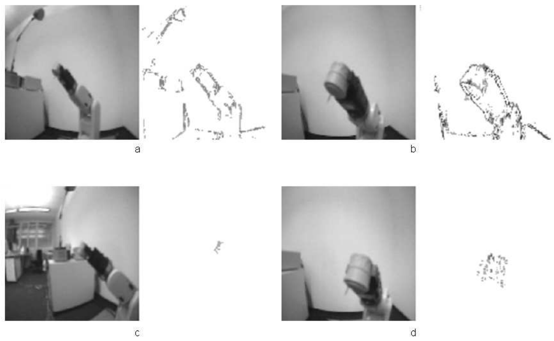


Figure 4. Motion detection. (a) Movement was detected from right to left. (b) Movement was detected from left to right. (c) and (d) Motion detectors reacting only to red objects moving in the environment.

- *Motion detection.* Motion detectors were created to detect movements of red objects in the environment. These motion detectors are based on the well-known elementary motion detector (EMD) of the spatio-temporal correlation type [10], a description of the model implemented, can be found in [7]. Motion detectors reacting to red objects moving to the right side of the image were mapped directly to neuronal units of the area *RedMovementToRightField* (see Figure 3b) and the motion detectors reacting to red objects moving to the left side of the image were mapped directly to neuronal units of the area *RedMovementToLeftField* (see Figure 3d). Both neuronal areas have a size of 8x8. The activities of the neurons in these areas were calculated as:

$$S_i = \begin{cases} 1.0 & : \ if \ \ |EMDOutput_i| > \theta_2 \\ 0.0 & : \ otherwise \end{cases} \tag{3}$$

Where S_i is the activity of the i-th neuron; $EMDOutput_i$ is the output of the motion detector at position i-th; and θ_2 is a threshold value.

- *Proprioceptive information.* The movements of each joint of the robot arm were encoded using eight neuronal units. During the experiments the size of the neural area *ProprioceptiveField* (see Figure 3c) was increased. The minimum size was 8x1 when it encoded the joint: $J0$, it had a medium size of 8x2 when it encoded the joints: $J0$ and $J2$, and it had a maximum size of 8x3 for encoding the joints: $J0$, $J1$, and $J2$. Joint $J0$ had a range of movements from -60 to 60 degrees, joint $J1$ moved in a range from -25 to 25 degrees, and joint $J2$ moved in a range from 0 to 100 degrees.

4.2 Neuronal field and motor field

The size of the neuronal area *NeuronalField* (see Figure 3e) was 8x8 and its neuronal units had a sigmoid activation function.
During the experiments the size of the neuronal area *MotorField* (see Figure 3f) was increased. The minimum size was 4x4 and the maximum was 16x16 and its neuronal units had a sigmoid activation function whose outputs were passed directly to the *MotorActivities* (see Figure 3g) for controlling the joints of the arm: $J0$, $J1$ and $J2$. The size of the neuronal area *MotorActivities* was 6x1.

4.3 Synaptic connections

Neuronal units in the areas *RedColorField*, *RedMovementToLeftField*, and *RedMovementToRightField* were connected retinotopically to the neuronal units in area *NeuronalField*. The neuronal units in the area *ProprioceptiveField* were fully connected to the neuronal units in area *NeuronalField*. The neuronal units in area *NeuronalField* were fully connected to the neuronal units in area *MotorField*, which in turn were fully connected to the *MotorActivities*.

4.4 Learning Mechanism

The active neurons controlling the robot arm were ”rewarded” if the movement of the arm brought the colored object closer to the center of the visual field and ”punished” otherwise. In this way the synaptic connections between the neuronal areas *NeuronalField* (see Figure 3e) and *MotorField* (see Figure 3f) were changed. A learning cycle (i.e., the period during which the current sensory input is processed, the activities of all neuronal units are computed, the connection strength of all synaptic connections are computed, and the motor outputs are generated) had a duration of approximately 0.35 seconds. For more details see [3] and [5, 6].

5 Simulating development in a real robot

Because we are dealing with embodied systems, there are two dynamics, the physical one or body dynamics and the control one or neural dynamics. There is the deep and important question of how the two can be coupled in optimal ways. It has been hypothesized that given a particular task environment, a crucial feature of adaptive behavior is a balance between the complexity of an organism’s sensor, motor, and control system (this is also referred to as principle of ecological balance) [13] and [12]. Here, we extended this principle to developmental time, and attempted to comply to it by simultaneously increasing the sensor resolution, the precision of the motors, as well as the size of the neural structure. Such concurrent changes are thought to simplify learning processes providing the basis for maintaining an adequate balance between the complexity of the three sub-systems, which reflects the development of biological systems.

5.1 Increasing the motor capabilities of the robot

The development of the robot’s controllability was achieved by an initial freezing of mechanical degrees of freedom and gradual releasing of them. At the beginning only joint $J0$ was used, during the second developmental stage two joints were used (i.e., $J0$ and $J2$) and during the third developmental stage three joints were used (i.e., $J0$, $J1$, and $J2$).

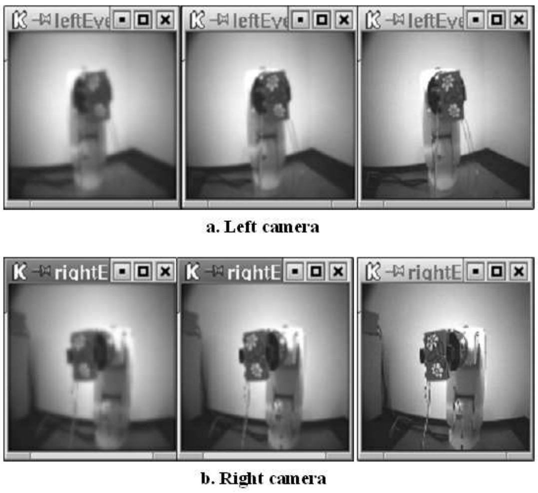


Figure 5. Gradual Increase of the sensory resolution. From left to right the image develops from blurred to high resolution.

5.2 Increasing the sensory capabilities of the robot

Increasing the resolution of the cameras was achieved by means of a gradual increase of the sharpness of a Gaussian blur lowpass filter applied to the original image captured by the cameras (see Figure 5(right)). Figures 5(left) and 5(center) show the result of applying a 5x5 and a 3x3 Gaussian kernel to the original image respectively.
The number of pressure sensors mounted on the gripper of the robot was also increased over time.

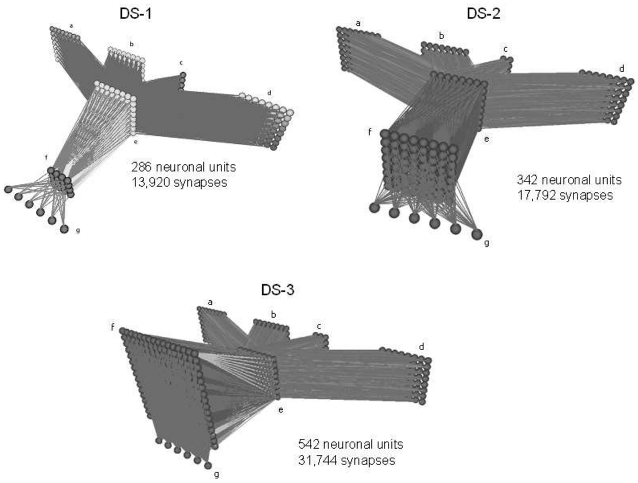


Figure 6. Gradual increase of the neural structure to cope with more sensory input and with more degrees of freedom of the motor system.

5.3 Increasing the complexity of the neural structure

In Figure 3 an overview of the neural network and its connections to the sensory-motor system is given. The neural network was gradually enhanced to cope with more sensory input and with more degrees of freedom of the motor system by (a) adding eight neuronal units to the area *ProprioceptiveField* (see Figure 3c) in order to encode another DOF and (b) making the size of the neuronal area *MotorField* (see Figure 3f) four times larger. The new weights were initialized randomly and the old weights were kept at their current values in order to preserve the previous knowledge acquired by the robot. The process is shown in Figure 6 and summarized in Table 1.

6 Developmental schedule

Development, in contrast to mere learning, implies on the one hand changes in the entire organism (not only

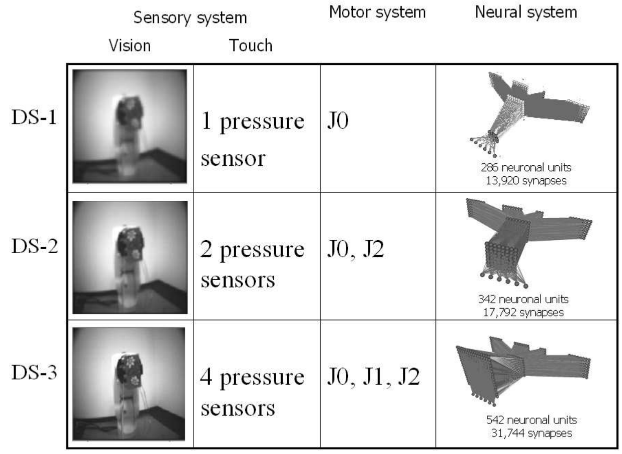


Figure 7. Configuration of the sensory, motor and neural components of the robot through the developmental approach. From top to bottom: DS-1 (immature state), DS-2 (intermediate state) and DS-3 (mature state).

the neural system) over time, and on the other hand a long-term perspective. The robot's movements were continuously shaped by the aforementioned learning mechanism, and "developmental" changes were triggered by the robot's internal performance evaluator (see definition of index "*P*" for the robot's task performance in Section 7). Such changes consisted in advancing the present developmental stage (DS-$_i$) to the next one. We defined a set of three different developmental stages (DS) in which the robot "grew up" as follows:

6.1 Developmental stage number 1 (DS-1)

At this stage, the sensory input to the robotic agent's neural structure consisted of a blurred, low resolution image (a 5x5 Gaussian kernel was applied to the original image captured by the cameras, see Figure 5(left)), and the activity of one pressure sensor. The neural network had 286 neuronal units and 13,920 synaptic connections, and controlled one single degree of freedom (i.e., joint $J0$). This developmental stage corresponds to the "*immature*" state of the robot. See Figure 7(DS-1).

6.2 Developmental stage number 2 (DS-2)

At this stage the robotic agent consisted of a medium level blurred image (a 3x3 Gaussian kernel was applied to the original image captured by the cameras, see Figure 5(center)), two pressure sensors, two DOF (i.e., joint $J0$ and $J2$), and the neural network had 342 neuronal units and 17,792 synaptic connections. This corresponds to the ”*intermediate*” state of the robot. See Figure 7(DS-2).

6.3 Developmental stage number 3 (DS-3)

At this stage the robotic agent consisted of the full high resolution image from the cameras (see Figure 5(right)), four pressure sensors, three DOF (i.e., $J0$, $J1$ and $J2$), and the neural network had 542 neuronal units and 31,744 synaptic connections. This corresponds to the ”*mature*” state of the robot. See Figure 7(DS-3).

6.4 Control setup

The control setup had the same configuration of the fully matured robotic agent at stage number 3.

The schedule on how the robot was changed over time was determined by the learning mechanism, every time that the robot was considered to have learned to solve the task its configuration was changed moving from one developmental stage to the next one. This was achieved as follows:

- The resolution of the camera image was increased.
- one or two pressure sensors were added.
- another degree of freedom came into operation and the size of the neuronal area: ”*ProprioceptiveField*” (see Figure 3c) was increased in 8 neuronal units.
- the size of the neuronal area: ”*MotorField*” (see Figure 3f) was increased by a factor of four, the new weights were initialized randomly and the old weights were kept at their current values in order to preserve the previous knowledge acquired by the robot.

Figure 7 presents a summary of the configuration of the robot at each developmental stage. The number of neuronal units in each neuronal area at each developmental stage can be found in Table 1.
Through this simulated development (from DS-1 to DS-3) the initial setup with reduced visual capabilities, noisy motor commands, low number of degrees of freedom, a few pressure sensors and a neural control architecture with a reduced number of neuronal units, was converted into an experimental setup with good vision, larger number of degrees of freedom, larger number of pressure sensors and a neural control architecture with a sufficient number of neuronal units.
At developmental stage number 3, the robotic agent reaches the same sensory, motor and neural configuration than the control setup. At this point, their performances could be compared to see whether the learning was affected or not by the developmental approach described above.

Table 1. Neural structure at each developmental stage

Neuronal Area	stage 1	stage 2	stage 3
RedColorField	64	64	64
RedMovementToRightField	64	64	64
ProprioceptiveField	8	16	24
RedMovementToLeftField	64	64	64
NeuronalField	64	64	64
MotorField	16	64	256
MotorActivites	6	6	6
Total neuronal units	286	342	542

7 Experiments and results

Figure 8 shows a typical experiment where the robot learned to move the object from the periphery of its visual field to the center of it by means of its robotic arm. To evaluate the change of the robot's task performance over time, at each time step i, we computed the cumulated distance covered by the center of the object projected onto one of the robot's cameras (x_i, y_i):

$$\hat{S} = \sum_{i=0}^{N-1} \sqrt{(}(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2) \tag{4}$$

Thus, (x_0, y_0) is the initial position of the object as perceived by the robot, and (x_N, y_N) = (x_c, y_c) is the center of the robot's visual field (assuming that the robot learns to perform the task).
The shortest possible path between (x_0, y_0) and (x_c, y_c) is defined as:

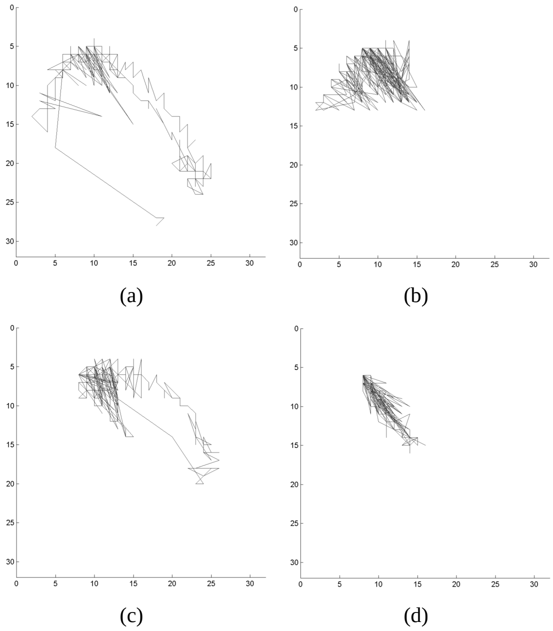
$$S = \sqrt{(}(x_0 - x_c)^2 + (y_0 - y_c)^2) \tag{5}$$

By using S and $\hat{S}$, we defined an index for the robot's task performance:

$$P = \frac{S}{\hat{S}} \tag{6}$$

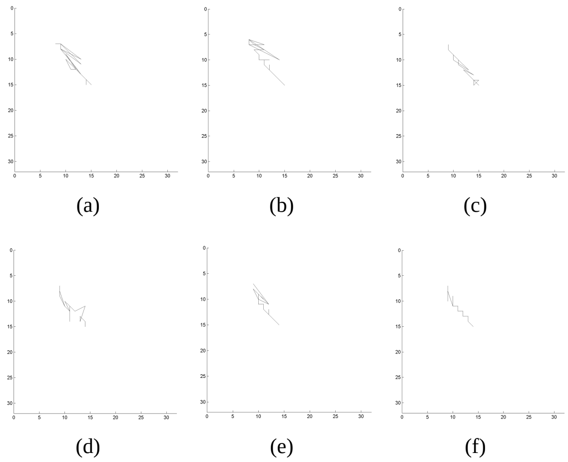
The closer P is to 1, the more straight the trajectory, and therefore the better the robot's behavioral performance.

Figure 9 shows how the robot's behavior improved over time for the last part of the experiment number 1 (see Figure 8 interval d.) and gives the performance measure over time.



(a) (b) (c) (d)

Figure 8. Experiment number 1. Learning to move a colored object from the upper left corner of the visual field to the center of it. Position of the center of the object in the visual field during the learning cycles in the interval (a) [1, 400]. (b) [401, 800]. (c) [801, 1200]. (d) [1201, 1602].



(a) (b) (c) (d) (e) (f)

Figure 9. Robot's internal performance evaluator *"P"* during the learning cycles in the interval (a) [1232, 1266], *P*=0.2898; (b) [1313, 1340], *P*=0.3574; (c) [1370, 1393], *P*=0.5114; (d) [1438, 1455], *P*=0.5402; (e) [1502, 1519], *P*=0.6569;(f) [1565, 1582], *P*=0.9176. (see Figure 8d).

A total of 15 experiments were performed with two types of robotic agents: one subjected to developmental changes (i.e., DS-1, then DS-2 and finally DS-3), and one fully developed since the onset (control setup). The results clearly show that the robotic agents that followed a developmental path took considerably less time to learn to perform the task. These robotic agents started with the configuration of the developmental stage number 1 and learned to solve the task during the learning cycle 483 ± 70 (where $\pm$ indicates the standard deviation), then they were converted to robotic agents with a configuration as described by the developmental stage number 2 which subsequently learned to solve the task around the learning cycle 1671 ± 102 and finally they become to be in the developmental stage number 3 (with the same configuration than the control setup) and solve the task around the learning cycle 4150 ± 149 (this is a cumulative value).
The control setup agents with full resolution camera images, four pressure sensor, three DOF (i.e., $J0$, $J1$ and $J2$), and a neural network with 542 neuronal units (randomly initialized synaptic connections) learned to solve the task around the learning cycle 7480 ± 105.
In other words, a reduction of about 44.5 percent in the number of learning cycles needed to solve the task can be observed in the case of robotic agents that followed a developmental approach when compared to the control setup agents.

8 Discussion and conclusions

We set out to investigate if the immaturity of sensory, motor, and neural system, which at first sight appears to be an inadequacy, might speed learning and task acquisition. In other words, we hypothesize that rather than being a problem, immaturity might effectively decrease or even eliminate excessive information and its potentially detrimental effects on learning performance.
This might be indeed the case as shown by the results presented in this paper. A system starting with low resolution sensors and low precision motor systems, whose resolution and precision are then gradually increased during development, learns faster than a system starting out with the full high resolution high precision system from scratch. For this particular case, by employing a developmental approach the learning was speeded up by 44.5 percent. To our knowledge this is the first time that this point is actually shown in a quantitative way.
There is a trade-off between finding a solution following a developmental approach and the potentially better solution, when starting out from the full high resolution high precision system from scratch.
Important is to keep in mind that the motor abilities should be increased gradually with the sensor abilities, since this significantly reduces the learning problem.

9 Future research

We will add proprioceptive information about the position of each motor of the active vision system and one possible task for the robot would be to not only bring the object to the center of the visual field, but also to normalize the size of the object in the camera image (i.e., a big object would be presented by the arm to the cameras further away than a smaller one) providing the robot with an ”Embodied concept of size”. In a future set of experiments we will put the developmental schedule under the control of an Artificial Evolutionary System.

Acknowledgments

Gabriel Gómez was supported by the grant NF-10-101827/1 of the Swiss National Science Foundation and the EU-Project ADAPT (IST-2001-37173). Max Lungarella was supported by the Special Coordination Fund for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science, and Technology of the Japanese government. Peter Eggenberger Hotz was sponsored by the EU-Project HYDRA (IST-2001-33060).

References

[1] N. Bernstein. *The coordination and regulation of movements.* Pergamon, Oxford, England., 1967.

[2] L. Berthouze and M. Lungarella. Motor skill acquisition under environmental perturbations: on the necessity of alternate freezing and freeing of degrees of freedom. *Adaptive Behavior (To appear)*, 2004.

[3] P. Eggenberger Hotz, G. Gómez, and R. Pfeifer. Evolving the morphology of a neural network for controlling a foveating retina - and its test on a real robot. In *Standish, R. K., Bedau, M. A., and Abbass, H. A., editors, Artificial Life VIII: Proceedings of the 8th International Conference on the Simulation and Synthesis of Living Systems, Sydney, Australia*, pages 243–251, 2002.

[4] J.L. Elman. Learning and development in neural networks: The importance of starting small. *Cognition*, 48:71–99, 1993.

[5] G. Gómez and P. Eggenberger Hotz. An evolved learning mechanism for teaching a robot to foveate. In *Sugisaka Masanori and Tanaka Hiroshi, editors (AROB 9): Proceedings of the 9th Int. Symp. on Artificial Life and Robotics, Beppu, Oita, Japan*, pages 655–658, 2004.

[6] G. Gómez and P. Eggenberger Hotz. Investigations on the robustness of an evolved learning mechanism for a robot arm. In *In Groen, F., Amato, N., Bonarini, A., Yoshida, E., and Krose, B., editors (IAS 8): Proceedings of the 8th International Conference on Intelligent Autonomous Systems, Amsterdam, The Netherlands*, pages 818–827, 2004.

[7] F. Iida. Biologically inspired visual odometer for navigation of a flying robot. *Robotics and Autonomous Systems*, 44:201–208, 2003.

[8] M. Lungarella and L. Berthouze. On the interplay between morphological, neural and environmental dynamics: a robotic case-study. *Adaptive Behavior*, 10:223–241, 2002.

[9] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini. Developmental robotics: a survey. *Connection Science 15*, 4:151–190, 2003.

[10] David Marr. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman and Company, 1982. MAR d 82:1 1.Ex.

[11] Y. Nagai, K. Hosoda, S. Morita, and M. Asada. A constructive model for the development of joint attention. *Connection Science 15*, 4:211–229, 2003.

[12] R. Pfeifer, F. Iida, and J. Bongard. New robotics: design principles for intelligent systems. *Artificial Life Journal (To appear)*, 2004.

[13] Rolf Pfeifer and Christian Scheier. *Understanding Intelligence*. MIT Press, 1999.

[14] A. Slater and S. Johnson. Visual sensory and perceptual abilities of the newborn: beyond the bloomig, buzzing confusion. In *Simion, F. and Butterworth, G., editors, The Development of Sensory, Motor and Cognitive Capabilities in Early Infancy: From Sensation to Cognition. Hove, Psychology Press*, pages 121–141, 1997.

[15] E. Thelen. Dynamics mechanism of change in early perceptuo-motor development. In *In McClelland, J. and Siegler, S., editors, Mechanims of cognitive development: Behavioral and neural perspectives. Proceedings of the 29th Carnegie Symposium on Cognition.*, 1999.

[16] G. Turkewitz and P.A. Kenny. Limitation on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. *Developmental Psychobiology*, 15:357–368, 1982.

A Theory of Developmental Architecture

Juyang Weng
Department of Computer Science and Engineering
Michigan State University
East Lansing, MI 48824 USA
weng@cse.msu.edu

Abstract

This paper presents a theory of developmental mental architecture for robots, motivated by neuroscience. Six types of architecture are presented, from the observation-driven Markov decision process as Type-1. From Type-1 to Type-6, the architecture progressively becomes more complete toward the necessary functions of autonomous mental development. Properties of each type are presented.

1. Introduction

A computational system can be specified at one of the *four levels of detail*: (1) constraint, (2) architecture, (3) algorithm, and (4) program, with increasing order of detail from one to the next. Studies in psychology often address issues at the constraint level while many engineering papers discuss systems at the algorithmic level. This paper deals with the *architecture* level. Mental architecture is a challenging and important subject, but there have been relatively few systematic (agent-wise) investigations. *The working of the mind is hard to understand without understanding autonomous mental development (AMD)*. Both are hard to understand without investigating mental architecture in a mathematically rigorous way.

Supervised and reinforcement learning, based on the Markov Decision Process (MDP) architecture (single- or multi-level), enables a robot to learn autonomously while the environment (including humans) provides labels [11] or rewards [7, 13]. However, the MDP architecture, as explained in the following sections, has fundamental limitations that prevent them to be effective for the developmental robots described in [19].

Several alternative general-purpose architectures have been proposed. Major remarkable ones include Soar proposed by Laird, Newell & Rosenbloom [9], ACT-R by Anderson [2], and the architecture by Albus [1]. Soar and ACT-R incorporated many useful concepts that are necessary for human intelligence. Albus' architecture outline is motivated by neural architecture. The subsumption architecture proposed by Brooks [3] is a biologically motivated architecture component well suited for what is now known as the behavior-based approach.

The architecture models discussed above do not directly address perception, such as vision and audition. Neisser [10] pointed out that any model of vision that is based on spatial computational parallelism alone is doomed to failure. He proposed a two-stage visual process which consists of a pre-attentive phase followed by an attentive phase. Feldman & Ballard [5] proposed a "100-step rule:" A biologically plausible algorithm for immediate vision (one that does not involve slower deliberate thinking) can require no more than 100 steps. John Tsotsos' study [14] on the complexity of *immediate vision* proposed a coarse architecture for a biologically motivated general purpose vision system (for immediate vision). All these architectures are nondevelopmental in the sense that the information processor is not generated through real-time interactions with the environment.

Recently, there has been an onset of efforts on computational studies of autonomous mental development (e.g., the workshop report in [20]). There is a lack of studies on the developmental mental architecture. This paper deals with this important issue. It does not describe developmental algorithms that use such architecture, but there has been experimental systems (e.g., [16]) with algorithm detail that support some architecture models described here. Studies of actual experimental systems cannot replace studies on architecture since the former does not provide properties of alternative architectures.

The history of studies on mental architecture has shown that this subject is hard to study and challenging to understand. This paper does not mean to solve all the problems and answer all the questions about this subject. It is a theoretical step toward the goal. In the following sections, I introduce a series of architectures, from simple to complex, along with the associated properties.

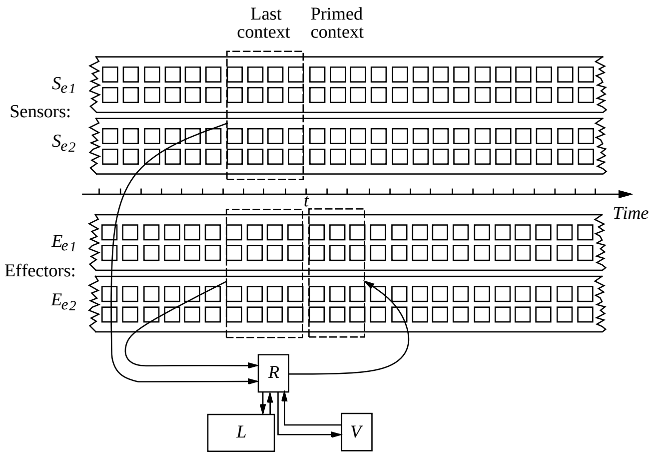


Figure 1. The Type-1 architecture of a multi-sensor multi-effector agent: Observation-driven Markov decision process. In the temporal sensory and effector streams, each square denotes a receptor (e.g., pixel) or motor and vertically aligned squares form a time frame. The Type-1 architecture takes the entire image frame without applying any mask. The block marked with L is a set of context states (prototypes), which are clusters of all observed context vectors $l(t)$.

2. Type-1: Observation-driven MDP

Definition 1 *The* internal environment *of an agent is the brain (or "the central nervous system") of the agent. The* external environment *consists of all the remaining parts of the world, including the agent's own body (excluding the brain).*

Definition 2 *An external sensor S_e and an internal sensor S_i are sensors that sense the external and internal environments, respectively. An external effector E_e and an internal effector E_i are effectors that act on the external and internal environments, respectively.*

Fig. 1 illustrates a multi-sensor multi-effector model of agent. The agent $A(t)$ operates at equally spaced discrete time instances $t = 0, 1, ...$ We assume that an image is produced at each time instance by the sensor, independent of the sensing modality, visual, auditory, touch, etc. Without loss of generality, we assume that the agent has two external sensors and two external effectors. Each external sensor S_{ei}, $i = 1, 2$, senses a random multi-dimensional sensory frame $x_e(t) = (x_{e1}(t), x_{e2}(t))$ at each time instance t and the sensed signal is fed into the agent. Each external effector E_{ei}, $i = 1, 2$, receives from the agent an effector frame $a_e(t) = (a_{e1}(t), a_{e2}(t))$ at each time instance t. Note that we change a variable of a vector to its subscript (e.g., change $x(t)$ to x_t) when it is convenient.

Definition 3 (MDP) *The Markov decision process (MDP) is as follows. Suppose $S = \{1, 2, ..., n\}$ is a set of n predefined symbolic states that is used to model a part of the world. The state s_t at time t is a random variable taking one of the values in S. Its prior probability distribution is $P(s_0)$. The action a_t is the action of the agent at time t. Let H_t be the random history from time $t = 0$ up to time $t - 1$:*

$$H_t = \{s_{t-1}, s_{t-2}, ..., s_0, a_{t-1}, a_{t-2}, ..., a_0\}.$$

If its conditional state transitional probability $P(s_t \mid H_t)$ satisfies

$$P(s_t \mid H_t) = P(s_t \mid l_t)$$

where l_t is the short last k frames of the history

$$l_t = \{s_{t-1}, s_{t-2}, ..., s_{t-k}, a_{t-1}, a_{t-2}, ..., a_{t-k}\},$$

we call it the k-th order MDP [7, 13].

In many applications, the state of the world is not directly observable by the agent, or observable but with noise.

Definition 4 (Partially Observable MDP) *If the state s_t of the world is not totally observable to the agent. Instead, there is an observation x_t at time t that depends on the state s_t by an observation probability $P(x_t \mid s_t)$, the process is called partially observable MDP or POMDP [7, 13] (or HMM [11]).*

In contrast, consider the following observation driven Markov Decision Process.

Definition 5 (Type-1) *Let $x_t \in \mathcal{X}$ and $p_t \in \mathcal{P}$ be the observations and outcome covariates (i.e., random vectors) at time t, respectively. Let H_t be the random vector of the entire history:*

$$H_t = \{x_t, x_{t-1}, ..., x_0, p_{t-1}, ..., p_0\}.$$

If its state transitional probability $P(s_t \mid H_t)$ satisfies

$$P(p_t \mid H_t) = P(p_t \mid l_t)$$

where l_k is the last k observations:

$$l_t = \{x_t, x_{t-1}, ..., x_{t-k}, p_{t-1}, ..., p_{t-k}\}$$

as shown in Fig. 1, we call the process as the k-th order Observation-driven MDP (ODMDP) [4]. The Type-1 mental architecture is a k-th order ODMDP.

In the developmental ODMDP, the random observations in l_t across time $t = 0, 1, ..., t$ are the source from which the agent automatically generates states in the form of clusters $l \in \mathcal{L}$, where $\mathcal{L}$ consists of all possible observations of the last contexts $\mathcal{L} = \{l_t \mid 0 \leq t\}$. The predicted consequence p_t consists of predicted action a_t and the predicted value v_t, $p_t = (a_t, v_t)$.

The following are the major differences between a POMDP (or HMM) and an ODMDP:

1. The POMDP is world-centered, where each state corresponds to an object or event of the world (e.g., a corner). The ODMDP is mind-centered (from sensors), where each state corresponds to an observation from the environment (e.g., a view of the corner with other background objects).

2. The states s_t of POMDP are hand-specified but the states of ODMDP can be automatically generated (developed) on-the-fly. With the POMDP, the mean of each state must be specified so that the initial estimates of the three probability distributions can be provided.

3. In the POMDP, there are two layers of probability: the state transition probability $P(s_t \mid x_t, s_{t-1})$ and the state observation probability $P(x_t \mid s_t)$, while the observation-driven MDP has only one layer of probability: $P(p_t \mid l_t)$, making a more efficient learning algorithm possible.

Higher cerebral cortices realize the regressor R. It is a great challenge to incrementally generate R from real-time experience (cortical development). We implemented the regressor R using the Incremental Hierarchical Discriminant Regression (IHDR) [18, 17]. Given any observed (last) context $l(t)$, the regressor R produces multiple consequences (primed contexts) $p_1(t), ..., p_k(t)$ having a high probability:

$$\{p_1(t), ..., p_k(t)\} = R(l(t)). \quad (1)$$

Thus, the regressor R is a mapping from the space of the last context $\mathcal{L}$ to the power set of $\mathcal{P}$:

$$R : \mathcal{L} \mapsto 2^{\mathcal{P}}. \quad (2)$$

R is developed incrementally through the real-time experience.

The value system $V(t)$ (called motivational system in neuroscience [8]) selects a desirable context from multiple primed ones:

$$V(R(l(t))) = V(\{p_1(t), p_2(t), ..., p_k(t)\}) = p_i(t) \quad (3)$$

where $1 \leq i \leq k$ and k varies according to experience. The value function selects the best consequence $p_i(t)$ that has the best value $v_i(t)$ in $p_i(t) = (a_i(t), v_i(t))$: $i = \arg\max\{v_1(t), v_2(t), ..., v_k(t)\}$. We applied the real-time Q-learning algorithm [15] to estimate the value v_i of each consequence $p_i(t)$, $i = 1, 2, ..., k$, by handling delayed rewards [16]. Therefore, the value system V is a mapping from the power set of $\mathcal{P}$ to the space of $\mathcal{P}$:

$$V : 2^{\mathcal{P}} \mapsto \mathcal{P}. \quad (4)$$

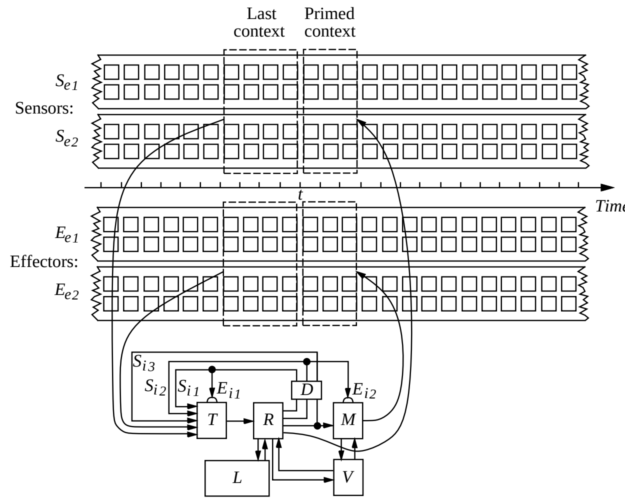


Figure 2. Progressive additions of architecture components from Type-2 to Type-5. Type-2: adding T and E_{i1}. Type-3: Adding M and E_{i2}. Type-4: Adding S_i and primed sensation. The block marked with D is a delay module, which introduces a unit-time delay. Type-5: Developmental T, R, M and V.

3. Type-2: Observation-driven Selective MDP

The Type-1 mental architecture is sensory nonselective in the sense that it is not able to actively select a subpart of relevant information from the sensory frame (intra-modal attention) or to attend a particular modality but not the other (inter-modal attention).

Given a d-dimensional input vector x, the attention can be modeled by an attention mask m, where m is a d-dimensional vector whose elements are either 0 or 1. Suppose that the input vector is $x = (x_1, x_2)$ and the mask is $m = (m_1, m_2)$. Then the corresponding attended input vector is $x' = x \otimes m = (x_1 m_1, x_2 m_2)$, where $\otimes$ denotes vector pointwise product. Not all the masks are admissible. For example, the set of admissible masks consists of circles with different radiuses ρ at different center positions (r_0, c_0) of the image frame. Then, the attention selection effector has three degrees of freedom: (r_0, c_0, ρ).

Definition 6 (Type-2) *The* Type-2 mental architecture *is a Type-1 architecture, with the addition of an attention selector*

$$T : \mathcal{Y} \times \mathcal{A}_i \mapsto \mathcal{L},$$

as shown in Fig. 2, where $\mathcal{Y}$ is the space of all possible pre-attention contexts $\mathcal{Y} = \{l(t) \mid 0 \leq t\}$, $\mathcal{A}_i$ is the space of all possible attention selections for T, and $\mathcal{L}$ is the space of attention-masked last contexts.

In order to investigate the properties of different architectures, we define a concept called *higher architecture*.

Definition 7 (Higher architecture) *Given a set D of tasks, we say that a developmental architecture A_2 is* higher *than another developmental architecture A_1, if given the same teaching environment E, the architecture A_2 requires statistically fewer teaching examples than A_1, expected over the environment E and over the tasks in D.*

As a convention, we regard environment as part of the specification of a task. For example, a task is more challenging if the environment of the task execution is uncontrolled.

Theorem 1 (Existence of higher architecture) *There is at least one class D of tasks and the associated teaching environment E in which the Type-2 architecture is higher than the Type-1 architecture.*

Due to the limit of space, here only a sketch of proof is provided. Construct a set D of tasks whose goal is to classify sensory information in set C (e.g., human bodies). Without loss of generality, assume that at any time, only one of the attention squares of sensor S_{e1} contains an element (e.g., human body) in C and the other window does not (e.g., natural background that is free of human bodies). The Type-2 architecture enables the teacher to teach the agent to pay attention to S_{e1}, but the Type-1 architecture cannot.

Clearly, the major rule of attention is to generalize better in new settings. E.g., in order to understand new settings where a familiar human face appears in new backgrounds, the agent architecture must enable the agent to pay attention to the human face and the background components separately. The higher architecture enables better generalization, but requires additional skills (e.g., attention selection). A higher architecture typically requires more sophisticated learning.

4. Type-3: Observation-driven Selective Rehearsable MDP

The Type-2 architecture does not have a motor mapping M. Therefore, it cannot autonomously rehearse an action sequence to evaluate its consequences without actually carrying out the action sequence. The rehearse is autonomous in that there is no pre-defined program segments that specify when and how to rehearse.

Definition 8 (Type-3) *The* Type-3 mental architecture *is a Type-2 mental architecture, with the addition of an action releaser M:*

$$M : \mathcal{P} \times \mathcal{A}_i \mapsto \mathcal{P},$$

as shown in Fig. 2, where $\mathcal{P}$ is the space of all possible predicted consequences, $\mathcal{A}_i$ is the space of all possible attention selections for M.

The action releaser M is a special case of the more general motor mapping (corresponding to the motor cortex) which also generates representation for frequently practiced action sequences (e.g., using the principal component analysis PCA or independent component analysis ICA) so that smooth action sequences can be generated.

With a traditional MDP with hand-designed states, it is possible to compute all the possible next states and perform planning. The Q-learning method uses the estimated action value $Q(s, a)$ of action a at state s to select the best action $a^* = \max_{a \in A} Q(s, a)$, from the set A of all the possible actions. This best next action a^* maximizes the expected rewards in the future. This kind of approach has two fundamental problems. First, the value system is rigid. No matter what value model is used (finite horizon, time discount model, etc.), the agent cannot autonomously change the way the value is determined [7, 13]. For example, if the time discount model is used, the agent is short-sighted. It prefers small rewards in the near further to faraway but important reward. Second, the agent is not able to learn to predict events (not just value) using the learned experience. For example, fed well and sleep well can be a reasonable goal for a human infant, but a human adult has a more sophisticated value system.

The Type-3 architecture does not suffer from these limitations. However, as long as the predicted action (e.g., drop a cup) is released, the effect that it causes to the external world will result (the broken cup). Can we design an architecture that enables the robot to *autonomously* "consider" and "plan" a significant amount of time ahead before it releases the action? The next type makes it possible.

5. Type-4: Observation-driven SASE MDP

Type-4 architecture is Self-Aware and Self-Effecting (SASE). The term "self" here means the brain, instead of the body of the agent.

Definition 9 (Awareness) *The awareness of a task b in an (internal and external) environment E by an agent A is the capability of the agent to (1) sense various context states s of task b from E and (2) recall the predicted multiple contexts (primed contexts) $p = R(s)$ using the regressor R.*

By definition, the agent must use its sensors, the entry point of its sensory architecture (the input of T), to sense the contexts. In the above definition for awareness, we consider a particular b and an environment E. This is because any awareness has a scope. A person who is aware of the boiling temperature of water in a domain (e.g., in a normal environment) may not necessarily be aware of the boiling temperature of water in another domain (e.g., lower in a low pressure environment). With the above definition, we are ready to address the issue of self-awareness and self-effecting.

Theorem 2 (Necessary conditions of self-awareness) *Suppose an agent is aware of its mental activities (sensations and actions) about a task b in an environment E. Then the following must be true: (1) It senses such mental activities using its sensors. (2) It feeds the sensed signal into its perceptual entry point just like that for external sensors.*

Proof: Point (1) is true because, according to the definition for the awareness of an object, the agent must sense the object using its sensors. Point (2) is true because the status of the object must be sensed and fed into the entry point for sensors for proper perception and recall of the primed contexts. □

Based on the last two theorems, let us examine the issue of self-awareness more closely. If an agent runs a learning algorithm (e.g., the Q-learning algorithm to be explained later) but it does not sense the voluntary decision process using its sensors which are linked to its entry point for sensors, the agent is not aware of its own algorithm. For the same reason, humans do not sense the way their primary cortex works and, therefore, normally they are not aware of their own earlier visual processing. However, the voluntary part of the mental decision process does require a conscious, willful decision.

The traditional model of agent has a fundamental flaw. The model is for an agent that perceives only the external environment and acts on the external environment. It does not sense its internal "brain's" activities. In other words, its internal decision process is neither a target of its own cognition nor a subject for the agent to explain when the agent is sufficiently mature.

The human brain allows the thinker to sense what he is thinking about without performing an overt action. For example, visual attention is a self-aware and self-effecting internal action (see, e.g., Kandel, et al.[8], pp. 396 - 403). Motivated by neuroscience, the mathematical model of the *self-aware and self-effecting* (SASE) agent, shown in Fig. 3, is defined as follows.

Definition 10 *A self-aware and self-effecting (SASE) agent has internal sensors S_i and internal effectors E_i for this internal (brain) environment, in addition to its external sensors S_e and external effectors E_e for its external environment (outside brain). The regressor R takes signals from S_i and S_e and generates internal and external actions for E_i and E_e, respectively.*

The major design principles for a SASE agent include:

1. A SASE agent must have a sensor for each of its voluntary external effectors, so that it can sense what each effector is doing. For example, the muscle spindles sense the tension of human muscles to tell the position of the arms.

2. A SASE agent must have internal effectors and internal sensors for its voluntary internal effectors. For example, it needs a set of internal attention effectors to select the most relevant part of sensory information for later processing, which eventually leads to voluntary actions for attention selection.

3. A SASE agent needs pre-motor areas, where it stores information about the control of the effectors, but the signal in the pre-motor area is not sent to the effectors unless an action release signal is issued. The effector signals in the pre-motor areas are also sensed by internal sensors so that the robot can "talk to itself" internally.

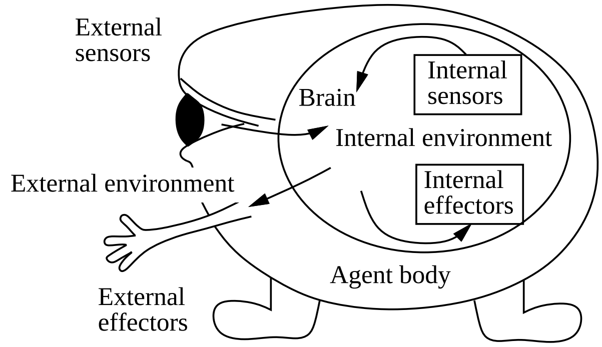


Figure 3. A self-aware self-effecting (SASE) agent. It interacts with not only the external environment but also its own internal (brain) environment: the representation of the brain itself.

It is important to note that not all the internal brain representations are sensed by the brain itself. Early processing actions are typically not sensed.

Definition 11 (Type-4) *The* Type-4 mental architecture *is a Type-3 mental architecture, but additionally, the internal voluntary decision is sensed by the internal sensors S_i and the sensed signals are fed into the entry point of sensors, i.e., the entry point of the attention selector T. In order to recall the effects of the voluntary actions, not only the expected reward value is estimated by the value system, but also the primed context which includes not only the primed action, but also the primed sensation.*

The architecture illustrated in Fig. 2 is a Type-4 architecture. Two voluntary internal actions are modeled by E_{i1} for attention selection, and by E_{i2} for action release. Both internal actions are sensed by the internal (virtual) sensors

S_{i1} and S_{i2}, respectively. The rehearsed external action (not released) is sensed by the virtual internal sensor S_{i3}.

The regressor R maps each attended context $l \in \mathcal{L}$ to a set of multiple primed contexts, from which the value system selects a single primed context $p \in \mathcal{P}$. In other words, the composite function of R followed by V gives a mapping: $V \circ R : \mathcal{L} \mapsto \mathcal{P}$. With a SASE agent, both external context (sensed by S_e) and internal context (sensed by S_i) are available in l.

Through a consecutive time series $t = 1, 2, ..., k$, the composite function $V \circ R$ performs a series of reasoning, represented by the regression sequence:

$$s = ((l_1, p_1), (l_2, p_2), ..., (l_k, p_k)) \tag{5}$$

where each regression pair (l_i, p_i) is an input-output pair of the composite function $V \circ R$, $p_i = V \circ R(l_i)$, $i = 1, 2, ..., k$. The link between two consecutive regression pairs can be realized by two paths, the external path and the internal path, denoted by e and i respectively. Symbolically, the reasoning process can be represented by the following composite reasoning sequence:

$$s = \left((l_1, p_1), \left[\begin{array}{c} e_1 \\ i_1 \end{array} \right], (l_2, p_2), ..., (l_k, p_k), \left[\begin{array}{c} e_k \\ i_k \end{array} \right] \right) \tag{6}$$

where

$$\left[\begin{array}{c} e_i \\ i_i \end{array} \right], i = 1, 2, ..., k$$

represents the parallel external and internal paths. Whether the result of external and internal paths are taken into account at any time t by the regressor depends on the attention selection in T.

Definition 12 (External and external reasoning process) *There are three types of reasoning processes, external, internal, and mixed, corresponding to the attention in which the attention module T attends to external, internal or both, respectively.*

From the above discussion, we have the following summary:

- Type-1 through Type-3 architectures allow the agent to perform external reasoning processes, but not internal reasoning as defined above.

- A Type-4 architecture is able to execute external, internal, and mixed reasoning processes.

Theorem 3 *The Type-4 architecture allows internal reasoning to realize the following kinds of learning (1) nonassociative learning, (2) classical conditioning, and (3) instrumental conditioning.*

Proof: First we prove the nonassociative learning[12]. The nonassociative learning occurs when the agent is exposed to stimulus because of the history of similar or dissimilar stimuli. Sensitization and habituation are two well known examples of nonassociative. In Eq. (6), the nonassociative learning can be accomplished by the link (l_i, p_i) realized by the composition of regression R and the value system: $p_i = V \circ R(l_i)$. The value system plays a central role. For example, the action (e.g., looking at another direction) that is predicted to generate more novel stimuli then alternative action (e.g., continue looking after repeated exposure to the similar stimuli), the former action is selected by the value system V from the alternative actions predicted by R.

Next, we prove the case for classical conditioning. In classical conditioning, a conditional stimulus CS (e.g., tone) is repeatedly paired with unconditional stimulus US (e.g., food) that elicits unconditional response UR (e.g., salivation). In this case, l_i = CS, l_{i+1} = US, and p_{i+1} = salivation, for all the time instances i where the event occurs. The Q-learning used by the value system V back-propagates repeatedly the primed action p_{i+1} through time i, so that l_i primes p_i = salivation even in the absence of l_{i+1}.

Finally, we establish the case for instrumental conditioning. When l_i stimulus is present, two actions a_1 and a_2 are predicted, $(a_1, a_2) = R(l_i)$. According to past experience, a_1 has a low value and a_2 has a higher value, using, e.g., Q-learning by the value system V. Thus, a_2 is selected by V. $\square$.

For a realization of the nonassociative learning, the classical conditioning, and the instrumental conditioning, using the Type-4 architecture, see Huang & Weng [6], Zhang & Weng [21], and Huang & Weng [6], respectively. The instrumental conditioning has been known as reinforcement learning in the machine learning community and has been very widely studied using the traditional MDP architecture [7] [13]. A major contribution here is that a *single architecture* realizes all three types of learning.

There are many more complex internal mental activities enabled by the architecture. As an example, we addresses a complex activity known as *autonomous planning*. Planning has been conducted extensively using the traditional MDP architecture, based on the Q-learning mechanism (i.e., time discounted value propagation) [15]. However, Q-learning based planning has a major drawback: It prefers immediate small rewards to future large rewards. One can program the planner in such a way so that only the final goal produces a reward and intermediate goals do not. However, such a task-specific setting is too inflexible for the general setting of mental development, where various rewards are generated from the real world at different stages and it is impossible for the programmer to write a different program for a different planning task (due to the task non-specific nature of autonomous mental development Weng et al. [19].)

Theorem 4 *The Type-4 architecture allows internal reasoning to realize autonomous planning.*

Proof. Autonomous planning requires first an accumulation of experiences so that alternative condition-action pairs are learned. Suppose that there are two plans according to the experiences: The execution path of the plan (a) is recalled as:

$$(l_1, p_{a,1}), (l_{a,2}, p_{a,2}), ..., (l_{a,i}, p_{a,i}))$$

and that of the plan (b) is recalled as:

$$(l_1, p_{b,1}), (l_{b,2}, p_{b,2}), ..., (l_{b,j}, p_{b,j})).$$

Both lead to a completion of the task. Both plans are recalled sequentially using only the internal path, i path instead of e in Eq. (6). Finally the value of $p_{a,i}$ is compared with that of $p_{b,j}$. The value system decides which value is better and so chooses the corresponding plan (a) or (b). The association of a to the primed action in $p_{a,1}$ and b with that in $p_{b,1}$ is represented by "talking to itself:" For example, the selected plan in $p_{a,i}$ as part of the last context in l, which primes the first action in $p_{a,1}$. The similar process takes place for plan (b). □

I expect that early demonstration of autonomous planning is possible in a restricted (simplified) natural setting. Anywhere any-time planning in uncontrolled natural settings is possible after a significant amount of "living experience." It is important to note that attention (intra-model and intermodal) at each sensorimotor path plays a powerful role of generalization for new and unexpected settings (see the task-transfer and action chaining in [21]).

One might think at this point that the internal process looks like "thinking." Yes, the Type-4 architecture is capable of thinking. However, the internal process defined here is not equivalent to autonomous thinking which is fundamental to human intelligence. A necessary piece for thinking is autonomous mental development, as explained below. The skills of thinking must be autonomously developed.

6. Type-5: Developmental observation-driven SASE MDP

Definition 13 (DOSASE MDP) *The developmental observation-driven SASE MDP (DOSASE MDP) has an architecture Type-4 or higher, that satisfies the following requirements:*

1. *During the programming time, the tasks that the agent will learn are unknown to the programmer.*
2. *The agent $A(t)$ starts to run at $t = 0$ under the guidance of its developmental program P_d. After the birth, the brain of the agent is not accessible to humans.*
3. *Human teachers can only affect the agent $A(t)$ as a part of its environment through its sensors and effectors recursively: At any time $t = 0, t = 1, ...,$ its observation vector at time t is the last context $l(t)$. The output from $A(t)$ at time t is its selected primed context $p(t) \in \mathcal{P}$. $A(t-1)$ is updated to $A(t)$, including T, R (and L), M, and V.*

In contrast with the traditional MDP, the DOSASE MDP $(A(t), P_d)$ is developmental in the sense that the developing program P_d does not require a given estimate of the *a priori* probability distribution $P(l)$ for all $l \in \mathcal{L}$, nor even a given set of states. Consequently, P_d does not require a given estimate for the state observation probability $P(l_t \mid l_{t-1})$ nor that for the state observation probability $P(x_t \mid l_t)$.

When the number of states is very large, it is practically sufficient to keep only track of the states that have a high probability, instead of estimating probability of all the states, which is too computationally expensive to reach the real-time speed. In HMM, this technique is called beam forming.

7. Type-6: Multi-level DOSASE MDP

The Type-5 architecture has only one sensorimotor level, although each mapping T, R, and M have multiple levels in their own internal structure. We call it a sensorimotor level because the pathway from T through R up to M corresponds to a pathway from sensory input to motor output. The primed context of such a level can be fed into another sensorimotor level for the following reasons:

1. Abstraction. While a low level is tightly linked to fine time steps, the higher levels become more "abstract," in the sense that the higher level clusters of context states are coarser in temporal granularity and grouped more according to actively attended events.
2. Self-generated context: Allow voluntarily generated motor actions to serve as context input to the higher level. Thus the agent is able to "talk to itself" at a more abstract level.
3. Enabling a higher degree of sensory integration. It is not practical to integrate all the receptors in a human body by a single attention selection module T, because otherwise, e.g., the attention is too complex.

Definition 14 (Type-6) *The Type-6 mental architecture is composed of several levels of Type-5 architecture. The primed contexts from a lower-level system is fed into the sensory input of the higher-level system.*

Fig. 4 illustrates the Type-6 architecture. The input to the attention selector T at level 2 includes the primed context

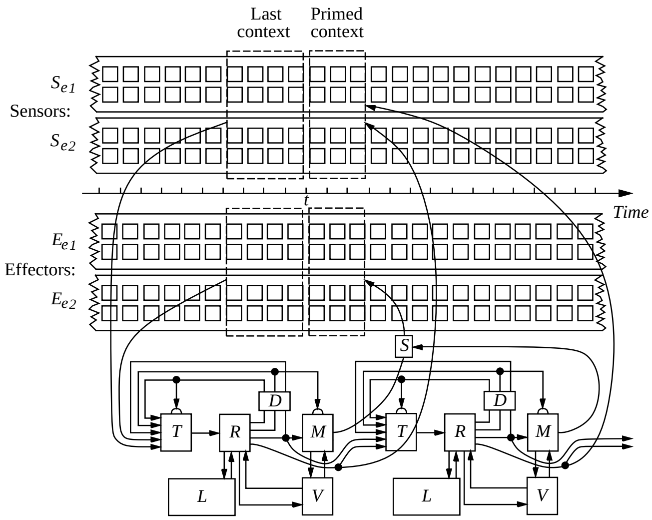


Figure 4. The Type-6 architecture.

$p(t) = (x_p(t), a_p(t))$ from level 1, where $x_p(t)$ and $a_p(t)$ are primed sensation and primed action, respectively. One or multiple levels can feed their primed contexts into the next higher level for sensory integration.

We have systematically introduced six types of architectures. Although the order at which new capabilities are added to the previous type is primarily a design choice, the order used here is motivated by a relatively large payoff in capability with a minimal addition of the architecture complexity.

8. Conclusions

The observation-driven MDP (Type-1), seems more suited for autonomous mental development than the traditional MDP. This paper provides Type-1, through Type-5 (DOSASE MDP), up to Type-6 (multilevel DOSASE MDP). A DOSASE MDP can perform nonassociative learning, classical conditioning, instrumental conditioning and planning. The realization of higher capabilities has yet to be demonstrated.

References

[1] J. S. Albus. Outline for a theory of intelligence. *IEEE Trans. Systems, Man and Cybernetics*, 21(3):473–509, May/June 1991.

[2] J. R. Anderson. *Rules of the Mind*. Lawrence Erlbaum, Mahwah, New Jersey, 1993.

[3] R. A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2(1):14–23, March 1986.

[4] D. R. Cox. Statistical analysis of time series: Some recent developments. *Scand. J. Statist.*, 8(2):93–115, 1981.

[5] J. A. Feldman and D. H. Ballard. Connectionist models and their properties. *Cognitive Science*, 6(3):205–254, 1982.

[6] X. Huang and J. Weng. Novelty and reinforcement learning in the value system of developmental robots. In *Proc. Second International Workshop on Epigenetic Robotics (EPIROB'02)*, pages 47–55, Edinburgh, Scotland, August 10 - 11 2002.

[7] L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[8] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, editors. *Principles of Neural Science*. McGraw-Hill, New York, 4th edition, 2000.

[9] J. E. Laird, A. Newell, and P. S. Rosenbloom. Soar: An architecture for general intelligence. *Artificial Intelligence*, 33:1–64, 1987.

[10] U. Neisser. *Cognitive Psychology*. Appleton-Century-Crofts, New York, 1967.

[11] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of IEEE*, 77(2):257–286, 1989.

[12] L. R. Squire and E. R. Kandel. *Memory: From Mind to Molecules*. Scientific American Library, New York, 1999.

[13] R. S. Sutton and A. Barto. *Reinforcement Learning*. The MIT Press, Cambridge, Massachusetts, 1998.

[14] J. K. Tsotsos. Analyzing vision at the complexity level. *Behavioral and Brain Sciences*, 13:423–469, 1990.

[15] C. Watkins. Q-learning. *Artificial Intelligence*, 8:55–67, 1992.

[16] J. Weng. Developmental robotics: Theory and experiments. *International Journal of Humanoid Robotics*, 1(2):199–235, 2004.

[17] J. Weng and W. Hwang. Online image classification using IHDR. *International Journal on Document Analysis and Recognition*, 5(2-3):118–125, 2002.

[18] J. Weng and W. S. Hwang. An incremental learning algorithm with automatically derived discriminating features. In *Proc. Asian Conference on Computer Vision*, pages 426 – 431, Taipei, Taiwan, Jan. 8-9 2000.

[19] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291(5504):599–600, 2001.

[20] J. Weng and I. Stockman. Autonomous mental development: Workshop on development and learning. *AI Magazine*, 23(2):95–98, 2002.

[21] Y. Zhang and J. Weng. Action chaining by a developmental robot with a value system. In *Proc. IEEE 2nd International Conference on Development and Learning (ICDL 2002)*, pages 53–60, MIT, Cambridge, Massachusetts, June 12-15 2002.

A Virtual Reality Platform for Modeling Cognitive Development

Hector Jasso
Department of Computer Science and Engineering
UC San Diego, USA
hjasso@cs.ucsd.edu

Jochen Triesch
Cognitive Science Department
UC San Diego, USA
triesch@cogsci.ucsd.edu

Abstract

We present a virtual reality platform for developing and evaluating embodied models of cognitive development. The platform facilitates structuring of the learning agent, of its visual environment, and of other virtual characters that interact with the learning agent. It allows to systematically study the role of the visual and social environment for the development of particular cognitive skills in a controlled fashion. We describe how it is currently being used for constructing an embodied model of the emergence of gaze following in infant-caregiver interactions and discuss the relative benefits of virtual vs. robotic modeling approaches.

1. Introduction

Recently, the field of cognitive science has been paying close attention to the fact that cognitive skills are unlikely to be fully specified genetically, but develop through interactions with the environment and caregivers. The importance of interactions with the physical and social environment for cognitive development has been stressed by connectionist [7] and dynamical systems [17] approaches.

Developmental schemes are also being proposed in the field of intelligent robotics [1, 3, 18]. Instead of building a fully working robot, a body capable of interacting with the environment is given general learning mechanisms that allow it to evaluate the results of its actions. It is then "set free" in the world to learn a task through repeated interactions with both the environment and a human supervisor.

Our motivation is to develop embodied models of cognitive development, that allow to systematically study the emergence of cognitive skills in naturalistic settings. We focus on visually mediated skills since vision is the dominant modality for humans. The kinds of cognitive skills whose development we would ultimately like to model range from gaze and point following and other shared attention skills over imitation of complex behaviors to language acquisition. Our hope is that embodied computational models will help to clarify the mechanisms underlying the emergence of cognitive skills and elucidate the role of intrinsic and environmental factors in this development.

In this paper, we present a platform for creating embodied computational models of the emergence of cognitive skills using computer-generated *virtual environments*. These virtual environments allow the semi-realistic rendering of arbitrary visual surroundings that make it easy to relate model simulations to experimental data gathered in various settings. Our platform facilitates structuring of the graphical environment and of any social agents in the model. Typically, a single developing infant and a single caregiver are modeled, but arbitrary physical and social settings are easily accommodated. To illustrate the features of our our platform, we show how it can be used to build an embodied model of the emergence of gaze following in infant-caregiver interactions. This effort is a component of a larger research project studying the emergence of shared attention skills within the MESA (Modeling the Emergence of Shared Attention) project at the University of California San Diego[1].

The remainder of the paper is organized as follows. Section 2 describes our modeling platform and the underlying software infrastructure. Section 3 shows how it is currently being used to build an embodied model of the emergence of gaze following in mother infant interactions. Finally, we discuss our work and the relative benefits of virtual vs.

[1] `http://mesa.ucsd.edu`

Figure 1. Left: various views of a virtual living room used to model the emergence of gaze following. From top left, clockwise: caregiver's view, birds eye view, lateral view, and infant's view. Right: Saliency maps generated by analyzing the infant's visual input (lower left image in left half of figure). Top row, left to right: red, green, blue. Bottom row, left to right: yellow, contrast, face position. Bars on left of each saliency map indicate the intrinsic reward of this feature and the current habituation level.

robotic modeling approaches in Section 4.

2. The Platform

2.1. Platform Overview

The platform allows the construction of semi-realistic models of arbitrary visual environments. A virtual room with furniture and objects can be set up easily to model, say, a testing room used in a controlled developmental psychology experiment, or a typical living room. These visual environments are populated with virtual characters. The behavior and learning mechanisms of all characters can be specified. Typically, a virtual character will have a vision system that receives images from a virtual camera placed inside the character's head. The simulated vision system will process these images and the resulting representation will drive the character's behavior [15]. Figure 1 shows an example setting.

An overview of the software structure is given in Figure 2. The central core of software, the "Simulation Environment," is responsible for simulating the learning agent (infant model) and its social and physical environment (caregiver model, objects, . . .). The Simulation Environment was programmed in C++ and will be described in more detail below. It interfaces with a number of 3rd party libraries for animating human characters (BDI DI-Guy), managing and rendering of the graphics (SGI OpenGL Performer), and visual processing of rendered images to simulate the agents' vision systems (OpenCV).

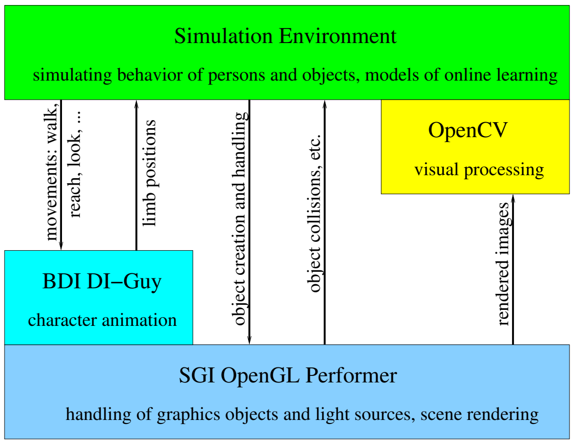


Figure 2. Overview of software structure.

The platform currently runs on a Dell Dimension 4600 desktop computer with a Pentium 4 processor running at 2.8GHz. The operating system is Linux. An NVidia GeForce video graphics accelerator speeds up the graphical simulations.

2.2. Third Party Software Libraries

OpenGL Performer. The Silicon Graphics *OpenGL Performer*[2] toolkit is used to create the graphical environment for running the experiments. OpenGL Performer is

[2] `http://www.sgi.com/products/software/performer/`

a programming interface built atop the industry standard *OpenGL* graphics library . It can import textured 3D objects in many formats, including OpenFlight (.flt extension) and 3D Studio Max (.3ds extension). OpenGL is a software interface for graphics hardware that allows the production of high-quality color images of 3D objects. It can be used to build geometric models, view them interactively in 3D, and perform operations like texture mapping and depth cueing. It can be used to manipulate lighting conditions, introduce fog, do motion blur, perform specular lighting, and other visual manipulations. It also provides virtual cameras that can be positioned at any location to view the simulated world.

DI-Guy. On top of OpenGL Performer, Boston Dynamics's *DI-Guy* libraries[3] provide lifelike human characters that can be created and readily inserted into the virtual world. They can be controlled using simple high-level commands such as "look at position (X, Y, Z)," or "reach for position (X, Y, Z) using the left arm," resulting in smooth and lifelike movements being generated automatically. The facial expression of characters can be queried and modified. DI-Guy provides access to the character's coordinates and link positions such as arm and leg segments, shoulders, hips, head, etc. More than 800 different functions for manipulating and querying the characters are available in all. Male and female characters of different ages are available, configurable with different appearances such as clothing style.

OpenCV. Querying the position of a character's head allows us to dynamically position a virtual camera at the same location, thus accessing the character's point of view. The images coming from the camera can be processed using Intel's *OpenCV* library[4] of optimized visual processing routines. OpenCV is an open-source, extendable software intended for real-time computer vision, and is useful for object tracking, segmentation, and recognition, face and gesture recognition, motion understanding, and mobile robotics. It provides routines for image processing such as contour processing, line and ellipse fitting, convex hull calculation, and calculation of various image statistics.

2.3. The Simulation Environment

The Simulation Environment comprises a number of classes to facilitate the creation and running of simulations. Following is a description of the most important ones.

The Object Class. The OBJECT class is used to create all inanimate objects (walls, furniture, toys, etc.) in the simulation. Instances of the OBJECT class are created by giving the name of the file containing the description of a 3D geometrically modeled object, a name to be used as a handle, a boolean variable stating whether the object should be allowed to move, and its initial scale. The file must be of a format readable by OpenGL Performer, such as 3D Studio Max (.3ds files) or OpenFlight (.flt files). When an OBJECT is created, it is attached to the Performer environment. There are methods for changing the position of the OBJECT, for rotating it, and changing its scale. Thus, it can easily be modeled that characters in the simulation can grasp and manipulate objects, if this is desired.

The Object Manager Class. The OBJECT MANAGER class holds an array of instances of the OBJECT class. The OBJECT MANAGER has methods for adding objects (which must be previously created) to the scene, removing them, and querying their visibility from a specific location. The latter function allows to assess if, e.g., an object is within the field of view of a character, or if the character is looking directly at an object.

The Person Class. The PERSON class is used to add any characters to the simulation. These may be rather complicated models of, say, a developing infant simulating its visual perception and learning processes, or they may be rather simplistic agents that behave according to simple scripts. To create an instance of the PERSON class, a DI-Guy character type must be specified, which determines the visual appearance of the person, along with a handle to the OpenGL Performer camera assigned to the character. The BRAIN type and VISION SYSTEM type (see below) must be specified. If the character's actions will result from a script, then a filename with the script must be given. For example, such a script may specify what the character is looking at at any given time. One BRAIN object and one VISION SYSTEM object are created, according to the parameters passed when creating the PERSON object. The PERSON object must be called periodically using the "update" method. This causes the link corresponding to the head of the character to be queried, and its coordinates to be passed to the virtual camera associated with the character. The image from the virtual camera in turn is passed to the character's VISION SYSTEM, if the character has any. The output of the VISION SYSTEM along with a handle to the DI-Guy character is passed to the BRAIN object, which will decide the next action to take and execute it in the DI-Guy character.

The Brain class. The BRAIN class specifies the actions to be taken by an instance of the PERSON class. The space of allowable actions is determined by the DI-Guy character type associated with the person. The simplest way of how a BRAIN object can control the actions of a PERSON is by following a script. In this case the PERSON will "play back" a pre-specified sequence of actions like a tape recorder. More interestingly, a BRAIN object can contain a simulation of the person's nervous system (at various levels of abstraction). The only constraint is that this simulation has to run in discrete time steps. For example, the BRAIN object may

[3] `http://www.bdi.com`
[4] `http://www.intel.com/research/mrl/research/opencv/`

instantiate a reinforcement learning agent [14] whose state information is derived from a perceptual process (see below) and whose action space is the space of allowable actions for this character. An "update" method is called every time step to do any perceptual processing, generate new actions, and possibly simulate experience dependent learning.

The actions used to control a character are fairly high-level commands such as "look to location (X,Y,Z)," "walk in direction Θ with speed v," or "reach for location (X,Y,Z) with the left arm," compared to direct specification of joint angles or torques. Thus, this simulation platform is not well suited for studying the development of such motor behaviors. Our focus is on the development of higher-level skills that use gaze shifts, reaches, etc. as building blocks. Thus, it is assumed that elementary behaviors such as looking and reaching have already developed and can be executed reliably in the age group of infants being modeled — an assumption that of course needs to be verified for the particular skills and ages under consideration. The positive aspect of this is that it allows to focus efforts on modeling the development of higher level cognitive processes without having to worry about such lower-level skills. This is in sharp contrast to robotic models of infant development, where invariably a significant portion of time is spent on implementing such lower level skills. In fact, skills like two-legged walking and running, or reaching and grasping are still full-blown research topics in their own right in the area of humanoid robotics.

The Vision System class. The VISION SYSTEM class specifies the processing to be done on the raw image corresponding to the person's point of view (as extracted from a virtual camera dynamically positioned inside the person's head). It is used to construct a representation of the visual scene that a BRAIN object can use to generate behavior. Thus, it will typically contain various computer vision algorithms and/or some more specific models of visual processing in human infants, depending on the primary goal of the model.

If desirable, the VISION SYSTEM class may also use so-called "oracle vision" to speed up the simulation. Since the simulation environment provides perfect knowledge about the state of all objects and characters in the simulation, it is sometimes neither necessary nor desirable to infer such knowledge from the rendered images through computer vision techniques, which can be difficult and time consuming. Instead, some property, say the identity of an object in the field of view, can simply be looked up in the internal representations maintained by the simulation environment — it functions as an oracle. This simplification is desirable if the visual processing (in this case object recognition) is not central to the developmental process under consideration, and if it can be assumed that it is sufficiently well developed prior to the developmental process being studied primarily. In contrast, in a robotic model of infant development, there is no "oracle" available, which means that all perceptual processes required for the cognitive skill under consideration have to be modeled explicitly. This is time-consuming and difficult.

Main Program and Control Flow. The main program is written in C++ using object-oriented programming. OpenGL Performer is first initialized, and a scene with a light source is created and positioned. A window to display the 3D world is initialized, and positioned on the screen. Virtual cameras are created and positioned in the world, for example as a birds eye view or a lateral view. Cameras corresponding to the characters are created but positioned dynamically as the characters move their heads. Each camera's field of view can be set (characters would usually have around a 90^o field of view), and can be configured to eliminate objects that are too close or too far. All cameras created are linked to the window that displays the 3D world. Environment settings such as fog, clouds, etc. can be specified. The DI-Guy platform is then initialized, and a scenario is created. The scenario holds information about all the characters, and must be used to create new characters. New instances of the PERSON class are created, and their activities are specified by periodically giving them new actions to perform. The level of graphical detail of the characters can be specified to either get fairly realistically looking characters or to speed up processing.

Statistics gathering. Throughout the session, statistics are gathered by querying the different libraries: DI-Guy calls can be used to extract the position of the different characters or the configuration of their joints. The OBJECT MANAGER can be used to query the position of objects and their visibility from the point of view of the different characters. In addition, the internal states of all characters' simulated nervous systems are perfectly known. This data or arbitrary subsets of it can easily be recorded on a frame by frame basis for later analysis. These statistics are useful for analyzing long-term runs, and allow to evaluate whether the desired behavior is being achieved and at what rate. We point out that every simulation is perfectly reproducible and can be re-run if additional statistics need to be collected.

3. A First Example: Gaze Following

The motivation for constructing the platform was to facilitate the development of embodied models of cognitive and social development. To illustrate how the platform can be used through a concrete example, we will outline how we are currently developing an embodied model of the emergence of gaze following [5]. Gaze following is the capacity to redirect visual attention to a target when it is the object of someone else's attention. Gaze following does not occur at birth, but instead develops during a child's first 18 months of life.

The model we are developing is aimed at testing and refining the *basic set hypothesis* [8], which states that the following conditions are sufficient for gaze following to develop in infants: a) a reward-driven general purpose learning mechanism, b) a structured environment where the caregiver often looks at objects or events that the infant will find rewarding to look at, c) innate or early defined preferences that result in the infant finding the caregiver's face pleasant to look at, and d) a habituation mechanism that causes visual reward to decay over time while looking at an object and to be restored when attention is directed to a different object. Recently, Carlson and Triesch [4] demonstrated with a very abstract and simplified computational model, how the basic set may lead to the emergence of gaze following and how plausible alterations of model parameters lead to deficits in gaze following reminiscent of developmental disorders such as autism or Williams syndrome.

In our current work, we want to investigate if the basic set hypothesis still holds for a more realistic situation, where learning takes place in a complex naturalistic environment. The platform is configured for an experimental setup consisting of a living room with furniture and a toy, all of them instantiations of the OBJECT class and built from 3D Studio Max objects. Two instantiations of the PERSON class are created, one for the caregiver and one for the baby. The caregiver and learning infant are placed facing each other. The caregiver instantiates a BRAIN object controlling its behavior. A single toy periodically changes location within a meter of the infant, and its position is fed to the caregiver's BRAIN. In a first version of the model, the caregiver's BRAIN will simply cause the character to look at the position of the interesting toy with fairly high probability (75%). No visual system is given to the caregiver.

The baby instantiates a VISUAL SYSTEM object that models a simple infant vision system. In particular, it evaluates the *saliency* of different portions of the visual field [9], it recognizes the caregiver's head, and it discriminates different head poses of the caregiver. Saliency computation is based on six different features, each habituating individually according to Stanley's model of habituation [13]. The feature maps (see Figure 1) are: red, green, blue and yellow color features based on a color opponency scheme [12], a contrast feature that acts as an edge detector by giving a high saliency to locations in the image where the intensity gradient is high, and finally a face detector feature that assigns a high saliency to the region of the caregiver's face, which is localized through orace vision. The saliency of the face can be varied depending on the pose of the caregiver's face with respect to the infant (infant sees frontal view vs. profile view of the caregiver). A similar scheme for visual saliency computation has been used by Breazeal [2] for a non-developing model of gaze following, using skin tone, color, and motion features.

Image Scale	Vision	Map Display	Animation
80×60	0.0226	0.0073	0.0476
160×120	0.0539	0.0092	0.0431
240×180	0.0980	0.0121	0.0522
320×240	0.1507	0.0113	0.0422
400×300	0.2257	0.0208	0.0507
480×360	0.3025	0.0276	0.0539

Table 1. Simulation times (sec.)

The infant's BRAIN object consists of a two-agent reinforcement learning system similar to that used in [4]. The first agent learns to decide when to simply look at the point of highest saliency (reflexive gaze shift) or whether to execute a planned gaze shift. The second agent learns to generate planned gaze shifts based on the caregiver's head pose. The infant should learn to direct gaze to the caregiver to maximize visual reward, and habituation will cause him/her to look elsewhere before looking back to the caregiver. With time, the infant learns to follow the caregiver's line of regard, which increases the infant's chance of seeing the interesting toy. However, the caregiver's gaze does not directly index the position of the object, but instead only specifies a direction with respect to the caregiver but not the distance from the caregiver. One goal of the current model is to better understand such spatial ambiguities and how infants learn to overcome them [11].

3.1. Platform Performance

To illustrate the performance of the platform given our current hardware, we made a number of measurements to establish the computational bottlenecks for this specific model. The time spent for each frame was divided into three separate measures for analysis: the time to calculate the feature maps (Vision), the time to display them (Map Display), and the time for the DI-Guy environment to calculate the next character positions and display them (Animation). Table 1 shows how the times vary with the resolution of the infant's vision system. As can be seen, most time is spent on simulating the infant's visual processing. Real time performance is achievable if the image resolution is not set too high.

4. Discussion

The platform presented here is particularly useful for modeling the development of *embodied* cognitive skills. In the case of the emergence of gaze following discussed above, it is suitable because the skill is about the inference

Property	Robotic Model	Virtual Model
physics	real	simplified or ignored
agent body	difficult to create	much easier to simulate
motor control	full motor control problem	substantially simplified
visual environment	realistic	simplified computer graphics
visual processing	full vision problem	can be simplified through oracle vision
social environment	real humans	real humans or simulated agents
real time requirements	yes	no, simulation can be slowed down or sped up
data collection	difficult	perfect knowledge of system state
reproducibility of experiments	difficult	perfect
ease-of-use	very difficult	easy
development costs	extremely high	very modest

Table 2. Robotic vs. virtual models of infant cognitive development.

of mental states from bodily configurations, such as head and eye position, which are realistically simulated in our platform.

4.1. Virtual vs. Robotic Models

Recently, there has been a surge of interest in building robotic models of cognitive development. Compared to the virtual modeling platform presented here, there are a number of important advantages and serious disadvantages of robotic models that we will discuss in the following. A summary of this discussion is given in Table 2.

Physics. The virtual simulation is only an approximation of real-world physics. The movements of the characters do not necessarily obey physical laws but are merely animated to "look realistic." For the inanimate objects, we currently do not simulate any physics at all. In a robotic model, the physics are real, of course. The justification of neglecting physics in the virtual model is that the cognitive skills we are most interested in are fairly high-level skills, i.e., we simply do not want to study behavior at the level of muscle activations, joint torques, and frictional forces, but at the level of primitive actions such as gaze shifts, reaches, etc., and their coordination into useful behaviors.

Agent body. In the virtual modeling platform, we can choose from a set of existing bodies for the agents. These bodies have a high number of degrees of freedom, comparable to that of the most advanced humanoid robots. Further, since physics is not an issue, we are not restricted by current limitations in robotic actuator technology. Our characters will readily run, crawl, and do many other things.

Motor control. Our interface to the agents in the model allows us to specify high-level commands (walk here, reach for that point, look at this object). The underlying motor control problems do not have to be addressed. In contrast, for a robotic model the full motor control problem needs to be solved, which represents a major challenge. Clearly, the platform should not be used to study the specifics of human motor control but it makes it much easier to focus on higher level skills. At the same time, perfect control over individual joint angles is possible, if desired.

Visual environment. The simulated computer graphics environment is of course vastly simpler than images taken by a robot in a real environment. For example, shadows and reflections are not rendered accurately, and the virtual characters are only coarse approximations of human appearance. Clearly, again, such a modeling platform should not be used to, say, study the specifics of human object recognition under lighting changes. The skills we are most interested in, however, use object recognition as a basic building block (e.g., the ability to distinguish different head poses of the caregiver with a certain accuracy). We believe that the details of the underlying mechanism are not crucial as long as the level of competence is accurately captured by the model.

Visual processing. In the virtual modeling platform we can vastly simplify perceptual processes through the use of oracle vision. In a robotic model, this is not possible and the perceptual capabilities required for some higher level cognitive skills may simply not have been achieved by contemporary computer vision methods.

Social environment. A robotic model can interact with a real social environment, i.e., one composed of real human beings. In our virtual modeling platform we could achieve this to some extent by using standard Virtual Reality interfaces such as head mounted displays in conjunction with motion tracking devices. In such a setup a real person would control a virtual person in the simulation, seeing what the virtual person is seeing through the head mounted display. However, the ability to experiment with vastly simplified agents as the social environment allows us to systematically study what aspects of the social environment, i.e., which behaviors of caregivers, are really crucial for the development of specific social skills [16]. This degree of control over the social environment cannot be achieved with human subjects. Also, the social agents may be programmed to exhibit behavior that replicates important statistics of caregiver behavior observed in real infant caregiver interactions. For example, Deák et al. are collecting such statistics from videos of infant-caregiver dyad interactions [6]. We are planning on developing caregiver models that closely replicate the observed behaviors.

Real time requirements. A robotic model must be able to operate in real time. This severely limits the complexity of the model. Perceptual processes in particular are notoriously time consuming to simulate. In the virtual model, we are not restricted to simulating in real time. Simulations may be slowed down or sped up arbitrarily. In addition, the availability of oracle vision allows to save precious computational resources.

Data collection. In the virtual model it is trivial to record data about every smallest detail of the model at any time. This is much harder to achieve in a robotic model interacting with real human caregivers. In particular, the exact behavior of the caregiver is inherently difficult to capture. Useful information about the caregiver behavior can be recovered by manually coding video records of the experiment, but this information is not available at the time of the experiment.

Reproducibility of experiments. Along similar lines, the virtual modeling platform allows perfect reproducibility of experiments. Every last pixel of the visual input to the learning agent can be recreated with fidelity. This is simply impossible in a robotic model.

Ease-of-use. Not having to deal with robotic hardware shortens development times, reduces maintenance efforts to a minimum, and makes it much easier to exchange model components with other researchers. Also, recreating the specific setup of a real-world behavioral experiment, only requires changing a configuration file specifying where walls and objects are, rather than prompting a renovation.

Development costs. Finally, robotic models are much more expensive. Most of the software components used in our platform (Linux OS, SGI OpenGL Performer, Intel OpenCV) are freely available to researchers. The lion share of the costs is the price of the BDI DI-Guy software.

All these benefits may make a virtual model the methodology of choice. Even if a robotic model is ultimately desirable, a virtual model may be used for rapid proto-typing. We see the use of virtual and robotic models as complementary. In fact, we are pursuing both methodologies at the same time in our lab [10].

4.2. Possible Extensions

There are several extensions to our platform that may be worth pursuing. First, we have only considered monocular vision. It is easy to incorporate binocular vision by simply placing two virtual cameras side by side inside a character's head. Foveation could also be added to the characters' vision systems. Second, in order to model language acquisition, a simulation of vocal systems and auditory systems of the characters could be added. Even in the context of non-verbal communication, a caregiver turning his head to identify the source of a noise may be a powerful training stimulus for the developing infant. Third, the platform is not restricted to modeling human development, but could be extended to model, say, the development of cognitive skills in a variety of non-human primates. To this end the appropriate graphical characters and their atomic behaviors would have to be designed. Fourth, on the technical side, it may be worth investigating in how far the simulation could be parallelized to run on a cluster of computers.

4.3. Conclusion

In conclusion, we have proposed a research platform for creating embodied virtual models of cognitive development. We have outlined how the platform may be used to model the emergence of gaze following in naturalistic infant-caregiver interactions. The virtual modeling platform has a number of important advantages compared to robotic modeling approaches. The relative benefits of virtual models over robotic models on the one hand or more abstract computational models on the other hand need to be evaluated on a case-by-case basis.

5. Acknowledgments

This work is part of the MESA project at UCSD. We acknowledge funding from the UC Davis MIND Institute[5] and the National Alliance for Autism Research[6]. Tone Milazzo developed the first version of the platform. Tone Milazzo and Jochen Triesch developed the initial object-oriented programming scheme. The following people contributed with valuable input: Gedeon Deák, Javier Movellan, Hyundo Kim, Boris Lau, and Christof Teuscher.

References

[1] M. Asada, K. F. MacDorman, H. Ishiguro, and Y. Kuniyoshi. Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Robotics and Autonomous Systems*, 37:185–193, 2001.

[5] `http://www.ucdmc.ucdavis.edu/mindinstitute/`
[6] `http://www.naar.org`

[2] C. Breazeal. *Designing Sociable Robots*. MIT Press, Cambridge, MA, USA, 2002.

[3] R. A. Brooks, C. Breazeal, R. Irie, C. C. Kemp, M. Marjanovic, B. Scassellatti, and M. M. Williamson. Alternative essences of intelligence. In *Proc. of the American Association of Artificial Intelligence*, pages 961–968, 1998.

[4] E. Carlson and J. Triesch. A computational model of the emergence of gaze following. In H. Bowman and C. Labiouse, editors, *Connectionist Models of Cognition and Perception II: Proceedings of the Eighth Neural Computation and Psychology Workshop, University of Kent, UK, 28–30 August 2003*, volume 15, pages 105–114. World Scientific, 2003.

[5] G. O. Deák, R. Flom, and A. D. Pick. Perceptual and motivational factors affecting joint visual attention in 12- and 18-month-olds. *Developmental Psychology*, 36:511–523, 2000.

[6] G. O. Deák, Y. Wakabayashi, and H. Jasso. Attention sharing in human infants from 5 to 10 months of age in naturalistic conditions. In *Proc. of the 3rd International Conference on Development and Learning (ICDL'04)*, La Jolla, California, USA, 2004.

[7] J. L. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett. *Rethinking Innateness*. MIT Press, Cambridge, MA, USA, 1996.

[8] I. Fasel, G. O. Deák, J. Triesch, and J. Movellan. Combining embodied models and empirical research for understanding the development of shared attention. In *Proc. of the 2nd International Conference on Development and Learning (ICDL'02)*, pages 21–27, Los Alamitos, California, USA, 2002.

[9] L. Itti and C. Koch. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10-12):1489–1506, 2000.

[10] H. Kim, G. York, G. Burton, E. Murphy-Chutorian, and J. Triesch. Design of an antropomorphic robot head for studying autonomous mental development and learning. In *Proc. of the IEEE 2004 International Conference on Robotics and Automation (ICRA 2004)*, New Orleans, LA, USA, 2004.

[11] B. Lau and J. Triesch. Learning gaze following in space: a computational model. Proceedings of the Third International Conference on Development and Learning (ICDL'04), La Jolla, California, October 20–22, 2004. (This volume.).

[12] T. W. Lee, T. Wachtler, and T. J. Sejnowski. Color opponency is an efficient representation of spectral properties in natural scenes. *Vision Research*, 42(17):2095–2103, 2002.

[13] J. C. Stanley. A computer simulation of a model of habituation. *Nature*, 261:146–148, 1976.

[14] R. S. Sutton and A. G. Barto. *Reinforcement Learning: an Introduction*. MIT Press, Cambridge, MA, USA, 1998.

[15] D. Terzopoulos, X. Tu, and R. Grezeszczuk. Artificial fishes: Autonomous locomotion, perception, behavior, and learning in a simulated physical world. *Artificial Life*, 1(4):327–351, 1994.

[16] C. Teuscher and J. Triesch. To care or not to care: Analyzing the caregiver in a computational gaze following framework. Proceedings of the Third International Conference on Development and Learning (ICDL'04), La Jolla, California, October 20–22, 2004. (This volume.).

[17] E. Thelen and L. Smith. *A Dynamic Systems Approach To the Development of Cognition and Action*. MIT Press, Cambridge, MA, USA, 1994.

[18] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291:599–600, 2001.

SPARSE REGRESSION VIA THE WINNER-TAKE-ALL NETWORKS

Nan Zhang, Shuqing Zeng and Juyang Weng

Michigan State University
Department of Computer Science and Engineering
East Lansing, MI, 48823
email:{nanzhang, zengshuq, weng}@cse.msu.edu

ABSTRACT

Sparse representation is a desirable property for machine learning architectures, because it improves the generalization capability of the learning system. In supervised learning, the goal is to derive a mapping based on the training samples. The generalization capability is accomplished by reducing the complexity of the model, which is characterized by the number of non-zero parameters. This problem is formalized as ℓ_0 norm minimization, which is called sparseness. It is known that ℓ_0 norm minimization is NP-hard. Thus, it is usually approximated by assuming a super-Gaussian priori and applying a MAP (maximum a posteriori) procedure. If the Laplace priori is used, this problem is boiled down to LASSO regression, a minimization of residual error with an ℓ_1 norm regularization term.

In this paper, we propose a new perspective to achieve sparseness via the winner-take-all principle for the linear kernel regression and classification task. We form the duality of the LASSO criteria, and transfer an ℓ_1 norm minimization to an ℓ_∞ norm maximization problem. Two solutions are proposed: it can be solved by standard quadratic programming with linear inequality constraints. We show that the number of parameters to be estimated in the ℓ_∞ normed space is half of the parameters in the solution suggested by David Gay for ℓ_1 normed space. Second, we introduce a novel winner-take-all neural network solution derived from gradient descending, which links the sparse representation and the competitive learning scheme. This scheme is a form of unsupervised learning in which each input pattern comes through learning, to be associated with the activity of one or at most a few neurons. However, the lateral interaction between neurons in the same layer is strictly preemptive in this model. This framework is applicable to a variety of problems, such as Independent Component Analysis (ICA), feature selection, and data clustering.

1. INTRODUCTION

The central problem of supervised learning or regression can be formulated as function approximation. In either case we have pairwise correspondence of samples $\mathbf{x}$ and $\mathbf{y}$ from two sample space $\mathbf{X}$ and $\mathbf{Y}$, and the task is to find a function $f(\cdot)$, such that $\mathbf{y} = f(\mathbf{x})$. More precisely, if the model of the function is chosen, the function can be written as $\mathbf{y} = f(\mathbf{x}, \beta)$, where β is the parameter vector of the model. For example, the linear kernel regression assumes such function is a linear combination of a set of basis functions, i.e. $\mathbf{y} = \sum_i \beta_i h_i(\mathbf{x}) = \mathbf{h}(\mathbf{x})^\top \beta$, where $\beta = [\beta_1, ...\beta_d]^\top \in \Re^d$, $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), ...h_d(\mathbf{x})]$ is a set of basis functions. $h_i(\mathbf{x}) = K(\mathbf{x}, \mathbf{x}_i)$, where $K(\cdot)$ is a certain symmetric kernel function.

Typically, it is assumed that the output variable $\mathbf{y}$ from the training set was contaminated by additive white Gaussian noise, *i.e.* $\mathbf{y}_i = f(\mathbf{x}_i, \beta) + w_i$, for $i = 1, ..., n$, where $[w_1, ..., w_n]$ is a set of i.i.d. white Gaussian random variable with variance σ^2. Thus, the conditional probability $p(\mathbf{y}|\beta)$ is Gaussian, *i.e.* $p(\mathbf{y}|\beta) = \mathbf{N}(\mathbf{y}|\mathbf{h}(\mathbf{x})^\top \beta, \sigma^2 \mathbf{I})$. We write $\mathbf{H} = \mathbf{h}(\mathbf{x})^\top$, where $\mathbf{H}$ is called design matrix.

Simply apply Maximum Likelihood Estimator (MLE), we get least square error estimation, $\hat{\beta} = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{y}$. Note there is not any preference on β, so its prior is a uniform distribution. With a zero-mean Gaussian prior for β with variance A, the estimation is turned into *maximum a posteriori* (MAP) process. The prior of β then becomes an ℓ_2-norm regularization term in the log-likelihood, where it will prefer a small β. When $A = \mu^2 I$, it is called *ridge regression* [1].

Other β prior can also be applied. If sparseness is preferred, then Laplacian prior can be adopted for β, *i.e.*

$$p(\beta|\alpha) = \left(\frac{\alpha}{2}\right)^k \exp(-\alpha \|\beta\|_1),$$

where α is a parameter of the Laplacian pdf, and $\|\mathbf{x}\|_i$, $i = 0, ...\infty$ is the so called ℓ_i-norm. Laplacian distribution features heavy tail and has a high concentration at near-zero area, which means that most of the β's components will be zero, and the probability of having a large value is relatively high, comparing to the Gaussian distribution with the same variance. Utilizing the same MAP process, the estimation of β is given by

$$\hat{\beta} = \arg\min_{\beta} \left\{ \|\mathbf{y} - \mathbf{H}\beta\|_2^2 + t\,\|\beta\|_1 \right\}, \qquad (1)$$

where $t = 2\sigma^2\alpha$ is a control parameter, which can favor either the squared error term or the ℓ_1-norm regularization term. This criterion is also known as LASSO [2]. It is worth noting here that due to the non-Gaussian prior, the MAP estimation is not equivalent to the Bayesian estimation as it is in the ridge regression. So the estimation is biased. To make a unbiased estimation, one needs to integrate in all β space, which is computationally prohibitive. However, if the posterior concentrates at certain point, then this biased estimation may only have a small variance from the unbiased estimation, which is desirable. This is one reason why a concentrated sparse prior is preferred. Some researchers introduced "hyper-parameter" to further steepen the prior, like in Relevance Vector Machine (RVM) [3] and Figueiredo's work [4].

Another reason that a sparse representation is desirable is because it improves the generalization of a learning system. For example in supervised learning, the goal is to infer a mapping based on the training samples. The generalization capability is accomplished by reducing the complexity of the model, which is characterized by the number of non-zero parameters. This problem is formalized as ℓ_0-norm minimization. It is known that ℓ_0-norm minimization is NP-hard [5]. However, it is established that the solution of the ℓ_1 problem is the same as the ℓ_0 problem if certain condition is satisfied. So, ℓ_1-norm problem is still an important issue.

In this paper, we proposed a method to solve the ℓ_1-norm version of the problem. We construct the dual problem of LASSO criterion as in Eq. (1) and use a gradient-based method to find the solution. This method is by no means claimed to be superior to the quadratic programming (QP) based method, however it opens a perspective to address the problem differently. We have shown that the proposed method actually applies the competitive learning principle. And this nature can also motivate other biologically plausible models for solving similar problems.

The reminder of paper is organized as follows: In Section 2 we formulate the dual problem of the LASSO, and propose a solution base on gradient descend. We reformulate the proposed algorithm in Section 2.5. The experiment results and comparison with existing sparse optimization algorithms is in Section 3. Section 4 provides conclusions.

2. METHOD

2.1. Duality

First, we will construct the dual problem of Eq. (1). Fig. 1 illustrates this problem. The Eq. (1) can be geometrically

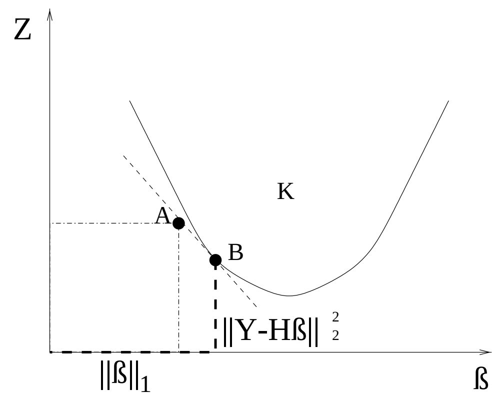


Fig. 1. Geometry explanation of duality. Point B is the closest point in K to origin.

explained as the minimum ℓ_1-norm between the origin and the convex set K. This distance, according to duality theory, is equal to the maximum ℓ_∞-norm distance between the origin and the plane that separate the origin and the convex set K. In Fig. 1, the parabola denotes the convex set K. Point B is the closest point to the origin, in terms of ℓ_1-norm. Actually the ℓ_1-norm of B consists of two parts, the quadratic form of Eq. (1) and $t\,\|\beta\|_1$ in Eq. (1), as illustrated in Fig. 1.

Point A is on the tangent plane of point B, and A is the closest point to the origin, in terms of ℓ_∞-norm. The duality theory establishes that the ℓ_∞ norm reaches its maximum value while the ℓ_1-norm reaches its minimum value. Thus, with a few manipulation, we get the following theorem to formulate this intuitive geometric explanation.

Theorem 2.1 *If we have*

$$\hat{\beta}' = \arg\max_{\beta} - \frac{\begin{bmatrix} \frac{\partial Z}{\partial t\beta} \\ -1 \end{bmatrix}^T \begin{bmatrix} t\beta \\ Z \end{bmatrix}}{\left\| \begin{bmatrix} \frac{\partial Z}{\partial t\beta} \\ -1 \end{bmatrix} \right\|_\infty}$$

where $Z = \|\mathbf{y} - \mathbf{H}\beta\|_2^2$, then $\hat{\beta}' = \hat{\beta}$, and $\hat{\beta}$ is the same as in Eq. (1).

The generalized proof can be found in [6].

2.2. Derivations

Let $Z = \|\mathbf{y} - \mathbf{H}\beta\|_2^2 = \beta^\top\mathbf{H}^\top\mathbf{H}\beta - 2\mathbf{y}^\top\mathbf{H}\beta + \mathbf{y}^\top\mathbf{y}$, and denote

$$J(\beta) = \begin{bmatrix} \frac{\partial Z}{\partial t\beta} \\ -1 \end{bmatrix}^\top = \begin{bmatrix} 2/t(\mathbf{H}^\top\mathbf{H}\beta - \mathbf{H}^\top\mathbf{y}) \\ -1 \end{bmatrix}^\top \qquad (2)$$

Now we can formulate the dual problem of the original LASSO criterion,

$$E(\beta) = -\frac{J(\beta)\begin{bmatrix} t\beta \\ Z \end{bmatrix}}{\|J(\beta)\|_\infty}, \tag{3}$$

Therefore,

$$\frac{\partial Z}{\partial t\beta} = \frac{2}{t}(\mathbf{H}^\top \mathbf{H}\beta - \mathbf{H}^\top \mathbf{y}). \tag{4}$$

$$\frac{\partial J}{\partial \beta} = \begin{bmatrix} & 0 \\ \frac{2}{t}\mathbf{H}^\top \mathbf{H} & \vdots \\ & 0 \end{bmatrix}. \tag{5}$$

Letting $\mathbf{C} = \mathbf{H}^\top \mathbf{H} = [\mathbf{c}_1, ..., \mathbf{c}_n]$, and plugging with Eq. (5) and Eq. (4), we get

$$\frac{\partial E}{\partial \beta} = \frac{2\mathbf{C}\beta}{\|J(\beta)\|_\infty} - \frac{\left[\beta^\top \mathbf{C}\beta - \mathbf{y}^\top \mathbf{y}\right]}{\|J(\beta)\|_\infty^2} \cdot \frac{\partial \|J(\beta)\|_\infty}{\partial \beta}.$$

Now we proceed to compute the partial derivative of $\|J(\beta)\|_\infty$ w.r.t. β.

$$\|J(\beta)\|_\infty = \max_i \left\{ \left| \frac{2}{t}\left(\beta^\top \mathbf{c}_i - \mathbf{y}^\top \mathbf{h}_i\right) \right|, 1 \right\}$$
$$= \sqrt{\max_i \left\{ \left| \frac{2}{t}\left(\beta^\top \mathbf{c}_i - \mathbf{y}^\top \mathbf{h}_i\right) \right|^2, 1 \right\}}$$

So,

$$\frac{\partial \|J(\beta)\|_\infty}{\partial \beta} = \frac{1}{2}\|J(\beta)\|_\infty^{-1} \cdot \frac{\partial}{\partial \beta} \max_i \left\{ \left| \frac{2}{t}\left(\beta^\top \mathbf{c}_i - \mathbf{y}^\top \mathbf{h}_i\right) \right|^2, 1 \right\}. \tag{6}$$

The last partial derivative term in Eq. (6) needs some special treatment. The difficulty of the analysis lies in the discontinuity caused by the maximum function. This problem can be circumvented by the use of the following equality. Let $\{a_i\}$ be a set of positive real scalars; then it generally holds that

$$\max_i \{a_i\} \equiv \lim_{r \to \infty} \left[\sum_i a_i^r\right]^{\frac{1}{r}}.$$

This is just another identity of the ℓ_∞-norm, which is differentiable. We have not yet get the strict derivation of this method. In fact, a similar technique can be seen in [7], in which Kohonen derived the vector quantization (VQ) algorithm base on this idea. In [8], a competitive learning algorithm has been derived from a maximum criteria function. Based on aforementioned observation, we conjecture that the order of the partial derivative and the max function can be exchanged. This leads to the following updating rules.

$$\frac{\partial}{\partial \beta} \max_i \left\{ \left| \frac{2}{t}\left(\beta^\top \mathbf{c}_i - \mathbf{y}^\top \mathbf{h}_i\right) \right|^2, 1 \right\} =$$
$$\begin{cases} \frac{8}{t}\left(\beta^\top \mathbf{c}_m - \mathbf{y}^\top \mathbf{h}_m\right) \cdot \mathbf{c}_m, \text{if } \frac{4}{t^2}\left(\beta^\top \mathbf{c}_m - \mathbf{y}^\top \mathbf{h}_m\right)^2 > 1 \\ [0, 0...0]^\top, \text{otherwise.} \end{cases}, \tag{7}$$

where

$$m = \arg\max_j \left(\left| \beta^\top \mathbf{c}_j - \mathbf{y}^\top \mathbf{h}_j \right| \right). \tag{8}$$

Rearrange these equations, we have the final updating rules,

$$\nabla \beta \propto \begin{cases} \frac{2\mathbf{C}\beta}{\|J(\beta)\|_\infty} - \frac{4\left[\beta^\top \mathbf{C}\beta - \mathbf{y}^\top \mathbf{y}\right]}{t\|J(\beta)\|_\infty^3} \cdot \left(\beta^\top \mathbf{c}_m - \mathbf{y}^\top \mathbf{h}_m\right)\mathbf{c}_m \\ \frac{2\mathbf{C}\beta}{\|J(\beta)\|_\infty} \end{cases} \tag{9}$$

The switching condition of these two updating rules is the same as that in Eq. (7)

Remarks: Eq. (8) defines a competitive learning process. The algorithm will find the winner and use the updating rule with the winners value.

The aforementioned method will search the optimal value of β to minimize the criteria function in Eq. (1). With the ℓ_1-norm constrain, some of the components β_i of vector β will gradually decay, however due to the nature of the gradient method, they will not be exact zero. So an additional step that set β_i that is closed to zero to zero will help the convergence of the algorithm.

The choice of learning rate in any gradient algorithm is important. Here we used a dynamic learning rate, called amnesic average. Suppose there are two statistical estimators Γ_1 and Γ_2 for estimating parameter θ. If $E\|\Gamma_1 - \theta\|^2 < E\|\Gamma_2 - \theta\|^2$, Γ_1 is said to be more statistically efficient than Γ_2.

We consider $(\mathbf{w}_i^\top \mathbf{x})\mathbf{x}$ with $\|\mathbf{w}_i\| = 1$ as an "observation." The goal is to get the mean of this observation, while $\mathbf{w}_i$ is estimated incrementally. It is known that for many distributions, the sample mean is the most efficient estimator for the mean of the random variable. When the distribution is unknown, the sample mean is the best linear estimator, which results in the minimum error variance. For many distributions, the sample mean reaches of approaches the Cramér-Rao bound (CRB).

Then an efficient estimator is one that has the least variance from the real parameter $\mathbf{W}$, and its variance is bounded below by the CRB. Thus, we estimate an independent component vectors $\mathbf{w}_i$ by the sample mean of the observation $(\mathbf{w}_i^\top \mathbf{x})\mathbf{x}$.

The sample mean uses a batch method. For incremental estimation, during which $\mathbf{W}$ is continuously improved, we use what is called an amnesic mean [9].

$$\bar{x}^{(n)} = w_1(n)\,\bar{x}^{(n-1)} + w_2(n)\,x_n, \tag{10}$$

where $\bar{x}^{(n)}$ is the mean at the n-th iteration, x_n is the n-th sample, and $w_1(n)$ and $w_2(n)$ are defined by

$$w_1(n) = \frac{n-1-\mu(n)}{n}, \tag{11}$$

and

$$w_2(n) = \frac{1+\mu(n)}{n}. \tag{12}$$

$\mu(n)$ is a non-negative small function that discounts old estimate and gives more weight to the new observation $\mathbf{x}_n$ at time n. When $\mu(n) \equiv 0$, $\bar{\mathbf{x}}^{(n)}$ is exactly the sample mean.

The algorithm is guided by the statistical efficiency, but it is not absolutely the most efficient one, because

1. the true distribution of the observation is unknown;
2. the distribution changes with $\mathbf{W}$ being incrementally estimated and therefore;
3. amnesic average is used to gradually discount "old" observations, which reduces the statistical efficiency moderately.

2.3. Quadratic Programming

Quadratic programming (QP) is used for finding the solution of LASSO regression. The method suggested by David Gay [2] transforms the LASSO to a QP problem with $2n$ variables and $(2n+1)$ constraints, where n denotes the number of the kernels. In this subsection, we are proposing the dual-QP algorithm, which converts the original problem into a problem with fewer variables (n) and $2n$ constraints.

The dual problem of LASSO regression is described in Eq. (3) and we recapitulate it as finding $\hat{\beta}$, s.t.

$$\hat{\beta} = \arg\max_{\beta} -\frac{J(\beta)\begin{bmatrix} t\beta \\ Z \end{bmatrix}}{\|J(\beta)\|_\infty} \tag{13}$$

This optimization problem is equivalent to

$$\hat{\beta} = \arg\max_{\beta} -J(\beta)\begin{bmatrix} t\beta \\ Z \end{bmatrix}$$

$$= \arg\min_{\beta} -[t^2\beta^\top(\beta\mathbf{H}^\top\beta\mathbf{H})\beta - \mathbf{y}^\top\mathbf{y}]$$

s.t.

$$\|J(\beta)\|_\infty = 1 \tag{14}$$

The constraint Eq. (14) can be written as

$$\max(\|\frac{2}{t}\mathbf{H}^\top\mathbf{H}\beta - \mathbf{H}^\top\mathbf{y}\|_\infty, 1) = 1$$

and is simplified to be

$$\|\frac{2}{t}\mathbf{H}^\top\mathbf{H}\beta - \mathbf{H}^\top\mathbf{y}\|_\infty \le 1 \tag{15}$$

Written in component form, Eq. (15) becomes

$$\max_i \mid \frac{2}{t}\mathbf{c}_i^\top\beta - \alpha_i \mid \le 1, \text{ for } i = 1, ..., n$$

where

$$\mathbf{H}^\top\mathbf{H} = \begin{bmatrix} \mathbf{c}_1^\top \\ \mathbf{c}_2^\top \\ \vdots \\ \mathbf{c}_n^\top \end{bmatrix} \text{ and } 2\mathbf{H}^\top\mathbf{y} = \begin{bmatrix} \alpha_1^\top \\ \alpha_2^\top \\ \vdots \\ \alpha_n^\top \end{bmatrix}$$

The inequality can be expressed in

$$-1 \le \frac{2}{t}\mathbf{c}_i^\top\beta - \alpha_i \le 1$$

We write the inequality in vector form

$$2\mathbf{H}^\top\mathbf{y} - 1 \le 2\mathbf{H}^\top\mathbf{H}\beta/t \le 1 + 2\mathbf{H}^\top\mathbf{y}$$

Thus, the equivalent QP problem becomes

$$\hat{\beta} = \arg\min[\beta^\top\mathbf{H}^\top\mathbf{H}\beta]$$

s.t.,

$$-2\mathbf{H}^\top\mathbf{H}\beta/t \le 1 - 2\mathbf{H}^\top\mathbf{y}$$
$$2\mathbf{H}^\top\mathbf{H}\beta/t \le 1 + 2\mathbf{H}^\top\mathbf{y}$$

It is worthy noting that the above QP has n variables and $2n$ constraints.

2.4. Filter Development

Let us reconsider the original linear kernel regression problem in a special case. Suppose the design matrix $\mathbf{H}$ is orthogonal n by n matrix, and there is no additive noise. So, $\beta = \mathbf{H}^\top\mathbf{y}$. Also we are not considering optimization with respect to β, but instead, with respect to the design matrix $\mathbf{H}$. Then, we get the criteria function adapted from Eq. (1)

$$\hat{\mathbf{H}} = \arg\min_{\mathbf{H}}\{\left\|\mathbf{H}^\top\mathbf{y}\right\|_1\}.$$

Using the same gradient technique described in section 2.2, we can get the updating rule of design matrix $\mathbf{H}$,

$$\nabla\mathbf{h}_j \propto 2\delta_{cj}(\mathbf{h}_c^T\mathbf{y})\mathbf{y},$$

where $c = \arg\max_i \left\{\frac{|\mathbf{h}_i^\top \cdot \mathbf{y}|}{\|\mathbf{h}_i\|}\right\}$, and δ_{cj} is the Kronecker delta: $\delta_{cj} = \{1 \text{ if } c = j, 0 \text{ otherwise}\}$. This is precisely a "winner-take-all" algorithm. Only the winning kernels (neurons) will get updated. For more details about this method, see [10].

2.5. Gradient Sparseness Optimization Algorithm

Algorithm 1 Gradient Sparseness Optimization Algorithm

Preprocessing: Get the training set $(\mathbf{x}_i, \mathbf{y}_i)$, $i = 0, 1, 2, ..., n$. For each i, $\mathbf{y}_i = \mathbf{y}_i - \bar{\mathbf{y}}$, where $\bar{\mathbf{y}} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{y}_i$.

$\beta_i = 0$, $i = 0, 1, 2, ..., k$.

Initialize the design matrix $\mathbf{H}$, and compute $\mathbf{C} = \mathbf{H}^\top \mathbf{H}$.

repeat

  $k = 1$.

  Compute $m = \arg\max_j \left(\left|\beta^\top \mathbf{c}_j - \mathbf{y}^\top \mathbf{h}_j\right|\right)$.

  if $\frac{4}{t^2}\left(\beta^\top \mathbf{c}_m - \mathbf{y}^\top \mathbf{h}_m\right)^2 > 1$ **then**

    Update β with:

$$\beta_{\text{new}} = w_1(k)\beta_{\text{old}} + w_2(k)[\frac{2\mathbf{C}\beta}{\|J(\beta)\|_\infty} - \frac{4\left[\beta^T \mathbf{C}\beta - \mathbf{y}^T\mathbf{y}\right]}{t\,\|J(\beta)\|_\infty^3} \cdot \left(\beta^T \mathbf{c}_m - \mathbf{y}^T \mathbf{h}_m\right)\mathbf{c}_m],$$

    where w_1 and w_2 is the same as in Eq. (11) and Eq. (12).

  else

    Update β with:

$$\beta_{\text{new}} = w_1(k)\beta_{\text{old}} + w_2(k)\frac{2\mathbf{C}\beta}{\|J(\beta)\|_\infty}$$

  end if

  $k = k + 1$.

until Objective function in Eq. (3) reaches the target value, or $\Delta\beta$ is less than some threshold.

We summarize the training procedure in Algorithm 1. Each iteration of this algorithm is computational efficient, because it only involves matrix multiplication and maximum function. The time complexity of each iteration is $O(kn)$, where k is number of basis vector and m is the dimension of each basis vector.

3. EXPERIMENTS

3.1. Kernel Regression

Our first experiment illustrates the performance of the proposed algorithm in kernel regression. The regression model is

$$\mathbf{y} = f(\mathbf{x}, \beta) = \beta_0 + \sum_{i=1}^{k} \beta_i K(\mathbf{x}, \mathbf{x}_i),$$

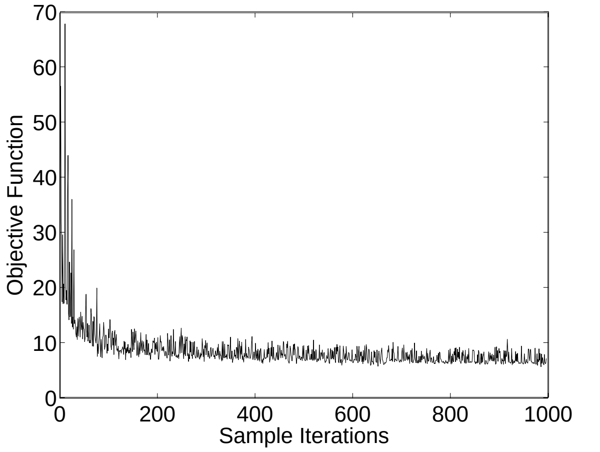


Fig. 3. Objective function vs. number of iterations. We utilize a dynamic learning rate mechanism called Amnesic Average. Interested reader can refer to [8] for more details.

where $K(\mathbf{x}, \mathbf{x}_i) = exp\{-\frac{\|\mathbf{x}-\mathbf{x}_i\|^2}{2\sigma^2}\}$ is the kernel function, $\mathbf{x}_i$ and σ are the kernel parameters. The function to be approximated is $1-d$ sinc function $y = \sin(x)/x$. We randomly collected 150 samples and added Gaussian noise with variance 0.01. The first row in Fig. 2 shows the fitting results of proposed method, ridge regression and the proposed dual-QP LASSO regression algorithm, respectively. The dots are samples with noise, and the dashed lines are the ground truth sinc function. Solid lines show the approximation results. The circled dots correspond to the kernels with nonzero weight, a.k.a the "supporting kernels". In the second row, we use the bar figure to display the weights of those kernels. As it clearly indicated, both proposed gradient sparseness method and dual-QP LASSO achieve sparseness. The ℓ_∞-norms are also marked on these figures. The proposed method in our testing performs better than LASSO regression.

Fig. 3 shows the convergence procedure. The proposed method takes about 200 iterations to converge. Fig. 4 illustrate how the control parameter $t = 2\sigma^2\alpha$ affects the mean square error and model sparseness. We conducted 20 tests with t ranging from 0.3 to 1.5. As indicated in the figure, greater t makes the model fit well but increases its complexity, and vice versa.

3.2. Classification

The experiment addressed the kernel-based classifier for two-class problems: A special case of the regression problem with $y \in \{+, -\}$. The classifier is formulated as the following two functions:

$$p(+ \mid \mathbf{x}) = \psi(\mathbf{H}\beta_+)$$
$$p(- \mid \mathbf{x}) = \psi(\mathbf{H}\beta_-)$$

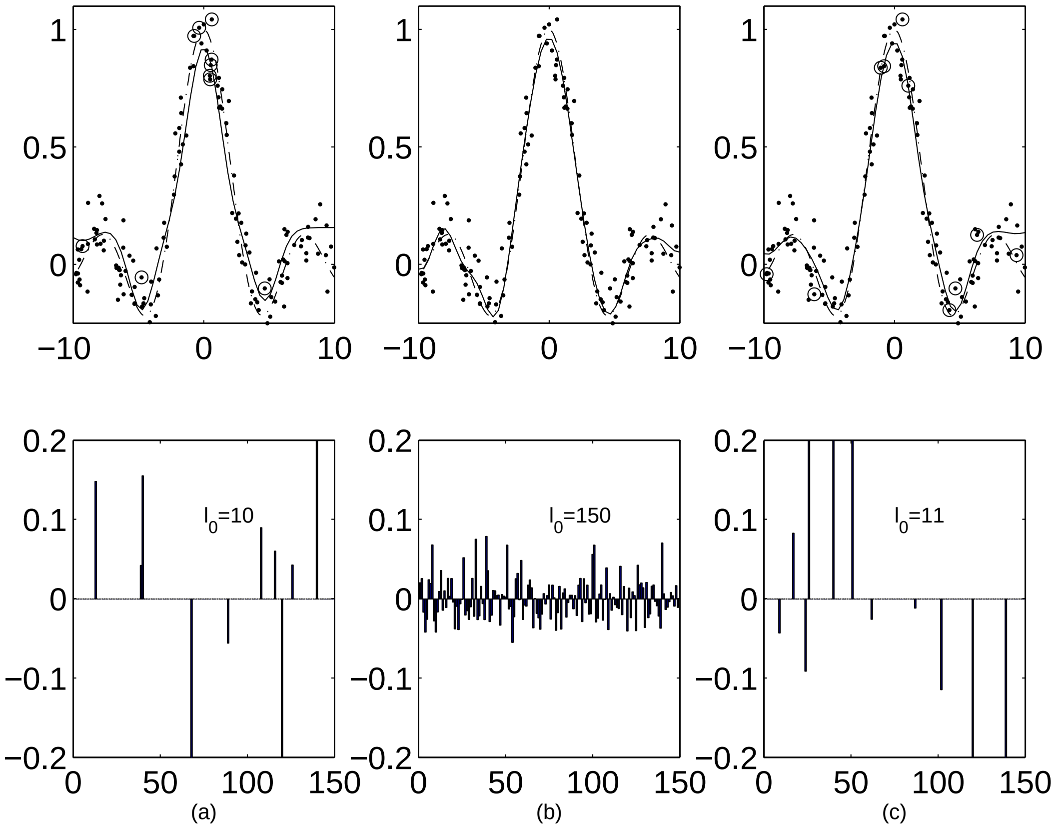


Fig. 2. Kernel regression results. The dashed lines are true sinc functions. Solid lines are approximation results. (a) Proposed gradient method. (b) Ridge regression. (c) LASSO regression.

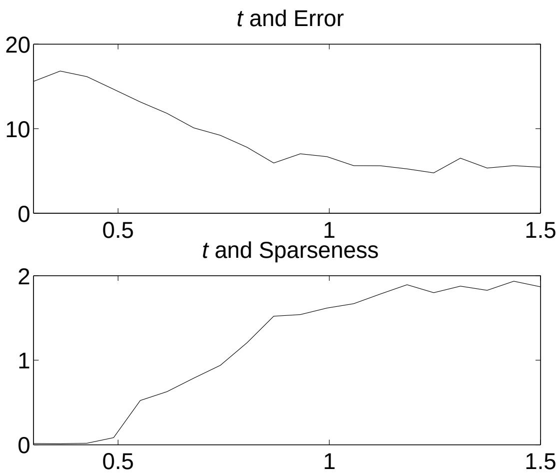


Fig. 4. The effect of control parameter t over mean square error and sparseness.

where ψ denotes the logistic function. If $p(+ \mid \mathbf{x}) > p(- \mid \mathbf{x})$, $\mathbf{x}$ belongs to the class $+$. Otherwise, $\mathbf{x}$ belongs to the class $-$.

We used two data sets from real-data problems: the *Pima indian diabetes*[1], which were collected from women of Pima heritage and the goal is to decide whether a subject has diabetes or not, based on 7 different tests; the *Wisconsin breast cancer* (WBC)[2], whose goal is to diagnose (benigh/malignant) based on the results of 9 measurements. In WBC, we removed the cases with missing attributes for simplicity. Table 1 shows the results of the proposed classifier on the 5-fold cross validation experiment. For comparison, we also include the results of the kernel-based logistic classifier. In both data sets, the proposed classifier is better. The performance is improved partially by setting some decayed kernel weight to be exact zero.

[1]Downloadable at www.stats.ox.ac.uk/pub/PRNN
[2]Downloadable at www.ics.uci.edu/ mlearn/MLSummary.html

Table 1. The result of the 5-fold cross-validation.

ROC/No. kernels	Pima	WBC
Logistic	0.750/200	0.772/455
Proposed classifier	0.965/70	0.980/253

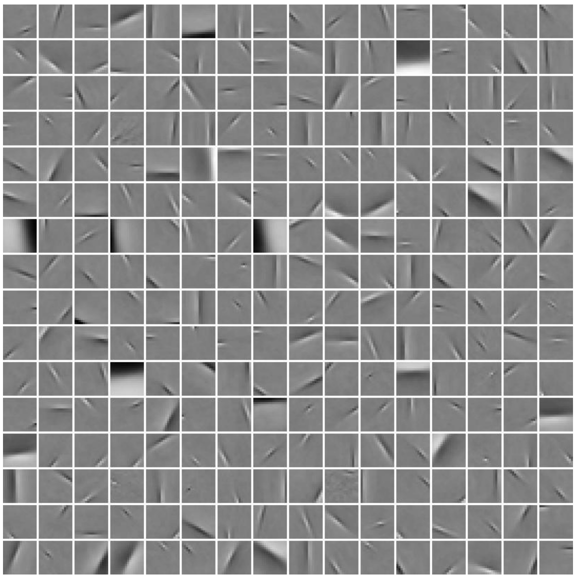


Fig. 5. Basis functions of natural images. The basis dimensionality is 16×16.

3.3. Filter Development

We also conducted experiment on filter development of natural image patches. 500,000 of 16×16 image patches were randomly taken from thirteen natural images. The image patches were subtracted by mean and pre-whitened, but the original dimensionality was kept. The proposed algorithm was applied to the pre-whitened data to update the design matrix $\mathbf{H}$. Each column of the matrix is shown in Fig. 5 by a 16×16 patch, as the features of the natural scenes. A total of 256 basis 256-dimensional vectors are displayed. They all look smooth and most of them are localized as expected. The entire learning procedure took less than 46 minutes on a Pentium III 700MHz PC with 512MB memory compared to over 10 hours of learning time for the Fixed Point algorithm [11] using 24% of the samples (disk thrashing is also a factor).

4. CONCLUSIONS

In this paper, we have formulated the dual problem of the ℓ_1 norm based sparse approximation. We show the geometric relation of the duality and solve the dual ℓ_∞ maximization problem by gradient and quadratic programming. The algorithm's performance is close to or better than the result of LASSO regression. Comparing with traditional method, the efficiency of this algorithm is quite remarkable.

5. REFERENCES

[1] A. Hoerl and R. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, pp. 55–67, 1970.

[2] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Statistical Society (B)*, vol. 58, pp. 267–288, 1996.

[3] M. Tipping, "Sparse bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[4] M.Ã.T. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1150–1159, 2003.

[5] E. Amaldi and V. Kann, "On the approximibility of minimizing non zero variables or unsatisfied relations in linear systems," *Theoretical Computer Science*, vol. 209, pp. 237–260, 1998.

[6] D.G̃. Luenverger, *Optimization by Vector Space Methods*, John Wiley and Sons, Inc., 1969.

[7] T. Kohonen, *Self-organizing Maps*, pp. 60–62, Springer-Verlag, New York, 2001, 3rd edition.

[8] J. Weng and N. Zhang, "A quasi-optimally efficient algorithm for independent component analysis," in *Proc. IEEE Int. Conf. on Acoustics Speech, and Signal Processing*, Montreal, Canada, 2004.

[9] J. Weng, Y. Zhang, and W.-S. Hwang, "Candid covariance-free incremental principal component analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 1034–1040, 2003.

[10] N. Zhang and J. Weng, "Sparse representation from a winner-take-all neural network," in *Proc. IEEE Int. Joint Conf. on Neural Networks*, Budapest, Hungary, 2004.

[11] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

Machine Emotional Intelligence: A Novel Method for Spoken Affect Analysis

Irina Gorodnitsky
Neo-dynamic Engineering & UCSD
San Diego, CA USA
igorodni@neo-dyne.net

Claudia Lainscsek
University of California, San Diego
9500 Gilman Dr. La Jolla, CA USA
clainscsek@ucsd.edu

Abstract

The paper presents a new text-independent system for extracting features that are shown to have the capability to provide discrimination of speaker emotional state. The system is a novel tool that nds low-level features in a particular projection space. The result is a representation of acoustic signals in form of timing sequences. The representation reveals characteristic structures in data. The key feature of the method is the learning of the timing information directly from the data. An analogy is pointed out between the timing representation and auditory-like signal encoding. The features are shown to contain information necessary to perform text-independent affect recognition. The experiments are performed using sentences from a German-language emotional speech corpus [1]. The corpus includes sentences spoken with two primary emotions: happiness and cold anger, and in a neutral speaking style, and two types of sentence accentuation: on a noun and on the nal verb. The features are shown to have basic discrimination properties similar to pitch but provide far better performance than pitch in noisy conditions.

1. Introduction

Human-computer interfaces, medical diagnostics, and consumer product design are just a few of the fields that can benefit from technology that can recognize and adapt to users emotional state. Effective human-computer interactions necessitate that machines posses at least a subset of emotion processing skill of humans. A computer that recognizes what a person says but ignores how the person says it will appear as if talking but never listening, thus likely to annoy the user.

In the medical field this technology is useful in evaluating patient emotional states and stress. In consumer product design, it can be used in assessing user acceptance of ever more complicated technologies by evaluating the level of frastration arising from interaction with a product.

Although humans can easily perceive emotions from auditory cues, corresponding machine recognition technology has been slow to develop. The difficulty in recognizing emotions is compounded by their wide range. Not only there are numerous emotional extremes that sometimes are not even discriminated consistently the same by human speakers, for example, joy-happiness-elation, but the expression of emotions spans a continuum between these extremes. Emotions that carry a similar connotation, i.e. negative (or positive) emotions, may have very dissimilar speech features, that reflect, among other things, very different levels of arousal. Anger and sadness are example of emotions that have a negative connotation but entirely different arousal level, high in anger and low in sadness [2]. Thus a fundamental question, what would constitute an effective definition system for emotions, remains open.

A widely accepted system is to divide emotions into primary and secondary. Secondary emotions are expressed through combinations of primary emotions. The list of primary emotions that is generally agreed on is: anger, disgust, fear, happiness, sadness, and surprise [3]. However, other classification systems, in particular ones based on physical properties, e.g. arousal level [2], also have been considered.

The difficulty in identifying emotions is also compounded by the localized and nonstationary nature of nature of emotional expression. A speaker may place the entire emotional expression on one word in the sentence, or spread it over some segment and even the duration of a sentence. Emotional expression can also vary with prosodic content, pitch and pronunciation. Some individuals have acoustic 'gestures' in their expression of affect that are unique to them and problematic for current automated recognition systems.

So far to be most effective feature for automated affect recognition has been pitch and measures derived from it. Pitch is the fundamental frequency of a signal, commonly denoted by F_0. Mean and standard deviation of F_0, mean duration, variability, slope, jitter (number of changes in sign of the pitch derivative in a window), and range have been used by a number of researchers, e.g. [4, 5, 6, 7, 8]. Notably, the best results in these papers were reported when pitch-based measures were combined with others, most common of which were mean energy, value of high frequency energy, energy per phone, articulation rate, speech rate (frequency of occurrence of unvoiced periods), speed of speaking (duration of inter-sentence silences), intensity, intensity variance and tremor (measure of tremor in the intensity over the intensity curve).

Speech rate, in particular, has been shown to be consistently modulated by certain emotions, highest in 'anger' and slowest in 'sadness'. Most recently, good recognition rates have been reported, using long-term (high-level) and short-term (low-level) features derived from these multiple metrics. Noam showed recognition rates with long-term features ranging from 55% to 98% [9]. Li and Zhao suggest that long-term features perform better then short-term features [10], but good recognition results have been shown with short-term features as well. Noquerias et. al. [11] used short-term features with recognition rates of 82.5%.

The present paper explores a novel acoustic-level feature for affect recognition from continuous speech. The motivation for the method is to model the physical mechanism involved in speech production that give rise to the acoustic gestures related to affect. To this end, we propose a new feature space derived through a projection of data as described in Section 3. The resulting approach is shown to exploit structures in signals which it learners directly from the signals. The structures are expressed in the form of a timing sequence. We refer to this representation as the Interval Domain (ID) representation. The modeled structures can be arbitrarily complex, containing any combination of linear (i.e. periodic) and nonlinear (aperiodic) components plus stochastic noise.

The feature space used is derived as a projection designed to minimize effect of noise. Being able to analyze an arbitrary structure embedded in noise, without a priori having to establish the nature of the signal, is one powerful feature of this technique.

The method is tested using sample sentences from a German-language emotional speech corpus [1]. It is shown to extract features correlated happy, and angry emotions and with a neutral speaking style. The performance comparison between pitch and new feature shows that the ID feature is comparable to pitch in noise free conditions, but that it drastically outperform pitch in noise. The ID-feature is shown to identify affect related gestures in -5dB SNR levels when pitch can no longer identify voiced parts of speech.

The mammalian auditory system is known to encode incoming acoustic signals as timing sequences of neuronal spikes. The mathematical principle for such encoding of a temporal signal as a timing sequence is not known. However, it is interesting to note that our mathematically derived approach results in the same general principle for representation in terms of a timing sequence that is used by the mammalian auditory system.

The paper does not address design of a recognition system based on the new features. The aim here is to understand the basic properties of the features first. Many factors go into design of a succesfull affect recognition systems including choce of scoring functions made of multiple signal features. Thus, the new feature may prove useful in combination with traditional features in particular where robustness to noise is required. Investigation of these issues is left for future work.

2. Emotional Speech Corpus

The experiments in this paper use seven sentences from the German language emotional speech corpus described in [1]. The entire corpus is comprised of 148 sentences with identical syntactic form (subject-auxiliary-NP-verb), where NP stands for 'the nominal phrase'. The 148 sentences are divided according to their lexical content. The lexical content, neutral, positive, or negative, was determined by having a group of subjects (n=20) rate the sentences. The sentences were recorded while spoken in Standard German by a trained female speaker. Each sentence was recorded and appeared in the database 6 times, using two forms of accentuations (on the NP and on the verb) and three forms of emotional state (happiness, neutral, and cold anger). Thus recorded utterances either matched sentence lexical content or mismatched it. The seven sentences were randomly chosen from the corpus and the six recordings of each of the sentence were provided to us.

One of the objectives in creating this corpus was to examine the connection between affect-dependent acoustic features and the neural responses of listeners, which were monitored using event-related brain potentials (ERPs). One of the study aims was to discriminate the different semantic conditions from listener responses. The use of the two forms of the accentuation were motivated by the hypothesis that the accented syllables are hyper-articulated while unaccented syllables are hypo-articulated. Vocal effort involved in hyper-articulation may produce measurable differences in acoustic features of emotional expression. The match/mismatch between the lexical content and the spoken affect was the other variable condition. The researchers in [1] hypothesized that the mismatch condition would produce a stronger emotive expression.

The objective of the present study is different from the goals in [1]. We are not concerned with discriminating among the different semantic conditions (match versus mismatch, NP versus final verb accent). The aim here is to recognize the expressed affect from spoken sentences. Thus we use only the speech corpus part of the data in our analysis. The different semantic conditions used in the original study yield an interesting dataset for testing affect-recognition in a variety of speech forms.

3. Timing sequence model for affect recognition

The method presented here is derived from the embedding concept of nonlinear dynamics theory [12]. The impetus for its development was the need to analyze real-life data that could originate in unknown environments that are complex, unstructured, and highly noisy.

At the heart of the method is identification of the deterministic structure in data that may be embedded in high amplitude, random noise. The formulation of the problem is straightforward. Given some data, we are interested in identifying the presence of all structures, periodic (linear) as well as aperiodic (nonlinear), they may contain. The problems associated with analysis of stochastically contaminated time series using classical nonlinear dynamics theory has been well described in the past. Casdagli at al. in [13] showed that given even arbitrarily small amounts of noise, some of the degrees of freedom of a system become completely unrecoverable. This means that classical embedding theory cannot be expected to be almost valid when data are almost deterministic. In other words, formal embedding reconstruction is not directly applicable to noisy data.

To deal with this problem several researchers have used projective schemes that identify a manifold in the embedding space (e.g. [14, 15]). The idea is that deviations of a trajectory in the embedding space from a manifold are caused by random noise in the data and the projection onto the manifold filters this noise, thus recovering the deterministic structure buried in data. Such projection techniques have been proposed and demonstrated on a number of signals including speech [14]. The projective techniques used in these works rely on recovery of the manifold from a reconstructed embedding.

The approach taken here is different. One of the immediate restrictions of the classical embedding theory is the fact that information contained in the embedded representation is critically influenced by the choice of embedding parameters and in particular the choice of the time delay values. There have been practical examples where the theoretically sufficient dimension can produce less optimal results than using a smaller dimension []. A relevant consensus from the existing works is that to optimize the embedding performance, including minimizing redundancy in the reconstruction, one should consider time delays that are variable in length, not integer multiples of a common lag [16].

We interpret this conclusion in a derivation where we learn the time delay parameters from data. Starting with the classical embedding theory, one can represent the state of an L-dimensional system from its output time series $x(t)$ by constructing an object in a space spanned by the time series and its time-delayed replicas. The object is a *phase trajectory* which can then be written as $\mathbf{x}(t) = \{x(t), x(t-\tau_1), \ldots, x(t-\tau_{D-1})\}$, which is a function of delayed coordinates. Based on the discussion above, we explicitly assume non-uniform delays $\tau_i,\ 1 \leq i \leq D-1$.

The importance of this construct is that it is diffeomorphic to the original phase space for sufficiently large D so that topological properties of the original high-dimensional system are preserved in the embedding under relatively loose restrictions. This means we can extend embedding theory to model the **dynamics** of the system output. Specifically, we express the evolution of the state vector $d\mathbf{x}(t)/dt$ as a function of the phase trajectories, i.e. $d\mathbf{x}(t)/dt = \mathbf{F}[\mathbf{x}(t), \mathbf{x}(t-\tau_1), \ldots]$. This formulation provided a novel data representation strategy. Instead of chosing delays to construct the embedding, we estimate the parameters of the deterministic function $\mathbf{F}$ from the data derivative $d\mathbf{x}(t)/dt$.

We estimate it in the following way. A general non-linear real-valued function can be extressed in a Taylor series expansion of functionals of increasing complexity around some fixed point. When the function $\mathbf{F}[\cdot]$ represents behavior of a dynamical system, that is, a time series model where the input is formed from past inputs $[x(t), x(t-\tau_1), \ldots]$, the expansion becomes a *Volterra series*. We have

$$\begin{aligned}\frac{d\mathbf{x}(t)}{dt} = \quad & \mathbf{x}_0 + \sum_{i=0}^{\infty} g_i \mathbf{x}_{\tau_i} + \sum_{i_1=0}^{\infty}\sum_{i_2=0}^{\infty} g_{i_1,i_2}\mathbf{x}_{\tau_{i_1}}\mathbf{x}_{\tau_{i_2}} + \cdots + \\ & \sum_{i_1=0}^{\infty}\sum_{i_2=0}^{\infty}\cdots\sum_{i_q=0}^{\infty} g_{i_1,i_2,\ldots,i_q}\mathbf{x}_{\tau_{i_1}}\mathbf{x}_{\tau_{i_2}}\ldots\mathbf{x}_{\tau_q}\end{aligned} \tag{1}$$

Equation (1) models the linear and non-linear data components as separate model terms. To find a model that is a projection onto a stable manifold, we consider low-order models made of a finite number of leading terms of equation (1). In other words we subselect candidate structures from equation (1) and fit them to data until we identify the smallest best fitting model.

We find the following low-order general structure to work well in many practical applications that we have attempted, including modeling affect expression

$$F(x,t) = \sum_k a_k x_{\tau_i}^l x_{\tau_j}^m, \tag{2}$$

where $x_{\tau_i} = x(t-\tau_i)$ is a delayed data vector with the delay τ_i , $i, j, l, m \in \mathbb{N}_0$ and $\tau_{i,j}$ permiting zero values, i.e. the signal itself.

This idea of restricted complexity of the model, i.e. leaving some of the dynamics unmodeled, plays a key role in the development of the practical algorithm. First of all, it allows us to reduce the computational load in this ill-posed problem to a manageable level so we can solve for the terms of the model. We describe this later when we present the final practical design model. Second, the unmodelled dynamics provides a means to control effect of noise on the estimation, much like the use of regularization in linear estimation.

The model in (2) permits polynomial functions with up to two non-zero delayed data vectors. The linear part of the equation can contain the scaled data itself plus up to two scaled delayed versions of the signal. The non-linear part of the equation permits any number of two term products of data and/or their delayed versions.

The model estimation problem reduces to a two part task: first select an appropriate low order model expansion and then fit the unknown parameters using the derivatives of the measured data. This estimation problem is non-trivial because its highly ill-posed and because the unknown parameters depend non-linearly on the data. We use a genetic algorithms (GA) to perform optimization here.

4. Emotive Feature Extraction

The analysis in this paper was done using sample sentences from the German-language emotional speech corpus [1] described in Section 2. Seven sentences, which were numbers 17, 43, 83, 85, 112, 123 in the corpus, were randomly chosen by a third party and provided to us for analysis. Of the seven sentences provided to us, three were rated as having positive lexical content, three sentences rated as having negative lexical content and one sentence rated as having neutral content. Each sentence was recorded and appeared in our database six times: with three types of affect and two types of accentuations as described in Section 2. Hence, we had a total of 42 records available for analysis.

Bellow is a list of the seven sentences. The number in front of each sentence indicates its position in the corpus and the sign '+', '-', or '0' indicates the positive, negative, or neutral rating of the sentence.

85 $+$ Sie hat es ans Licht gebracht.
(She brought some facts to light.)
103 $+$ Er hat um ihre Hand angehalten.
(He asked for her hand in marriage.)
112 $+$ Sie hat den Rekord gebrochen.
(She broke the record.)
17 $-$ Sie hat ihn mit der Waffe bedroht.
(She threatened him with the weapon.)
40 $-$ Er hat sie von der Klippe gestoben.
(He pushed her from the cliff.)
83 $-$ Er hat ihn ins Gesicht geschlagen.
(He slapped him in the face.)
123 0 Er hat den Brief geschrieben.
(He wrote the letter.)

Affect-recognition is typically investigated in the lexical context that is either nonspecific or consistent with the expresed affect. Yet, in practice, many instances can be found where the spoken affect mistmaches the lexical content, sarcastic expression being one example. Affect-lexicon mismatch in the given corpus provide an interesting study case in this respect. The authors in [] hypothesize that a speaker may use a stronger expression of affect in the mismatched condition, thus resulting in stronger acoustic features than in the neutral or matched condition. As shown below, we find strong evidence in our models that supports this.

4.1. Feature extraction model

Prior to analysis, the data were down-sampled from the original 44100Hz down to 8820Hz. The speech feature model was derived in two steps. Two, three, and four term 'best' models from the general model, Eq. (2) were estimated for the 42 sentence dataset using GA and the model coefficients were calculated for frames ranging from 20-ms to 100-ms with various overlaps. The model-frame-overlap combination which was consistently selected by our GA algorithm to produce the smallest fit error (RMS) in all 42 recordings was the following two delay 2nd order model

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1} x_{\tau_2}. \tag{3}$$

with 74.3-ms frames and 12-ms updates.

This model-frame-overlap setting was selected for the subsequent analysis. Parameters for all 42 records and all frames were re-calculated again using this setting. Analysis of the results is presented in the next subsection. Unlike typical derivation of a recognition model, no training for affect recognition occurred during this model selection. The model was strictly derived based on the smallest RMS error fit to continuous speech segments. Thus, this model is a general model of speech features represented in the ID space, in theory, could account for the speaker specific acoustic features, phonetic features, affect related features, and accentuation, and possibly other features of speech. Analysis of this model parameters and coefficients reveals that the coefficients a_1, a_2, a_3 each relate in a different way to the phonemes in the sentence. The relationship between the coefficients a_1 and a_2 reflects speaker acoustic characteristics and the parameter τ_2 is related to the emotive expression and to some degree to the sentence accentuation. Thus, the ID model provides a full spectrum of low-level features in speech signals.

4.2. Results from experiments: ID feature model

All seven sentences each spoken six different ways were analyzed using model Eq. (3). The present study focuses on the performance of the τ_2 parameter only. We summarize the results and present selected examples as space allows. The following questions are addressed in the study: 1) How consistent τ_2 responds to changes in affect. 2) The dependency of the response on the sentence lexical content and the placement of accentuation. 3) How the τ_2 feature compares to pitch. 4) Performance of the τ_2 feature and pitch in the presence of noise. Overall, the τ_2 values from model Eq. (3) were found to correlate consistently and in a meaningful way with the three expressed affects: neutral, happiness, and cold anger. This was consistent across all seven sentences and for both forms of accentuation. The observed level of discrimination of affect was similar to that of pitch in noise-free environments. τ_2 performed significantly better than pitch in noise.

Figure 1 shows the computed raw τ_2 values for each window for the sentence #17 spoken with negative affect and NP accentuation. The values form a clearly visible line, except during unvoiced segments, in which randomly distributed τ_2 values are seen. To aid in the future classification process, the unvoiced segments are identified and removed by setting τ_2 values in these segments to zero. Filtered τ_2 values are shown in Figure 1 as well. The employed filter is a moving data clustering predictive filter. Throughout the remainder of the paper we use filtered τ_2 values.

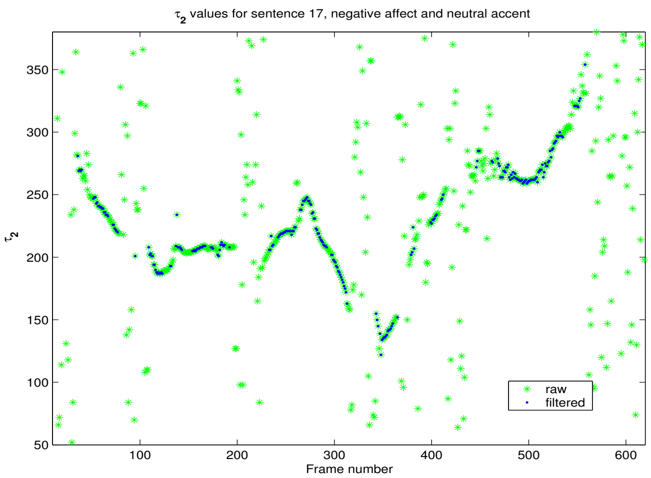


Figure 1. Raw and filtered τ_2 values for sentence number 17, negative affect, neutral accent.

The present section examines how consistent τ_2 responds to changes in affect. The results we present are consistent across all seven sentences and for both forms of accentuation. For reasons discussed below, we find least well defined separation in τ_2 for the three affect conditions when the sentence is lexically neutral (number #123) when NP is being accentuated.

The first experiment was performed using the first lexically non-neutral sentence (number #17, negative) with

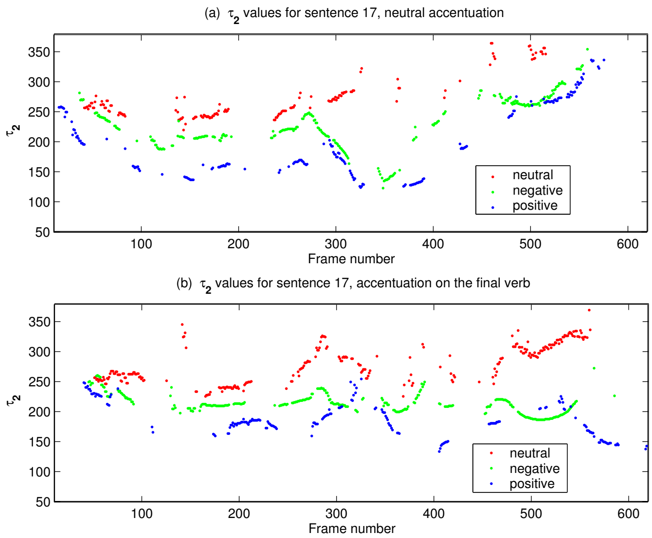


Figure 2. τ_2 values for the three forms of expressed affect in the lexically negative sentence number 17. (a) Neutral accent. (b) Accent on the final verb.

the three affect expressions and two forms of accentuation. Plots of the filtered τ_2 values are shown in Figure 2(a), neutral accentuation, and (b), accentuation on the final verb. Three clearly separated lines can be observed in the plots. In both plots, the three lines begin at about the same level but separate quickly. The upper line shows no significant change in amplitude and corresponds to the neutral expression, the middle line, to cold anger, and the bottom line to happiness. The relative differences in τ_2 are analogous to those in pitch values as shown later. In the NP accentuation case, the three affect conditions are separated throughout most of the sentence, coming together at the very end, at the point of unaccented verb. The τ_2 values settle at their characteristic level quickly and remain consistent at that level through the first half of the sentence, so that recognition of expressed affect could be made from the values in that half of the sentence. In the final verb accentuation case, the τ_2 values are also separated throughout most of the sentence, although slightly less pronounced in the first half of the sentence. In the second half of the sentence, the general three-level pattern remains but around frame 500 the values for the 'happiness' condition rise slightly above the values for the 'anger' condition and then decline back to their low levels. The rise may reflect the accentuation of the final verb rather than the difference in affect. Alternatively, the rise may be related to the change in affect, implying that affective expression may be described not solely by the τ_2 values but also the their dynamics of the τ_2 profile.

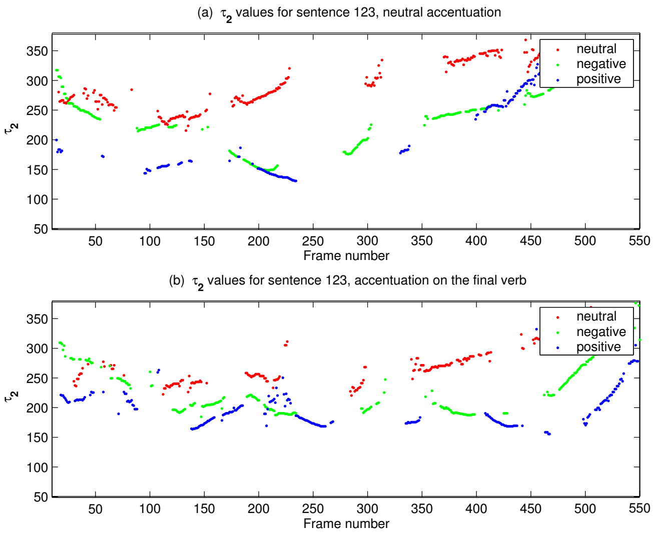


Figure 3. τ_2 values for the three forms of expressed affect in the lexically neutral sentence number 123. (a) Neutral accent. (b) Accent on the final verb.

Plots in Figure 2 demonstrate typical results obtained in all 7 sentences. The lexically neutral (number #123) sentence with NP being accentuated may be considered to provide a slight deviation from the typical results. Plots of the filtered τ_2 for this sentence are shown in Figure 3. The top to bottom line order is the same as for sentence #17 in Figure 2. The 'anger' line, however, follows the neutral line at the beginning and drops down at approximately the 180th frame, shortly prior to and during the accented part of the sentence only. On the other hand, the 'happiness' line stays clearly separated at a low level throughout the entire first part of the sentence, as in all other sentences. The late drop in the anger line, in fact, may facilitate recognition between anger and happiness. In the final verb accentuation case, the τ_2 values are separated throughout the sentence, characteristic of what is seen in the rest of the sentences for the verb accentuation case.

NP accentuation is the default accentuation in German for verb-final sentences. Thus this sentence utterance represents the most 'regular' form of speech. The response in τ_2 immediately prior to and during the accented part of the sentence only in the 'anger' case may support the hypothesis that vocal effort involved in the normal speech is less than in the cases of unusual constructs. Less vocal effort translates into less pronounced affect expression, hence it can only be recognized during the accented part of the sentence.

To assay the affect-lexicon independence we compare τ_2 values for the two primary emotions, happiness and cold

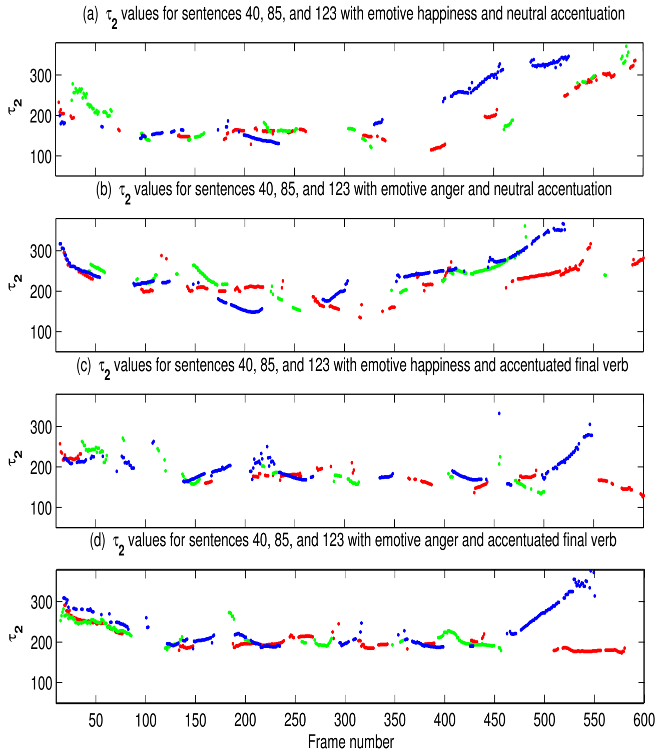


Figure 4. τ_2 values for three different lexical connotations, sentence numbers 40, 85, and 123 ploted for 'anger' and 'happiness' forms of affect and the two forms of accent. (a) Expression of happiness, neutral accent. (b) Expression of anger, neutral accent. (c) Expression of happiness, accented final verb. (d) Expression of anger, accented final verb.

anger, across the different lexical conditions. Overall, very little difference in τ_2 values was found accross the lexically different sentences for each form of affect expression. Figure 3 shows τ_2 plots for three lexically different sentences (numbers 40, 85, and 123), for two affect conditions, 'happiness' and 'anger', separated by the two forms of accentuation. The only obvious separation in τ_2 values can be seen in frames 325-350 in the top plot of Figure 3. This point is where the main emphasis is placed in the sentence, a point of hyper-articulation.

4.3. Results from experiments: comparison with pitch

A number of effective algorithms is available for computing pitch. For this paper, F_0s were estimated using the Tcl/Tk Snack audio analysis module on a Linux platform

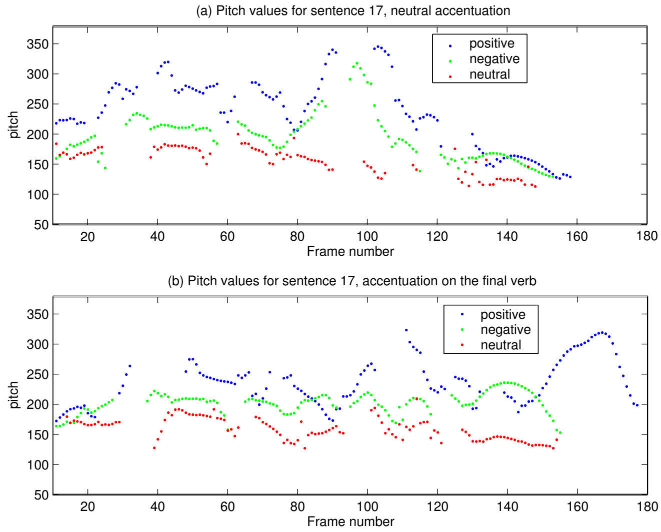


Figure 5. Pitch values for the three forms of expressed affect in the lexically negative sentence number 17. (a) Neutral accent. (b) Accent on the final verb.

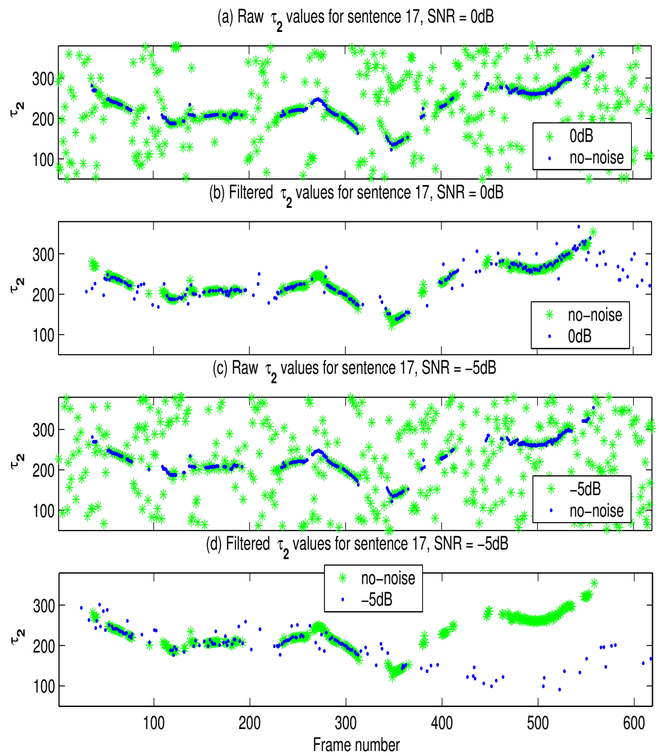


Figure 6. τ_2 values for noisy data, sentence number 17 with angry expression and neutral accent. (a) Raw τ_2 values for 0dB SNR and noise-free case. (b) Filtered τ_2 values for 0dB SNR and noise-free case. (c) Raw τ_2 values for -5dB SNR and noise-free case. (d) Filtered τ_2 values for -5dB SNR and noise-free case.

[17]. The ESPS method in Snack was used to compute pitch which integrates normalized cross-correlation pitch candidate generation with dynamic programming to select the optimal pitch track. The default settings of 10 msec and 7.5 msec were used respectively for the value spacing (-framelength) and the window length. Because ESPS method also calculates probability of voicing, the pitch values were selected only for the voiced speech segments and were set to 0 for the unvoiced speech segments.

Figure 5 show pitch curves for the same sentence number #17 as shown in Figure 2. The order of the curves is the same but inverted from the τ_2 curves of Figure 2. The lower τ_2 values correspond to higher F_0s. Otherwise, the correspondence is similar in noisy-free case and is consistent for all seven sentences.

4.4. Results from experiments: performance in noise

Performance in noise between F_0 and τ_2 is significantly different. Two experiments, with $0dB$ and $-5dB$ SNR levels are shown here. In both cases, white, Gaussian distributed noise of appropriate level was added to the speech signal. For consistency, the analysis is done using the same sentence number #17, first record, which was recorded with angery expression and neutal accentuation. The noisy signals were not filtered prior to analysis. F_0 and τ_2 values were computed, using the same methods as above, for the noisy signals. The derived τ_2 values were also filtered as described above. Figures 6 and 7 show plots of τ_2 and F_0, respectively, for $0dB$ and $-5dB$ SNR levels. For illustration, both raw and filtered τ_2 values are shown in Figure 6. In the $0dB$ SNR case, the filter is able to successfully recover the τ_2 values of the noise free case even through the raw τ_2 values are fairly noisy. In the $-5dB$ SNR case, clusters of raw τ_2 values are seen distributed everywhere in the plot space. The employed filter is able to recover the τ_2 values of the noise free case in the first half of the sentence, after which out filter estimates diverged from the noise-free values. Note that in the noise free case the affect is clearly differentiated only in the first half of this sentence as well, which reflects dimished vocal effort in the second part of the sentece. A more sophisticated filter could possibly track τ_2 throughout the sentence and could be attempted.

Pitch estimates were not consistent in the $0dB$ SNR case with many sections of the sentence determined to be unvoiced speech. In the $-5dB$ SNR case, the entire signal

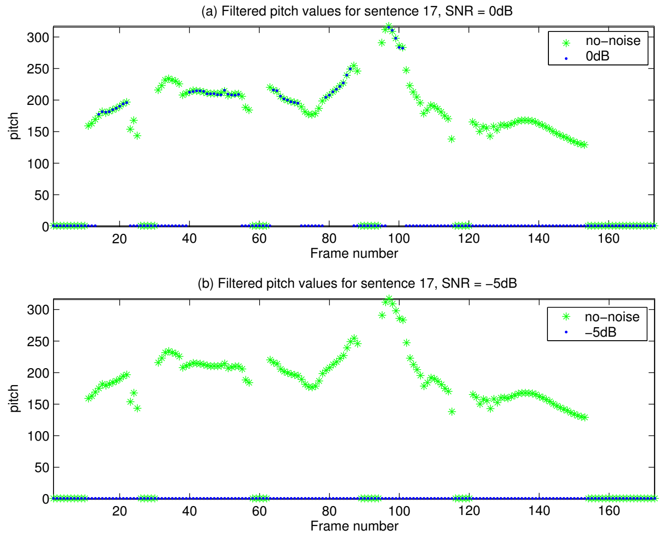


Figure 7. Pitch values for noisy data, sentence number 17 with angry expression and neutral accent. (a) Filtered pitch values for 0dB SNR and noise-free case. (b) Filtered pitch values for -5dB SNR and noise-free case.

was determined to be unvoiced speech. This was observed in all seven sentences.

The drastic difference in performance in noise between F_0 and the ID-feature is expected. The projection onto the manifold discussed in Section 3 is designed to mitigate noise in data. Since affect recognition in noise is a significant problem for current technology, we consider this result to be of great interest that warant further investigation of the proposed methodology.

5. Conclusions

We introduce the concept of timing domain representation, which is implemented by estimating directly from data the delay parameters of a projection onto a manifold in an embedding space. The resulting model estimates deterministic structures in data and the residual stochastic component. Affect recognition power of the model is demonstrated at the acoustic feature level. The model is shown to find features that distinguish neutral, happy, and angry emotions expressed by a speaker.

The concept of learning embedding parameters from data is in its infancy and it ushers in an entirely new approach. There are a number of areas that require further research, such as design of a classifier based on the presented model, training data requirements, and potential improvement of discrimination by adding conventional speech features to those identified here.

6. Acknowledgements

The authors thank Dr. Gernot Kubin for providing sentences from the emotional speech corpus.

References

[1] K. Alter, E. Rank, A. Kotz, U. Toepel, M. Besson, A. Schirmer, and A. Friederici. Affective encoding in the speech signal and in event-related brain potentials. *Speech Communication*, 40:61–70, 2003.

[2] J. Pittam and KR Scherer. The encoding of affect: a review and directions for future research. *4th Australian Int Conf Speech Science and Technology*, 1:744–9, 1992.

[3] R. Cornelius. *The science of emotion*. 1992.

[4] A. Iida, W. N. Campbell, S. Iga, F. Higuchi, and M. Yasumura. Acoustic nature and perceptual testing of corpora of emotional speech. In *Proc of the ICSLP 98*, 1998.

[5] B. Heuft, T. Portele, and Rauth. M. Emotions in time domain synthesis. In *Proc of the ICSLP 96*, 1996.

[6] J. E. H. Noad, S. P. Whiteside, and P. D. Green. A aacroscopic analysis of an emotional speech corpus. In *Proc EuroSpeech '97*, 1997.

[7] F. Dellaert, T. Polzin, and A. Waibel. Recognizing emotion in speech. In *Proc of the ICSLP 96*, 1996.

[8] N. Amir and S. Ron. Towards an automatic classification of emotions in speech. In *Proc of the ICSLP 98*, 1998.

[9] A. Noam. Classifying emotions in speech: A comparison of methods. In *EuroSpeech '01*, 2001.

[10] Y. Li and Y. Zhao. Recognizing emotions in speech using short-term and long-term features. In *Proc Int Conf on Speech and Language Processing*, pages 2255–58, 1998.

[11] A. Noquerias, A. Moreno, A. Bonafonte, and J. Marino. Speech emotion recognition using hidden markov models. In *EuroSpeech '01*, 2001.

[12] Floris Takens. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*, 898:366, 1981.

[13] M. Casdagli, S. Eubank, J.D. Farmer, and J. Gibson. State space reconstruction in the presence of noise. *Physica D*, 51:52, 1991.

[14] R. Hegger, H. Kantz, and L. Matassini. Denoising human speech signals using chaoslike features. *Physical Review Letters*, 84:3197–200, 2000.

[15] L. Matassini, H. Kantz, J. Holyst, and R. Hegger. Optimizing of recurrence plots for noise reduction. *Physical Review E*, 65, 2002.

[16] P. Grassberger, T. Schreiber, and C. Schaffrath. Nonlinear time sequence analysis. *Int. J. Bifurcation Chaos*, 1:521, 1991.

[17] K. Sjlander. The snack sound visualization module. http://www.speech.kth.se/SNACK/.

Cross-Task Learning By a Developmental Agent

Xiao Huang and Juyang Weng
Department of Computer Science and Engineering
Michigan State University
East Lansing, MI 48824
{huangxi4,weng}@cse.msu.edu

Abstract

Cross-task learning is essential for an artificial agent to explore in the real world. With this capability, an agent can learn multiple tasks and use acquired knowledge to learn new tasks. In this paper, we propose the Developmental, Observation driven, Self-Aware, Self-Effecting, Markov Decision Process (DOSASE MDP) model, which enables an agent to find the relatedness between different tasks and to develop cross-task learning capability by interacting with the environment. One challenge of this work is that no task is defined in advance. The trainer shapes the behavior of the agent to learn different tasks interactively and continuously through interactions. The development is conducted in real-time using high-dimensional input, which is another challenging issue. We tested the architecture on an artificial agent for vision-based navigation. The results show that DOSASE is an effective model for autonomous mental development (AMD) and dramatically reduces both time and space complexities for cross-task learning.

1 Introduction

Cross-task learning capability is very important for an artificial agent to explore in the real world. By cross-task, we mean that the same system must learn multiple tasks incrementally in the same mode, dealing with task specific contexts correctly. With this capability, the agent can learn different tasks and transfer learned knowledge to new tasks. Scientists in psychology proposed a lot of models to simulate human cognition in cross-task learning. Lovett [4] used an architecture called ACT-R to model how people organize knowledge and produce intelligent behaviors.They claimed that after separately fitting performance on a preliminary task, the model can make zero-parameter cross-task prediction of performance in on a second task. Computer scientists are also interested in building multiple-task learning models for artificial agents. Thrun [6] proposed the task-clustering algorithm, which learns the relationship between different tasks. When facing new a task, the algorithm first finds the most related task, then exploits information from this task. Pratt [5] and Caruana [1] investigated the similar problem: how to use information from one neural network to help a second network learn a related task. The limitations of these methods consist of: 1) human programmers know the task knowledge at the beginning. 2) in order to build up the relationship between different tasks, real-time learning is impossible. For a typical developmental agent, these limitations are not acceptable. A developmental agent has to run in real time and face different tasks, which are unknown to the agent before training.

In this paper, we propose the Developmental, Observation driven, Self-Aware, Self-Effecting, Markov Decision Process (DOSASE MDP) as the model of autonomous mental development (AMD) [8] to conduct cross-task learning. The advantages of the model include: 1) no task is defined at the beginning. By interacting with the trainer, the agent incrementally learns different tasks. 2) the agent can learn task-relatedness online in real-time. 3) the agent transfers learned knowledge to new tasks and dramatically reduces both time and space complexities. The model has been successfully tested on an artificial agent for vision-based navigation.

In the following section, we will first formulate what is cross-task learning. The DOSASE MDP model is introduced in section 3. Experimental results are shown in section 4 and then we conclude with a summary and discussion about future works.

2 Problem Description

To be precise in our further discussion, we need mathematical notations.

Definition 1 *Context: A context $c(t)$ of an agent is a stochastic process. It consists of two parts $c(t) = (x(t), a(t))$, where $x(t)$ denotes the sensory vector at time*

t, which collects all signals sensed by the agent at time t, $a(t)$ is the effector vector consisting of all the signals sent to the effectors of the agent at time t.

Definition 2 *Length of Context: The context related to the agent from the previous time t_1 up to a later time t_2 is a realization of a stochastic process $\{c(\tau)|t_1 \leq \tau \leq t_2\}$. The length of the context is $L = t_2 - t_1$.*

Definition 3 *Given an agent at time t_1, suppose that the agent produces different action contexts $\mathbf{a}_1$ and $\mathbf{a}_2$, from two different contexts $C_1 = \{\mathbf{c}(t) \mid t_1 \leq t \leq t_2\}$ and $C_2 = \{\mathbf{c}(t) \mid t_1 \leq t \leq t_3\}$, respectively. If $\mathbf{a}_1$ and $\mathbf{a}_2$ are considered different by a social group (human or robot), conditioned on C_1 and C_2, then we say that the agent* discriminates *two contexts C_1 and C_2 in the society. Otherwise, we say that the agent does not* discriminate *C_1 and C_2 in the society.*

For example, given a voice command "Go to the elevator," the speech signals of the command are different from different people. However, humans consider the commands are the same. In this case, the agent should not discriminate the above commands from different people.

Definition 4 *Cross-task Learning: 1). There are N tasks: $\Gamma = \{T_i|i = 1, 2, ..N\}$. 2). For each task, the learning goal is to generate the mapping from context to action: $M : C \rightarrow A$, where A is the action space, C is the context space, which could consists of different sensory input: vision, audition, touch etc. 3). Given the first N tasks, the agent uses acquired knowledge to speed up learning the N+1th task.*

A typical setting is shown in Fig. 1. There are two tasks: T_1 and T_2. For example, in autonomous navigation problem, T_1 is "go around" and T_2 is "go to the elevator." The contexts of each task are C_1 and C_2, respectively.

Definition 5 *Shared context: The overlapped trajectory between these two tasks is called shared context $C_{share} = C_1 \bigcap C_2$. Non-overlapped trajectory is called non-shared context $C_{non\_share} = C_1 \bigcup C_2 - C_{share}$.*

Give the above figure as an example, $C_1 = \{C_{11}, C_{12}, C_{13}, C_{14}, C_{15}\}$ and $C_2 = \{C_{21}, C_{22}, C_{23}, C_{24}, C_{25}\}$. In the task space, the shared context is $C_{share} = \{C_{12}, C_{14}\}$ and the non-shared context is $C_{non\_share} = \{C_{11}, C_{13}, C_{15}, C_{21}, C_{23}, C_{25}\}$.

Definition 6 *Merge point and break point: The point, where the contexts of two or more tasks begin to overlap, is call "merge point" (M_1, M_2). The point, where the contexts of two or more tasks begin to diverge is "break point" (B_1, B_2).*

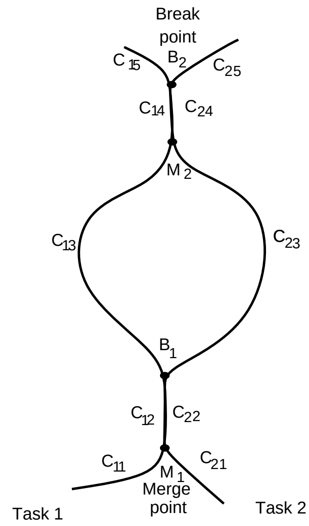


Figure 1. Typical setting of cross-task learning.

3 Architecture of DOSASE MDP

We propose the Developmental, Observation driven, Self-Aware, Self-Effecting, Markov Decision Process (DOSASE MDP) as a model of autonomous mental development to conduct cross-task learning. The architecture of this model is shown in Fig. 2. There are two level-building element (LBE) components. The first LBE online learns the association between task label and verbal commands. The second LBE learn the specific tasks (for example, different vision-based navigation tasks). A developmental agent is self-aware and self-effecting, which means the agent can sense both internal and external inputs and generate both internal and external actions. A typical internal action is thinking.

Definition 7 *Internal action and external action: Action consists of two component. That is $a(t) = (a_i(t), a_e(t))$, where $a_i(t)$ and $a_e(t)$ are the internal action and external action, respectively. $a_i(t)$ changes the internal state while $a_e(t)$ changes the external state.*

The input goes through a channel selector, which only pay attention to part of context. For example, the first LBE pays attention to auditory input while the second one pays attention to visual input. The input through the channel selector is fed into an observation-driven state transition function. After generating the current state, the cognitive mapping engine finds the best match and outputs the associated action. In the following subsections, we will discuss each component in details. It is worth noting that in the experiment only the second LBE is used for vision based navigation. The experimental result of voice command learning is not presented.

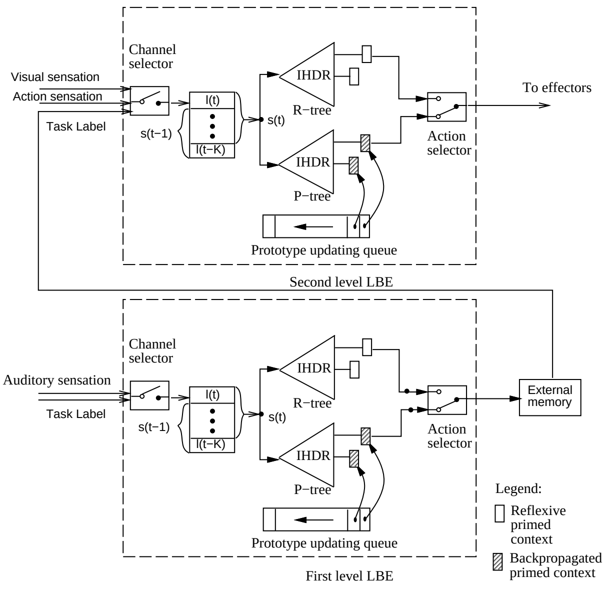


Figure 2. The cross-task learning system architecture.

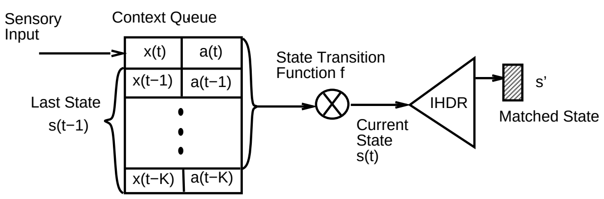


Figure 3. Observation-driven state transition function.

3.1 Observation-driven state transition function

After going through channel selector, the sensory input is fed into the observation-driven state transition component. Let's first define last context: $l(t) = f(a(t-1), x(t))$, which consists of last action $a(t-1)$ and current input $x(t)$. $l(t)$ is pushed into a context queue as shown in Fig. 3 and is combined with last state information to generate the current state. Mathematically, this is called observation-driven state transition function $g : S \times L \mapsto S$, where S is the state space, L is the context space and is generated by a queue of multiple sensory input. This is a general architecture. We can chose different lengths of the context queue. The observation-driven state transition function generates current state from last state and current context $l(t)$, which is defined as

$$s(t) = g(s(t-1), l(t)). \tag{1}$$

We should notice that $s(t)$ can be also defined as:

$$s(t) = g(s(t-1), l(t), l(t-1), ..., l(0)). \tag{2}$$

If we use Eq. 1, then we can treat it as a Markov decision processing. We call this model the Developmental, Observation Driven, Self-Aware, Self-Effecting, Markov Decision Process (DOSASE MDP). The difference between DOSASE MDP and POPDP are:

1. POMDP is a computational model but it is not a model generator. It requires humans to hand-design the model. However, DOSASE MDP contains both: a computational model and the model generator.
2. DOSASE MDP is a SASE model, but POMDP is not. By SASE, we mean that the model can sense inter input and issue internal action (for example, attention).
3. The state in the former is a symbol, which means that every state is different but there is no information about distance between states. The state in DOSASE MDP has a representation, which is subject to internal action.

3.2 Online learning of one task through cognitive mapping

Many problems like content-based retrieval, vision-based navigation can be formulated as a complicated function which maps high dimensional input and the current state to low-dimensional output signals. We use a decision tree to approximate this function, which is implemented by Locally Balanced Incremental Hierarchical Discriminating Regression (LBIHDR) [2] [3].

A detailed explanation is beyond scope. Basically, given a state s, the IHDR finds the best matched s' associated with action. The mapping is done through a coarse-to-fine tree structure. Each node of the tree is modeled by q Guassians. The original d-dimensional input space can be mapped into a q-1 dimensional discriminant subspace. We only conduct Linear Discriminant Analysis (LDA) in the very-low dimensional subspace, which saves tremendous computational cost. Each Gaussian is represented by its first two-order statistics: mean and covariance matrix. Mean is updated incrementally as follows:

$$\bar{s}^{(n+1)} = \frac{n-\mu}{n+1}\bar{s}^{(n)} + \frac{1+\mu}{n+1}s_{n+1} \tag{3}$$

where s_{n+1} is the $(n+1)$th state, $\bar{s}^{(n+1)}$ is the mean after this state is trained, μ is a parameter. If $\mu > 0$, the new input gets more weight than old inputs. We called this implementation the amnesic average. The covariance matrix can be updated incrementally by using the amnesic average too. In the testing phase, given s, we use K-nearest neighbor rule to find the best match in the leaf node.

$$s' = \arg\min_{1 \le i \le K} \parallel s'_i - s \parallel \tag{4}$$

Using IHDR, the relatedness of different task is measured by the similarity between the states generated for different tasks. If some of the states generated by two or more tasks are the same, then we say these tasks are related. That's why our architecture is better than the task-clustering method [6], which has to define different tasks in advance to measure the relatedness. Using the DOSASE MDP model, we don't need to know the task in advance but generating states for different tasks incrementally. IHDR intrinsically has incremental online learning capability to adapt to new inputs, which is a requirement for autonomous mental development. And because of its efficiency, it can learn high dimensional input, which outperforms the typical neural network algorithm [1].

3.3 One-task learning algorithm by a developmental agent

The algorithm to learn one task by a developmental agent is as follows:

Procedure 1 One task learning. Learn the association between context and action.

1: Collect the current context $c(t)$.
2: Go through channel selector to get sensory input $x(t) = c_v(t)$. (For vision-based navigation task.)
3: Push $x(t)$ to the context queue and generate generate last context $l(t)$.
4: Use observation-driven state transition function (Eq.(1)) to generate the current state $s(t)$.
5: Find the best match s' of $s(t)$. If they are similar, use amnesic average to update s'.
6: If imposed action is given, take this action. Otherwise, choose the action (a) associated with s'.
7: Go back to step 1.

3.4 Multiple task learning through DOSASE

How can this architecture learn multiples tasks? We know that the agent can learn one task with the model. Using the same architecture multiple tasks can be learned if the C_i and A_i are provided. Now, the problem is how to speed up learning if the agent acquires useful knowledge from former tasks? Given the tasks in Fig.1 as an example, if T_1("go around") has been learned, how the agent learns T_2 ("go to the elevator")? Let's take a look at the state space generated for each task.

A part of the state generated by these two tasks is shown in Fig. 4. The context of T_1 is $C_1 = \{c_1, c_2, c_3\}$. The state generated for T_1 is $S_1 = \{s_1, s_2, s_3\}$, where $s_j = \{s_{j1}, s_{j2}, ..., s_{jn_j}\}$. $j = \{1, 2, 3\}$ while n_j denotes the number of state generated for jth context. Now, the agent face T_2. The context is $C_a = \{c_a, c_b, c_c\}$. The state generated for T_2 is $S_2 = \{s_a, s_a, s_c\}$, where $s_m = \{s_{m1}, s_{m2}, ..., s_{mn_m}\}$. $m = \{a, b, c\}$ while n_m denotes the

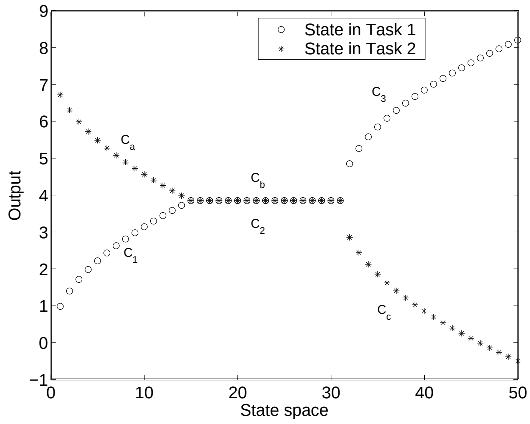


Figure 4. State space for 2 tasks.

number of state generated for mth context. The shared context of T_1 and T_2 is c_2 (c_b). The advantage of cross-task learning is that if the agent learns C_{share} for T_1, then there is no need to experience it again for T_2.

Lemma 3.1 *If the number of states generated for non-shared context of T_1 and T_2 is N_{Nshare} and the number of states generated for shared context is N_{share}: the ratio of the space complexity saved for multiple task learning is:* $\frac{N_{share}}{N_{Nshare}+N_{share}}$

Now let's think about the complexity in the training time space.

Lemma 3.2 *If the $N+1$th task shares context with the former N tasks: $C_{share} = (\bigcup_{i=1}^{N} C_n) \bigcap C_{N+1}$, the saved time for training the $N+1$th task is L_{share}.*

3.5 Attention mechanism for the break point set

Suppose the agent has been trained with C_1 and C_{share}, in the testing phase, the agent would go through the same context trajectory. However, there is a problem if the agent experience s_{bn_b} (the last state of context c_b) again. Since for different tasks the external actions a_e associated with s_{bn_b} are different. For T_1, the external action should make the state move from s_{2n_2} to s_{31} while for T_2, the external action should make the state move from s_{bn_b} (equal to s_{2n_2}) to s_{c1}. What should the agent do at the "break point?"

An attention mechanism is necessary to solve the problem. Remember that the agent learn the task label through the first LBE. If the agent get into the state s_{2n_2} (e.g. s_{bn_b}), there are two external actions a_{e1}, a_{e2}. Each external action is associated with a task label. Using the task label information, the agent knows what action should be taken to reach next state. Fig. 5 shows how the task label information helps discriminate state transition. Adding task label as

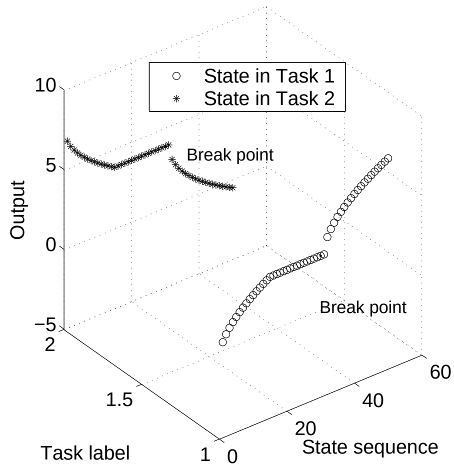


Figure 5. Discriminate state transition using multi-modal information.

another dimension, the break point is split into two points. By checking current task label, the agent knows which action to. The internal action generated by the attention mechanism is defined as:

$$a_i = \begin{cases} 1 & \text{if } T_i = 1 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

The output external action is $a = a_e \bigotimes a_i$, where $\bigotimes$ is component-wise multiplication. Only after checking the task label information, can an action be issued.

3.6 Cross-tasking learning algorithm

The algorithm is as follows:

Procedure 2 Cross-task learning.

1: Collect the current context $c(t)$. $x(t) = c(t)$.
2: Push $x(t)$ to the context queue and generate generate last context $l(t)$.
3: Use observation-driven state transition function (Eq. 1) to generate the current state $s(t)$.
4: Find the best match s' of $s(t)$. If they are similar, use amnesic average to update s'. Check the task label. If these two states have different task labels, generate a new action for s'.
5: If imposed action is given, take this action. Otherwise, check task label, then use Eq. 5 to pay attention to task label and take the corresponding action.
6: Go back step 1.

4 Experimental Results

In order to test the cross-task learning capability of the DOSASE MDP model , a simulation environment is developed. The simulation map is shown in Fig. 6. The artificial agent (rectangle with an arrow) can navigate through the white area. There is an elevator on the bottom right (white rectangle). In every state, the agent has three possible actions: go straight, turn left and turn right. Some of the example input images are shown in Fig 7. The second and the third images on the third row are the images near the elevator. The dimension of input image is 25×25.

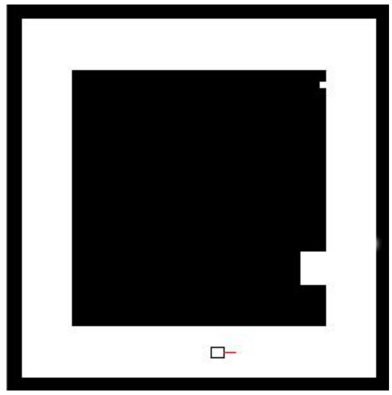


Figure 6. Simulation interface.

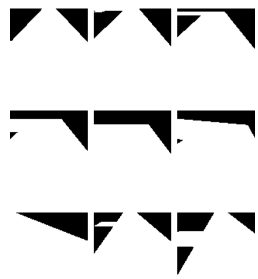


Figure 7. A subset of input images.

4.1 Trajectory and transferred knowledge

Only two tasks are considered: T_1 "go around" and T_2 "go the elevator." The contexts of these tasks are $C1$ and $C2$. The non-shared context is C'. The trajectories of two tasks are shown in Fig. 8. In the training stage, $C1$ (solid line) is trained and then C' (dot line from point 'A' to point 'B') is trained. C' includes context near the elevator. In the testing phase, the agent successfully finished these two navigation tasks. For task 2, its trajectory is different from that of task 1 around the "elevator" since this part is the non-shared. While another part of its trajectory is overlapped with that of task 1 because another part is not trained for task 2 but transferred from the knowledge of task 1.

Fig. 9 shows the task label of each sample when testing

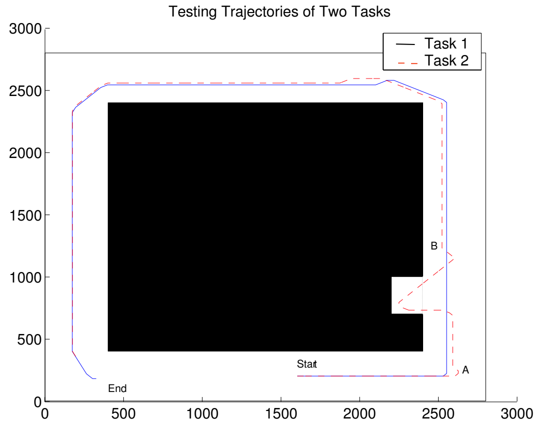


Figure 8. Trajectory of two tasks.

Table 1. Trasfered knowledge when testing on task 2.

Task	1	2	All	Transferred knowledge
No. of samples	2938	2442	5380	54.61%

on task 2. The task label in the testing phase is switching between 2 and 1. Tab. 1 shows that there are totally 5380 samples. 2442 of them are retrieved with task label 2 while 2938 of them are retrieved with task label 1, which means that for this test about 54.61% knowledge is transferred from task 1 to task 2.

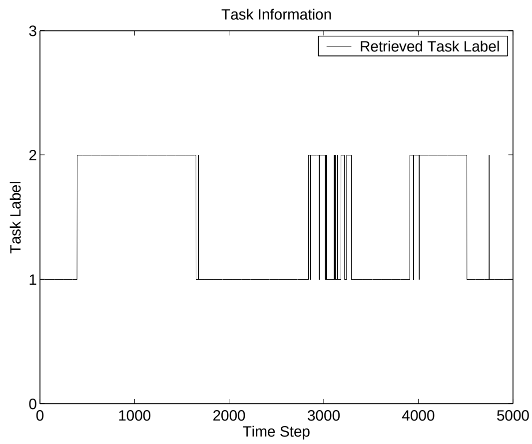


Figure 9. Task label information. The plot shows the retrieved task label of each sample when testing on task 2.

4.2 Complexity reduced in terms of space and time

The structures of IHDR trees for different tasks are shown in Fig. 10. The first two plots correspond to the IHDR tree architecture if task1 and task 2 are trained separately. The plots shows the number of states in different layers. The depth of both trees is 6. In the forth layer, the tree saves most of the states. The third plot shows the IHDR tree structure if the agent learns $C1$ first and then learns C'. As you can see, the depth of tree is still 6, which means that the system learns two tasks but the space complexity does not increase too much comparing with learning only one task.

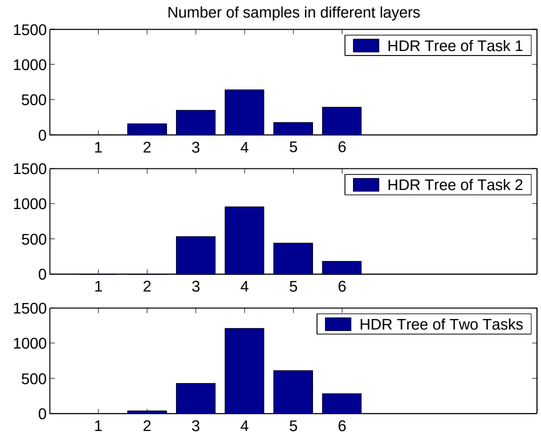


Figure 10. HDR tree structure. The first two plots show the number of states in different layers for task 1 and task 2, respectively. The third plot shows the tree structure if cross-task learning is conducted.

Table. 2 shows how cross-task learning reduces the complexity in terms of both space and time complexities. If we train 2 tasks separately, the size of the IHDR trees are 22.6M and 27.5M, respectively. If we train task 1 ($C1$), and then train the non-shared context (C') of $C1$ and $C2$, the size of the tree is 36.6M. The percentage of saved space is $\frac{22.6+27.5-36.6}{22.6+27.5} = 26.95\%$. The training time of task 1 and 2 is 147.42s and 143.75s, respectively. If cross-task learning is conducted, the training time is 205.83s. The percentage of saved time is about 29.31%. The experimental result shows that the model is effective in principle and cross-task learning really reduces a lot in terms of time and space complexities.

The training time of each step for task 1 is shown in Fig. 11. The average training time of each step is 0.043s

Table 2. Complexity reduced in terms of space and time.

Task	1	2	All	Complexity reduced
Tree size	22.6M	27.5M	36.6M	26.95%
Training time	147.42s	143.75s	205.83	29.31%

and the highest training time is 0.14s. Obviously, the system can work well in real time, which is a necessary condition to conduct autonomous mental development for an artificial agent.

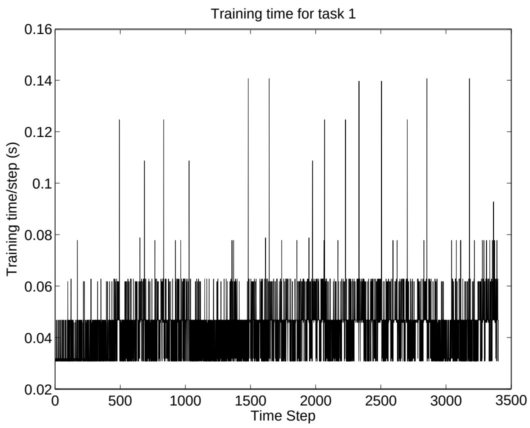


Figure 11. Training time of each step for task 1.

5 Conclusions

In this paper, we propose the DOSASE MDP architecture to model the cross-task learning in autonomous mental development. No prior knowledge is needed for each task. The model helps an agent to find the relatedness between different tasks by comparing the similarity of states generating for each task. The agent doesn't need to know all the tasks in advance. The learned knowledge in shared context can be transferred to new task, which speeds up the training procedure. The model is tested in a simulation environment for vision-based navigation. Two tasks have been incrementally learned in real time. There are big gains in terms of time and space complexities. The experimental result shows the effectiveness of the AMD paradigm. For the future work, we will test the model on a robot in the real world. This is a challenging problem since in the real world the number of state could be huge, which makes learning difficult. Also we would like to expand the model for learning using multi-modal inputs to see how a robot develops its visual and auditory capabilities at the same time.

References

[1] R. Caruana. Multitask learning. *Machine Learning*, 28:41–75, 1997.

[2] X. Huang and J. Weng. Locally balanced incremental hierarchical discriminant regression. In *Fourth International Conference on Intelligent Data Engineering and Automated Learning*. Hong Kong, 2003.

[3] W. Hwang and J. Weng. Hierachical discriminant regression. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(11):1277 –1293, 2000.

[4] M. Lovett, L. Daily, and L. Reder. A source activation theory of working memory: Cross-task prediction of performance in act-r. *Cognitive Systems Research*, 1:99–118, 2000.

[5] L. Y. Pratt, J. Mostow, and C. A. Kamm. Direct Transfer of Learned Information among Neural Networks. In *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI-91)*, pages 584–589. AAAI, July 1991.

[6] S. Thrun and J. O'Sullivan. Clustering learning tasks and the selective cross-task transfer of knowledge. In S. Thrun and L. Pratt, editors, *Learning To Learn*. Kluwer Academic Publishers, 1998.

[7] J. Weng and S. Chen. Vision-guided navigation using SHOSLIF. *IEEE Trans. Neural Networks*, 11:1511–1529, 1998.

[8] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291:599–600, 2000.

Developmental Connectivity Schemes and Their Performance Implications

A. Felch
UCI, afelch@uci.edu

R.H.Granger
UCI, granger@uci.edu

Abstract

The development of connectivity between brain networks (e.g., thalamo-cortical, cortico-thalamic, cortico-cortical) proceeds via a combination of axon and dendrite growth. Connectivity tends to be extremely sparse; it has been estimated that the probability of contact between two neocortical excitatory cells that are 0.2-0.3 mm apart is less than 0.1, and between two such cells that are more than 1mm apart, p < 0.01 (Braitenberg & Schuz, 1998). When one group or layer of neurons (A) generates a set of projections to another (B), interesting computational constraints can be educed as a function of characteristics of the originating and receiving networks. First some conditions are described that are clearly undesirable (e.g., if any cells in A produce no contacts in B, part of the "signal" from A presumably cannot be transmitted to B). Remaining conditions include a number of distinct cases with different information-theoretic utility, suggesting differential value for certain connectivity schemes over others. We characterize the tradeoffs among utility and costs and their dependence on different classes of strategies by which axons from A are assigned to dendrites in B. It is shown that hypergeometric distributions optimize a range of measures of these costs as compared to competing projection distributions.

Exact Inference in Robots Using Topographical Uncertainty Maps

Josh Susskind, John Hershey, Javier Movellan

Abstract

Many problems in social robotics require a real-time combination of incoming sensor information and prior information about likely behaviors of the objects in the world. For example, tactile, visual and acoustic information may all inform a distribution of beliefs about the location of humans with whom the robot may want to interact. When sensory information is not available, the uncertainty in this distribution should increase in a principled manner to reflect the fact that people are not static objects.

Bayesian filtering provides a principled approach to solve these problems and has thus become a method of choice in robotics. Most of the Bayesian filtering methods applied to robotics rely on analog hypothesis spaces and find approximate solutions to the resulting non-linear filtering problem using Monte-Carlo approximations (i.e., particle filters). Unfortunately, particle-filters tend to be very inefficient, thus greatly limiting the applicability of the approach. We propose an alternative approach based on digitizing the hypothesis space into a large number of hypotheses (on the order of 100,000). The approach has not been tried in the past because, in principle, solving the filtering equations requires order n-squared operations per time step, where n is the number of hypotheses. This means that as the hypothesis space expands, solving the filtering equations becomes rapidly prohibitive.

We show that in many problems, one can make use of the spatial-temporal structure of the hypothesis space and we propose an algorithm to solve the filtering operations in order n operations, vastly reducing the computational strain on the system. In practice, this allows handling hundreds of thousands of hypothesis in real time. We illustrate how the algorithm works for the problem of tracking human faces in real time. In this problem, possible object locations and scales (states) arrange in a three-dimensional topology (two dimensions for location and one for scale). Rectangular convolution kernels capture movement uncertainty over scales and locations. Interestingly the resulting architecture resembles the functional architecture of primary visual cortex, suggesting explanations for the computational role of forward, lateral and top-down connections in V1.

RobotCub: An Open Research Initiative in Embodied Cognition

Giulio Sandini, Giorgio Metta, and David Vernon*
DIST, University of Genova, Italy
sandini@dist.unige.it, pasa@dist.unige.it, vernon@ieee.org

We describe a new research initiative in embodied cognition that will create and exploit a 54 degree-of-freedom humanoid robot. It has the two-fold goal of (1) creating an open and freely-available humanoid platform – RobotCub – for research in embodied cognition, and (2) advancing our understanding of cognitive systems by exploiting this platform in the study of cognitive development.

We plan to construct an embodied system able to learn: i) how to interact with the environment by complex manipulation and through gesture production and interpretation; and ii) how to develop its perceptual, motor and communication capabilities for the purpose of performing goal-directed manipulation tasks.

The design of the humanoid robotic platform – RobotCub – is presently at its incipit. The final system will be made freely available to the scientific community through an open systems GNU-like general public license together with any software developed within this research initiative. RobotCub will have physical size and form similar to that of a two year-old child.

In addition, the project will further a research agenda in cognitive systems centered on manipulation in its widest sense including exploration, manipulation of objects, imitation, and communication through gestures. This agenda borrows heavily from experience in developmental psychology and cognitive neuroscience. The project includes interdisciplinary collaboration between neuroscientists and developmental psychologists on one side and roboticists and computer scientists on the other.

Our guiding philosophy – and the motivation for creating RobotCub – is that cognition cannot be hand-coded but has to be the result of a developmental process through which the system becomes progressively more skilled and acquires the ability to understand events, contexts, and actions, initially dealing with immediate situations and increasingly acquiring a predictive capability.

The RobotCub approach rests on three pillars: (1) its scientific stance on cognition: that cognition emerges through embodied development, (2) its research methodology: that cognition is best studied through a programme of progressive development, and (3) its research strategy: that progress in the global scientific community is best served by creating an open systems platform and by exploiting consequent synergies in that community.

To enable the investigation of relevant cognitive aspects of manipulation the design of the robot will be aimed at maximizing the number of degrees of freedom of the upper part of the body (head, torso, arms, and hands). The lower body (legs) will be designed to support crawling on four legs and sitting on the ground in a stable position with smooth autonomous transition from crawling to sitting. This will allow the robot to explore the environment and to grasp and manipulate objects on the floor. The sensory system will include a binocular vision system, touch, audition, and inertial sensors. Functionally, the system will be able to coordinate the movement of the eyes and hands, grasp and manipulate lightweight objects of reasonable size and appearance, crawl on four legs and sit.

Finally, we wish to emphasize again that one the principal goal of this initiative is to help foster the study of embodied cognition throughout the global research community by making the RobotCub humanoid and cognitive software freely available.

The content of this paper is based on the work of several people: Paolo Dario, Scuola S. Anna, Pisa; Rolf Pfeifer, University of Zurich; Claes von Hofsten, University of Uppsala; Luciano Fadiga, University of Ferrara; Kerstin Dautenhahn, University of Hertfordshire; Jose Santos-Victor, IST Lisbon; Darwin Caldwell and John Gray, University of Salford; Aude Billard and Auke Ijspeert, Ecole Polytechnique Federal de Lausanne; Francesco Becchi, Telerobot S.r.l.; Emilio Bizzi, European Brain Research Institute. Their contributions are gratefully acknowledged.

*David Vernon is on sabbatical leave at Etisalat University, UAE.

This work is funded by the European Commission's newly-established Cognition Unit, Directorate-General Information Society, as part of project no. IST-2004-004370: RobotCub — ROBotic Open-architecture Technology for Cognition, Understanding, and Behaviour.

Why do animals make their play more difficult?

Stan Kuczaj

Play is common in the young of many species, and it seems likely that play influences development and learning. One way in which play may influence learning is in the provision of contexts within which an organism can safely experience and benefit from moderately discrepant events. Such events are believed to facilitate learning because they consist of challenges that the organism can resolve, the successful resolution of the challenges resulting in cognitive growth. The possibility that play is linked to moderately discrepant events is explored by considering the play behavior of young dolphins and killer whales. Our systematic observations of these animals demonstrate that they consistently modify their play behavior to make the goal more difficult to achieve, demonstrating that the play activity is at least as important as the play outcome. This suggests that the animals are purposely producing their own moderately discrepant events, and that one of the functions of play is to provide cognitive stimulation. Such stimulation might result from an organism's own activities as well as the activities of play partners. Thus, play may have evolved to facilitate cognitive development in both solitary and social animals.

Social Dynamics: Signals and Behavior

Alex Pentland
MIT Media Lab, E15-387, 20 Ames St, Cambridge MA 02139

Abstract

Nonlinguistic social signals (e.g., `tone of voice') are often as important as linguistic content in predicting behavioural outcomes [1,2]. This paper describes four automated measure of such social signalling, and shows that they can be used to form powerful predictosr of objective and subjective outcomes in several important situations. Finally, it is argued that such signals are important determinants of social position.

1. Introduction

Animals communicate their social structure in many ways, including dominance displays, relative positioning, access to resources, etc. Humans add to that a wide variety of cultural mechanisms such as clothing, seating arrangements, and name-dropping. Most of these culture-specific social communications are conscious and are often manipulated.

However in many situations non-linguistic social signals (body language, facial expression, tone of voice) are as important as linguistic content in predicting behavioral outcome [1,2]. Tone of voice and prosodic style are among the most powerful of these social signals even though (and perhaps because) people are usually unaware of them [2]. In a wide range of situations (marriage counseling, student performance assessment, jury decisions, etc.) an expert observer can reliably quantify these social signals and with only a few minutes of observation predict about 1/3d of the variance in behavioral outcome (which corresponds to a 70% binary decision accuracy) [1]. It is astounding that observation of social signals within such a `thin slice' of behavior can predict important behavioral outcomes (divorce, student grade, criminal conviction, etc.) when the predicted outcome is sometimes months or years in the future.

Nonlinguistic vocal signaling is a particularly familiar part of human behavior. For instance, we speak of someone `taking change' of a conversation, and in such a case this person might be described as `driving the conversation' or `setting the tone' of the conversation. Such dominance of the conversational dynamics is popularly associated with higher social status or a leadership role. Similarly, some people seem skilled at establishing a `friendly' interaction. The ability to set conversational tone in this manner is popularly associated with good social skills, and is typical of skilled salespeople to social `connectors' [3].

The machine understanding has studied human communication at many time scales --- e.g., phonemes, words, phrases, dialogs --- and both semantic structure and prosodic structure has been analyzed. However the sort of longer-term, multi-utterance structure associated with social signaling has received relatively little attention [4]. In this paper I develop an automatic measurement method for quantifying some of these non-linguistic social signals, and describe how these measurements can be used to form powerful predictors of behavioral outcome in some very important types of social interaction: getting a date, getting a job, and getting a raise.

2. Measuring Social Signals

I have constructed measures for four types of vocal social signaling, which I have designated activity level, engagement, stress, and mirroring. These four measures were extrapolated from a broad reading of the voice analysis and social science literature, and we are now working to establish their general validity. To date they have been used to predict outcomes in salary negotiation, dating, friendship, and business preferences with accuracy comparable to that of human experts in analogous situations.

Calculation of the activity measure begins by using a two-level HMM to segment the speech stream of each person into voiced and non-voiced segments, and then group the voiced segments into speaking vs non-speaking [5]. Conversational activity level is measured by the z-scored percentage of speaking time plus the frequency of voiced segments.

Engagement is measured by the z-scored influence each person has on the other's turn-talking. When two people are interacting, their individual turn-taking dynamics influences each other and can be modeled as a Markov process [6]. By quantifying the influence each participant has on the other we obtain a measure of their engagement...popularly speaking, were they driving the conversation? To measure these influences we model their individual turn-taking by an Hidden Markov Model (HMM) and measure the coupling of these two dynamic systems to estimate the influence each has on the others' turn-taking dynamics [7]. Our method is similar to the classic method of Jaffe et al. [6], but with a simpler parameterization that permits the direction of influence to be calculated and permits analysis of conversations involving many participants.

Stress is measured by the variation in prosodic emphasis. For each voiced segment we extract the mean energy, frequency of the fundamental format, and the spectral entropy. Averaging over longer time periods provides estimates of the mean-scaled standard deviation of the energy, formant frequency and spectral entropy. The z-scored sum of these standard deviations is taken as a measure speaker stress; such stress can be either purposeful (e.g., prosodic emphasis) or unintentional (e.g., physiological stress caused by discomfort).

Mirroring behavior, in which the prosody of one participant is `mirrored' by the other, is considered to signal empathy, and has been shown to positively influence the outcome of a negotiation [8]. In our experiments the distribution of utterance length is often bimodal. Sentences and sentence fragments typically occurred at several-second and longer time scales. At time scales less than one second there are short interjections (e.g., `uh-huh'), but also back-and-forth exchanges typically consisting of single words (e.g., `OK?', `OK!', `done?', `yup.'). The z-scored frequency of these short utterance exchanges is taken as a measure of mirroring. In our data these short utterance exchanges were also periods of tension release.

2.1. Signaling Dynamics

These measures of social signaling can be computed on a conventional PDA in real-time, using a one-minute lagging window during which the statistics are accumulated. It is therefore straightforward to investigate how these `social signals' are distributed in conversation. In [9] we analyzed social signaling in 54 hours of two-person negotiations (described in more detail in the next section) on a minute-by-minute basis. We observed that high numerical values of any one measure typically occur by themselves, e.g., during periods in which participants showed high engagement they did not use high stress, etc., so that each participant exhibits four `social display' states, plus a `neutral' relaxed state in which the participant is typically asking neutral questions or just listening. The fact that these display states were largely unmixed provides evidence that they are measuring separate social displays.

The signaling state of the two participants was strongly coupled, so that (ignoring symmetries and outliers) the joint state space has only six states rather than the expected fifteen. For instance, when one participant displayed engagement, the other participant almost always followed suit (90% of the time), resulting in a highly engaged, roughly equal conversation. When one participant displayed mirroring behavior, the other would usually join in (74% of the time). When one participant became active, the other became neutral (75% of the time). However when one participant used stress, the result differed according to status. If the high status participant used stress, then low-status participant would usually (66% of the time) signal activity and only 11% of the time would the low-status participant also show stress. When the low-status participant used stress, the high-status participant would usually become active (54% of the time) but 24% of the time would respond with matching stress.

2.2. Negotiation Experiment

In this experiment we investigated what might be thought to be a prototypically rational form of communication: negotiating a salary package with your boss. The intuition is that negotiation participants who `take change' of the dynamics of the conversation, what might be described as `driving the conversation' will do better than those who are more passive.

In Pentland, Curhan, et al [10] we collected audio from forty-six gender-matched dyads (either male/male or female/female, 28 male dyads and 18 female dyads) that were asked to conduct a face-to-face negotiation as part of their class work. The mock negotiation involved a Middle Manager (MM) applying for a transfer to a Vice President's (VP) division in a fictitious company. Many aspects of the job were subject to negotiation including salary, vacation, company car, division, and health care benefits; these aspects were summed into an overall objective score based on their market value. Participants were offered a real monetary incentive for maximizing their own individual outcome in the negotiation. Subjects were first year business students at MIT Sloan School of Management, almost all with previous work experience.

Data collected included individual voice recordings of both parties in a closed room plus ratings of subjective features. There was no time limit set and the negotiations length ranged from 10 to 80 minutes in length, with an average duration of approximately 35 minutes, for a total of 54 hours of data.

Subjective features analyzed were the answers to the questions `What kind of impression do you think you made on your counterpart?' `To what extent did your counterpart deliberately let you get a better deal than he/she did?' and `To what extent did you steer clear of disagreements?

2.2.1. Results

Our hypothesis was that negotiation participants who showed the most engagement, stress and mirroring would do better than those who were more passive, i.e., that the time-averaged influence on each participant + amount of stress + amount of mirroring would predict the objective outcome of the negotiation. Following [1], we measured signaling in only the first five minutes of the negotiation and used that `thin slice' of behavior to predict the final negotiation outcome.

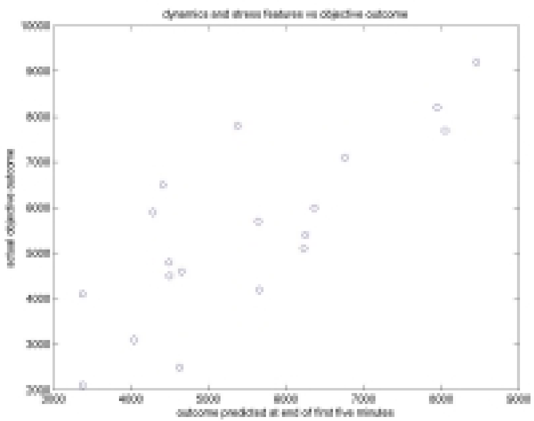


Figure 2: Outcome predicted from social signals at end of first five minutes of negotiation. (Female VPs shown).

This predictor had a strong (r= 0.57, p=0.001) correlation with the objective outcome of the negotiation. Thus the accuracy of this predictor is similar to that of human experts performing similar tasks [1].

Post-hoc analysis showed that the relationship differed for high- and low-status participants. For VPs, engagement + stress predicted almost half of their variation in outcome (r=0.75). For MMs, the mirroring measure alone predicted almost a third of the variation in their objective outcome (r=0.57).

The engagement measure had a significant positive correlation (r=0.63) with the subjective "impression I thought I made on my partner" rating and a with the "did your partner let you win" rating (r=0.65). The mirroring measure had a significant positive correlation with the extent to which participants said they were seeking to avoid disagreements (r=0.62).

2.3. Attraction Experiment

Speed dating is relatively new way of meeting many potential matches during an evening. Participants interact for five minutes with their `date', at the end of which they decide if they would like to provide contact information to him/her, and then they move onto the next person. A 'match' is found when both singles answer yes, and they are later provided with mutual contact information.

In Madan, Caneel and Pentland [11] we analyzed 57 five-minute speed-dating sessions. In addition to the `romantically attracted' question (where a postive answer from both participants resulting in sharing of contact information), participants were also asked two other yes/no questions: would they like to stay in touch just as friends, and would they like to stay in touch for a business relationship. These `stay in touch' questions were hypothetical, since contact information would not be exchanged in any case, but allowed us to explore whether vocal signals of romantic attraction could be differentiated from other types of attraction.

2.3.1. Results

Liniar regression was used to form predictors of the question responses using the values of the four social signaling measures. For each question the resulting predictor could account for more than 1/3rd of the variance, providing approximately 70% accuracy at predicting the questions response. This accuracy is comparable to that of human experts performing similar tasks [1].

For the females responses, for instance, the correlation with the ‘attracted’ responses were r=0.66, p=0.01, for the ‘friendship’ responses r=0.63, p=0.01, and for the ‘business’ reponses r=0.7, p=0.01. Corresponding values for the male responses were r=0.59, r=0.62, and r=0.57, each with p=0.01.

For the ‘attracted’ question the most predictive individual feature was the female activity measure. The engagement measure was the most important individual feature for predicting the ‘friendship’ and ‘business'’ responses. The mirroring measure was also significantly correlated with female ‘friendship’ and ‘business’ ratings, but not with with male ratings.

An interesting observation was that for the ‘attracted'’ question female features alone showed far more correlation with both male (r=0.5,p=0.02) and female (r=0.48, p=0.03) responses than male features (no significant correlation). In other words, female social signaling is more important in determining a couples ‘attracted’ response than male signaling.

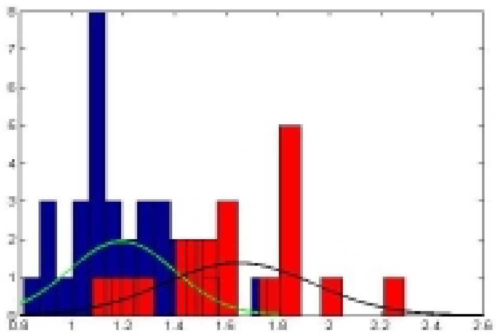


Figure 3: Frequency of female `attracted' responses (black=no) vs. predictor value. The cross-validated linear decision rule produces 71% accuracy.

Figure 3 illustrates a two-class linear classifier based on the social signaling measures; this classifier has a cross-validated accuracy of 71% for predicting the `attracted' response. The two fitted Gaussians are simply to aid visualization of the distributions' separability.

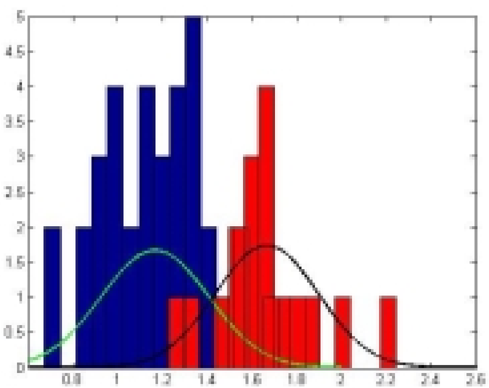


Figure 4. Frequency of female `business' responses (black=no) vs. predictor value. The cross-validated linear decision rule produces 74% accuracy..

Figure 4 illustrates a two-class linear classifier for the `business' responses, based on the social signaling measures; this classifier has a cross-validated accuracy of 74%. The two fitted Gaussians are simply to aid visualization of the distributions' separability.

2.4. Social Network Experiment

In Choudhury and Pentland [7] we collected audio data from 23 subjects from 4 different research groups over a period of 11 days, resulting in an average of 66 hours of data per subject. The subjects were a representative sample of the community, including students, faculty and administrative staff. During data collection users had a small audio recording device on them for six hours a day (11AM -5PM) while they were on the MIT campus. The data was automatically analyzed to detect the pair-wise conversations, and this was used to analyze the distribution of conversational statistics within the sampled conversations, including the engagement and activity measures

2.4.1. Results

Our first finding was that the Markov statistics describing individuals' turn-taking styles are distinctive and stable across different conversational partners, and that these turn-taking patterns are not just a noisy variation of the same average style ($p < 0.001$). Since these Markov statistics effectively determine the average values of the activity and engagement measures, the implication is that people have characteristic patterns of activity and engagement signaling.

Male and female patterns were extremely different, with only slight overlap between the range of parameters observed for males and those observed for females. Surprisingly, the total speaking time for males and females was nearly equal.

Choudhury [12] in her PhD thesis investigated the hypothesis that the `engagement' measure (e.g., the inter-Markov process influence parameter) would be correlated with information flow within their social network structure. To investigate this hypothesis she compared the engagement measure to the individual subject's betweeness centrality, which is a standard social science measure of how important an individual is to information flow within a social network [13]. The correlation value between this centrality measure and the influence parameter was 0.90 (p-value < 0.0004, rank correlation 0.92). Thus the amount of time an individual displayed engagement was a nearly perfect predictor of how much of a `connector' they were.

3. Discussion

In this paper I have proposed a method of measuring social signaling using non-linguistic vocal features, and shown that the measured signaling can be used to create powerful predictors of both objective and subjective outcome in social situations. In addition, at least some aspects of people's position in a social network appears to be signaled and negotiated via this same mechanism.

In our negotiation experiment we showed that signaling during the first five minutes of a negotiation account for more than 1/3d of the variation in objective outcome, and that the `winning' strategy is different for high-status vs low-status participants. For the high-status participants engagement and use of stress was most important. For low-status participants the use of mirroring was most important. The correlation between the engagement measure and the subjective questions concerning control, and between the mirroring measure and the subjective question concerning cooperation, support the validity of these features as measures of social signaling.

In our attraction experiment we showed that signaling during the first five minutes of conversation again accounted for more than 1/3d of the variation in outcome, and that the signaling and interaction is different for male and female participants. The high correlation between the female activity measure and the `attraction' responses aupports the validity of this feature as a measure of social signaling.

We found that in a research laboratory environment peoples' signaling mirrored the information flow within the social network. The more a person was a `connector' within the social network, the more they displayed engagement. This suggests that social signals are part of a continual, implicit `negotiation' between members of a social network that establishes each individual's appropriate position in the network.

Is this social signaling just a part of `normal' speaking prosody? Prosody is most commonly studied within the framework of speech understanding, where pitch, duration, and amplitude are used to modify, select, or emphasize the semantics conveyed by the words [2,4]. In contrast to this type of prosody the vocal features measured in these experiments occur at time scales that are far too long to be related to individual words or phrases.

The social signaling discussed in this paper instead seems to communicate and be involved in mediating social variables such as status, interest, determination, or cooperation, and arise from the interaction of two or more people rather than being a property of a single speaker. Semantics and affect are important in determining what signaling an individual will engage in, but they seem to be fundamentally different types of phenomina. The social signaling measured here seems to be a sort of `vocal body language' that operates relatively independently of linguistic or affective communication channels.

Finally, it is interesting that people in these experiments were only vaguely aware of their own vocal characteristics, and they were unable to articulate the connection between these characteristics and the behavioral outcome. It is interesting to speculate about what might happen if people were made more aware of their social signaling. One idea is to construct a small wearable `social signaling' meter that could provide users with real-time feedback. We are now beginning tests with such a meter and expect to be able to report the results by the time of the conference.

References

[1] Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. Psychological Bulletin, 111(2), 256-274.

[2] Nass, C., and Brave, S. (2004) Voice Activated: How People Are Wired for Speech and How Computers Will Speak with Us, MIT Press

[3] Gladwell, M. (2000) The Tipping Point: How little things can make a big difference. New York: Little Brown

[4] Handel, Stephen, (1989) Listening: an introduction to the perception of auditory events, Stephen Handel , Cambridge: MIT Press

[5] Basu, B., (2002) Conversational Scene Analysis, doctoral thesis, Dept. of Electrical Engineering and Computer Science, MIT. 2002. Advisor: A. Pentland

[6] Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. (2001). Rhythms of dialogue in early infancy. Monographs of the Society for Research in Child Development, 66(2), No. 264.

[7] Choudhury, T., and Pentland, A., (2004), NAASCOS, June 27-29, Pittsburgh, PA. PDF available at http://hd.media.mit.edu

[8] Chartrand, T., and Bargh, J., (1999) The Chameleon Effect: The Perception-Behavior Link and Social Interaction, J. Personality and Social Psychology, Vo. 76, No. 6, 893-910

[9] Khilnani, R. (2004) Temporal Analysis of Stages in Negotiation, MEng Project, Advisor: A. Pentland.

[10] Pentland, A., Curhan, J., Khilnani, R., Martin, M., Eagle, N., Caneel, R., Madan A (2004) ``Toward a Negotiation Advisor," UIST 04, Oct 24-27, ACM. PDF available at http://hd.media.mit.edu

[11] Madan, A., Caneel, R., and Pentland, A. (2004) GroupMedia: Distributed Multimodal Interfaces, ICMI, St. College, PA, Oct. 12-14, 2004. IEEE Press, PDF available at http://hd.media.mit.edu

[12] Choudhury, T., (2003) Sensing and Modeling Human Networks, Ph.D. Thesis, Dept. of Media Arts, and Sciences, MIT. Advisor: A. Pentland.

[13] Wasserman, S. and K. Faust, Social Network Analysis Methods and Applications. 1994: Cambridge University Press

Small-world Network Properties and the Emergence of Social Cognition: Evidence from Functional Studies of Autism

Matthew K Belmonte and Simon Baron-Cohen
Autism Research Centre
Departments of Psychiatry and Experimental Psychology
University of Cambridge

Abstract

Autism has been proposed to stem from a pervasive abnormality of neural information processing possibly involving decreased signal-to-noise in neural systems [6]. Such a final common pathway of dysfunction at the network level may admit a wide fan-in of causal factors at the level of cells, molecules, and genes, and could be expected to interact with normal neurodevelopmental programmes and gradients to fan out into a wide range of abnormalities of cognition and behaviour [1]. Autism may therefore serve as a useful test case for theories of normal brain and cognitive development at the junction of network modelling with developmental cognitive neuroscience. Behavioural and cognitive findings on autism have been dominated by a combination of facilitated processing of local features and impairment in complex, integrative cognition. These behavioural results are supported by functional imaging and quantitative EEG studies that find abnormally strong, unselective responses within sensory regions and abnormally weak functional correlations between distant regions (*e.g.* [5]). One possible explanation for this pattern is abnormal strengthening of local network connectivity at the expense of global connectivity – a combination that may sabotage the 'small-world' property [7] in which cortical networks combine strong local connectivity with short mean path length. Indiscriminately strong anatomical connectivity within local processing regions may actually decrease computational connectivity, if crosstalk is increased to a degree at which relevant stimuli cannot be discriminated from noise [2]. Network abnormalities of this sort may explain an autistic learning style founded on statistical association rather than on instructive focus on relevant stimuli. Aberrant patterns of cortical connectivity are suggested by reports of decreased size of cortical minicolumns [3], early developmental hyperplasia in short-range white matter but not longer-range white matter compartments [4], associations with neuroligins and other substances involved in synaptogenesis and synaptic modification, and comorbidity with disorders that involve abnormal excitability or alterations in synaptic structure, such as epilepsy and Fragile X syndrome. We describe fMRI findings from two visual selective attention tasks that bear on this issue of neural connectivity in the autistic brain. In Experiment 1, adults with autism spectrum disorders (ASDs) and normal controls covertly attended to one or the other stream of rapid serially presented colour stimuli in left and right hemifields, shifting attention in response to cues within the attended stream. In Experiment 2, 10-to-15-year-old boys with ASDs, their non-autistic brothers, and unrelated normal controls performed a cognitively demanding task combining selective attention to location, colour, and orientation. These studies revealed findings consistent with abnormalities of local and long-range connectivity, including (1) abnormally strong and unselective processing within visual areas of cortex, (2) abnormal deactivation of more anterior, integrative brain regions including inferior frontal gyrus, anterior cingulum, frontal pole, and medial temporal lobe, and (3) abnormally low functional correlations between anterior and posterior regions. In addition, other results may reflect cognitive strategies that attempt to compensate for abnormal excitability of sensory cortices, including (4) abnormally strong activation of parietal cortex during suppression of visual distractors and (5) deactivation of auditory cortex during difficult visual discriminations. Preliminary data suggest that non-autistic sibs share some of the anterior deactivations but none of the posterior hyperactivations; sibs are a useful contrast in disentangling the more subtle, primary effects of genetic susceptibilities from the complicated, secondary dysfunctions associated with the full syndrome of autism. Although this hypothesis of autism as a developmental effect of abnormal network properties remains speculative in the absence of any large-scale microanatomical studies directly examining autistic brains across anatomical regions and developmental periods, it provides a useful entry point for the study of social cognition, a complex capacity whose proper development may depend on a balance between connections that subserve local processing and global integration.

References

[1] M.K. Belmonte, E.H. Cook Jr., G.M. Anderson, J.L.R. Rubenstein, W.T. Greenough, A. Beckel-Mitchener, E. Courchesne, L.M. Boulanger, S.B. Powell, P.R. Levitt, E.K. Perry, Y. Jiang, T.M. DeLorey and E. Tierney, "Autism as a disorder of neural information processing: directions for research and targets for therapy," *Mol. Psychiatry,* vol. 9, no. 7, pp. 646-663, 2004. Unabridged edition at `http://www.cureautismnow.org/media/3915.pdf`

[2] M.K. Belmonte, G. Allen, A. Beckel-Mitchener, L.M. Boulanger, R.A. Carper and S.J. Webb, "Autism and abnormal development of brain connectivity," *J. Neurosci.,* in press 2004.

[3] M.F. Casanova, D.P. Buxhoeveden, A.E. Switala and E. Roy E, "Minicolumnar pathology in autism," *Neurology,* vol. 58, no. 3, pp. 428-432, 2002.

[4] M.R. Herbert, D.A. Ziegler, N. Makris, P.A. Filipek, T.L. Kemper, J.J. Normandin, H.A. Sanders, D.N. Kennedy, V.S. Caviness Jr., "Localization of white matter volume increase in autism and developmental language disorder," *Ann. Neurol.,* vol. 55, no. 4, pp. 530-540, 2004.

[5] M.A. Just, V.L. Cherkassky, T.A. Keller, N.J. Minshew, "Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity," *Brain,* in press 2004.

[6] J.L.R. Rubenstein and M.M. Merzenich, "Model of autism: increased ratio of excitation/inhibition in key neural systems," *Genes Brain Behav.,* vol. 2, no. 5, pp. 255-267, 2003.

[7] O. Sporns, G. Tononi and G.M. Edelman, "Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices," *Cereb. Cortex,* vol. 10, no. 2, pp. 127-141, 2000.

THE EMOTIONAL BRAIN IN AUTISM: EFFECTIVE CONNECTIVITY ANALYSIS OF BRAIN REGIONS INVOLVED IN EXPLICIT FACIAL EMOTION PROCESSING

B.Wicker*, B.Hubert*, P. Fonlupt[∞], B.Gepner°, C.Tardif°, C.Deruelle*.
*INCM, CNRS, Marseille, France.
°Hôpital Montperrin, Aix en Provence, France.
∞ INSERM U280, Lyon, France.

Email : bwicker@lnf.cnrs-mrs.fr

Although high-functioning individuals with autistic disorder (i.e. autism and Asperger syndrom) are of normal intelligence, they have life-long abnormalities in social communication and emotional behavior. However, the biological basis of socio-emotional difficulties in autism is still poorly understood. We investigated if high-functioning people with autistic disorder show neurobiological differences from controls when they explicitly or implicitly attributed hostile or friendly intentions to actors who directed their gaze towards or away from them.

Using fMRI, 14 healthy subjects and 12 high functioning patients were scanned while viewing video sequences of actor's faces displaying either happy or angry expressions. In 2 conditions subjects were required to actively engage in interpreting emotions from the perceived faces, whereas in 2 other conditions subjects had to judge the actor's age. Statistical random effect analysis using SPM99 examined contrasts between conditions. Effective connectivity using Structural Equation Modeling (SEM) was then performed on the data of each population.

Although occipital, temporo-ventral and right posterior superior temporal cortex were found commonly activated in both population, subjects with autistic disorders differed from controls in the activity of dorso-medial (DMPFC) and right lateral prefrontal (RLFPC) cortical areas when explicitly processing emotional expressions. Interestingly, effective connectivity analysis revealed a distinct pattern of interactions between brain regions of this functional network in the two populations. Whereas controls specifically exhibit functional relations between fusiform gyrus, superior temporal sulcus, DMPF and RLPFC, patients with autism do not (Fig 1.). These results suggest that high-functioning people with autistic disorders have biological differences from controls when processing emotional informations, and that these differences are most likely due to abnormal functional relations between brain structures known to be involved in social cognition.

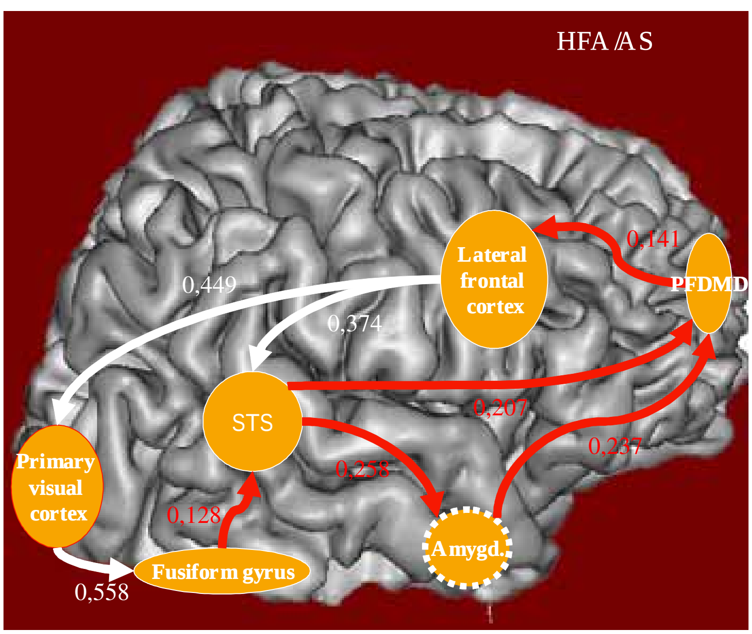


Fig. 1. Results of effective connectivity modeling on the data from the autistic population.

Pathological brain growth patterns in Autism, and catastrophic interference in establishing long-distance connectivity

John D. Lewis
Cognitive Science, UC San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0515, jdlewis@cogsci.ucsd.edu

Jeffrey L. Elman
Cognitive Science, UC San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0515, jelman@ucsd.edu

Abstract

The clinical onset of autism has recently been found to be preceded by a period of abnormally accelerated brain growth [1], apparently following delayed prenatal development. Findings from comparative neuroanatomy [2-4] and developmental neuroscience [5, 6] motivate the hypothesis that this growth trajectory will give rise to abnormalities in the development of cortico-cortical connectivity. Conduction delay is proportional to brain size [7], and so developmental growth abnormalities provide a force that should influence the development of cortical networks. The growth trajectory seen in children with autism should cause an initial increase in long-distance connectivity, and then a reduction. And the more extreme the under and overgrowth, the greater should be both the initial increase and the subsequent reduction. The loss of connectivity from birth to adolescence may thus be far more deviant than is revealed by comparisons of the connectivity that remains.

Substantial learning of linguistic and social relevance takes place during the last months of prenatal development and in the first few postnatal months [8]. A great deal of what is learned during this period is represented in neural assemblies that make substantial use of long-distance connections. According to our hypothesis, the prenatal growth trajectory in autism will initially lead to even greater than normal reliance on long-distance connectivity; but the conduction delay associated with the postnatal brain overgrowth will cause the postnatal loss of these connections. The changes in connectivity may thus explain the regression in functional development often reported in children with autism, and their subsequent failure to progress [9, 10].

To explore this hypothesis, neural networks which modelled interhemispheric interaction were grown at the rate of either typically developing children or severely autistic children. The long-distance connections were lesioned at `birth', restored, and lesioned again at `age 4'. The networks that modelled autistic growth were more effected by the lesions at `birth' — indicating a greater reliance on the long-distance connections — and those that modelled typical development were more effected by the lesions at `age 4'.

References

[1] E. Courchesne, R. Carper, and N. Akshoomoff, "Evidence of brain overgrowth in the first year of life in autism," *JAMA*, vol. 290, pp. 337-344, 2003.
[2] J. L. Ringo, R. W. Doty, S. Demeter, and P. Y. Simard, "Time is of the essence: a conjecture that hemispheric specialization arises from interhemispheric conduction delay," *Cerebral Cortex*, vol. 4, pp. 331-343, 1994.
[3] J. K. Rilling and T. R. Insel, "Differential expansion of neural projection systems in primate brain evolution," *neuroReport*, vol. 10, pp. 1453-9, 1999.
[4] R. Olivares, J. Montiel, and F. Aboitiz, "Species differences and similarities in the fine structure of the mammalian corpus callosum," *Brain, Behavior and Evololution*, vol. 57, pp. 98-105, 2001.
[5] A. S. LaMantia and P. Rakic, "Axon Overproduction and Elimination in the Corpus Callosum of the Developing Rhesus Monkey," *Neuroscience*, vol. 10, pp. 2156-2175, 1990.
[6] P. M. Thompson, J. N. Giedd, R. E. Blanton, C. Lindshield, R. P. Woods, D. MacDonald, A. C. Evans, and A. W. Toga, "Growth patterns in the developing brain detected by using continuum mechanical tensor maps," *Nature*, vol. 404, pp. 190-193, 2000.
[7] S. G. Waxman, "Conduction in myelinated, unmyelinated, and demyelinated fibers," *Archives of Neurology*, vol. 34, pp. 585-589, 1977.
[8] J. L. Elman, E. A. Bates, M. H. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett, *Rethinking Innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press, 1996.
[9] L. Kanner, "Autistic disturbances of affective contact," *Nervous Child*, vol. 2, pp. 217-250, 1943.
[10] C. Lord, C. Shulman, and P. DiLavore, "Regression and word loss in autistic spectrum disorders," *Journal of Child Psychology and Psychiatry*, vol. 45, pp. 936-955, 2004.

Project PRAKÄSH
Development of object perception following long-term visual deprivation

Pawan Sinha
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139
<psinha@mit.edu>

Abstract

Understanding how the human visual system learns to perceive objects in the environment is one of the fundamental challenges in neuroscience, and is also the motivation behind a new humanitarian and scientific initiative that we have launched in India, called 'Project Prakash'. This project involves a systematic study of the development of object perception skills in children following recovery from congenital blindness. Here we provide an overview of Project Prakash and also describe a specific study related to the development of face-perception skills following sight recovery. A few studies have reported profound impairments in face recognition following early visual deprivation. However, it is unknown how visual deprivation influences performance on the more basic task of face versus non-face classification. Here we report studies with two children, both of whom suffered from congenital blindness lasting at least the first 7 years of life. We assessed their face classification skills following surgical restoration of sight. For one child, the experiments were performed 1.5 months after surgery and for the other, four years post-surgery. Our results indicate that these children are able to detect faces and distinguish them from distracters with high reliability, comparable to control subjects. Furthermore, this ability appears to be based on the use of overall facial configuration rather than individual features – a finding that presents an interesting contrast to the hypothesis of piecemeal processing used to explain impairments in face identification. These results have implications for the nature of face-concept learning schemes in human and computational vision systems.

1. Introduction to Project Prakash

Through a process of extensive and continuous exposure, the brain comes to be able to parse complex visual scenes into distinct objects. Several issues about this process remain open. How much visual experience is needed for the development of this ability? What are the intermediate stages in the evolution of object representations? How critical is early visual experience for the development of object perception?

There are two dominant approaches for studying these questions: 1. experimentation with infants, and 2. experiments with adults using novel objects. These approaches have yielded valuable results, but their usefulness is limited by some significant shortcomings. For instance, infant experiments are operationally difficult and the development of object perception processes is confounded with the development of other brain subsystems such as those responsible for attention deployment and eye-movement control. Experiments with adults, on the other hand, are necessarily contaminated by the subjects' prior visual experience, even though the objects used as stimuli may be novel.

We have identified a unique population of children in India that allows us to adopt a very different approach. According to the WHO, India is home to the world's largest population of blind children. While the incidence of congenital blindness in developed nations such as the USA and UK is less than 0.3 per 1000 children, the incidence in India is 0.81/1000. Many of these children have treatable conditions, such as congenital cataracts. However, poverty, ignorance and lack of simple diagnostic tools in rural areas deprive these children of the chance of early treatment. Recently, in response to government initiatives for controlling blindness, a few hospitals have launched outreach programs to identify children in need of treatment and perform corrective surgeries at low cost. These initiatives are beginning to create a remarkable population of children across a wide age-range who are just setting out on the enterprise of learning how to see. We have launched Project Prakash with the goal of following the development of visual skills in these unique children to gain insights into fundamental questions regarding object concept learning and brain plasticity.

Such a population is not available in developed countries such as the United States. Given the extensive network of neonatal clinics and pediatric care in these countries, congenital cataracts are invariably treated surgically within a few weeks after their discovery. Consequently, in the developed world, it is rare to find an untreated case of blindness in a child of more than a few months of age. In India on the other hand, many children with congenital cataracts spend several years, or even their entire lives, without sight. The societal support and quality of life for blind children in India is extremely poor, leading to a life expectancy that is 15 years shorter than that of a sighted child. There is clearly a humanitarian need to help such children get treatment, and a key goal of Project Prakash is to help address this need. Furthermore, in tackling this need, the Project is presented with a unique scientific opportunity.

The scientific goal of Project Prakash is to study the development of low-level visual function (such as acuity, contrast sensitivity and motion perception), as well as object perception following recovery from congenital blindness. We are investigating the time-course of different object-perception skills as assessed behaviorally, the concurrent changes in cortical organization, and also the development of neural markers associated with object-perception. Of special interest to us is face perception, including face localization, identification and expression classification. Few object domains can rival the ecological relevance of faces. Much of the human social infrastructure is critically dependent on face-perception skills. We are studying both the deficiencies and proficiencies of children after onset of sight. The former allow us determine the visual skills that are susceptible to early visual deprivation while longitudinal studies of the latter yield insights about how face-perception develops and what the underlying processes might be.

We call this project 'Prakash', after the Sanskrit word for 'light', symbolizing the infusion of light in the lives of children following treatment for congenital blindness and also the illumination of several fundamental questions in neuroscience regarding brain plasticity and learning.

1.1. The broader impact of Project Prakash

The WHO estimates that the number of blind children globally increases by 500,000 every year. Significant advances have been made in pediatric eye-care to counter this problem. Treatments now exist to restore sight in a significant proportion of the afflicted children, such as those suffering from congenital cataracts. However, merely treating the eyes is not sufficient for ensuring restoration of normal visual function. An equally important requirement for sight recovery is that a child's brain be able to correctly process the visual information, after having been deprived of it for several years. Based on past animal studies of the consequences of visual deprivation on subsequent function [Wiesel and Hubel, 1963; Bauer and Held, 1975; Hubel et al., 1977; LeVay et al., 1980; Jacobson et al., 1981], we can expect that the treated children will exhibit visual deficits relative to normally reared children. However, we know very little about what the nature of these deficits will be, and Project Prakash is a step towards acquiring this information. Determining which skills the children are impaired at is crucial for creating effective rehabilitation schemes that would allow the children to be integrated into mainstream society and lead a normal active life. It is important to emphasize that although the patient population for this study will be drawn from India, the results will be relevant to child health in general. Furthermore, the spotlight this project is bringing to bear upon the problem of treatable childhood blindness is likely to strengthen outreach programs not just in India but globally.

Within this broad context of Project Prakash's motivations and goals, we have conducted several specific studies of object perception. Here we report an investigation of face-classification skills following recovery from blindness.

2. A study of face-classification following long-term visual deprivation

Past work has suggested that early visual deprivation profoundly impairs object and face recognition [Gregory, 1963; Valvo, 1971; Sacks, 1995; Fine et al., 2003]. Even relatively short periods of deprivation, ranging from the first 2 to 6 months of life, have been shown to have significant detrimental consequences on face recognition abilities [Le Grand et al., 2001]. However, we currently lack experimental data that address the more basic issue of the influence of early visual deprivation on face versus non-face discrimination (hereinafter also referred to as 'face classification'), i.e., can face discrimination skills be learned later in life? Results from infant studies of face perception are not too helpful in formulating a hypothesis in this context. While it is generally accepted that visual experience during the first 2 to 3 months of life is sufficient for the babies to exhibit a reliable preference for face-like patterns [Goren et al., 1975; Maurer and Barrera, 1981; Nelson and Ludemann, 1989; Johnson and Morton, 1991; Pascalis et al., 1995], it is not known whether similar learning processes continue to be available later in life. It is possible that long-term visual deprivation might permanently impair an individual's face-classification skills.

In order to investigate face/non-face classification skills following extended visual deprivation, we studied two children, SB and KK, who had both recovered sight after several years of congenital blindness. SB is a 10 year old boy who was born with dense bilateral cataracts. Prior to treatment, he showed no awareness of people's presence via visual cues and could orient to them only based on auditory cues. The cataracts severely compromised his pre-operative pattern vision. He was unable to discern fingers held up against a bright background beyond a distance of 6 inches. By comparison, subjects with normal acuity can perform this task at 60 feet and even an individual with 20/400 acuity, who would be classified as legally blind according to WHO guidelines, would be able to do this task at approximately 36 inches. It is an indicator of the poor state of awareness in rural India regarding childhood blindness, that when SB was brought in to a hospital, it was not to treat his eyes, but rather a leg injury he had suffered after tripping on an obstacle. After having been blind for 10 years, SB underwent cataract surgery in both eyes (the two procedures were a month apart). The opacified lenses were replaced with synthetic intra-ocular lenses (IOLs). Post-operative acuity in SB's eyes was determined to be 20/120, significantly below normal, but a great improvement over his original condition. SB's left eye currently exhibits significant strabismus.

KK is an 11 year old girl, also born with dense bilateral cataracts. Visual deprivation appears to have been severe right from birth since the white reflex in her eyes was pronounced even while she was an infant and KK did not exhibit any visually-guided responses. Furthermore, the nystagmus that KK currently exhibits also suggests severe visual deprivation during infancy. In tracing KK's family history, we found that her father had also been born with congenital cataracts. Thus, KK's blindness at birth was considered 'destined' (a blind father being expected to have a blind daughter) and no effort was made by her family to seek medical attention. It was only when KK was 7 years old that she happened to be examined by an ophthalmologist visiting her village as part of an outreach program. She was treated shortly thereafter and the opaque lens in her right eye was replaced with an IOL. Current visual acuity in this eye is approximately 20/120. Her left eye is still untreated and provides no useful vision.

With their guardians' permission, we conducted simple experiments to study SB and KK's face/non-face classification performance. The experiments were conducted six weeks post (first) surgery for SB and 4 years post-surgery for KK. Figure 1 shows SB and KK's eyes at the time of the study. SB's strabismus and KK's dense cataract in the left eye are evident in the images.

Figure 1. Views of SB's (top) and KK's eyes at the time our studies were conducted. Both have recovered functional vision in their right eyes. However, SB has significant strabismus in his left eye while KK continues to have a dense cataract in her left eye.

The first set of studies involved discriminating between face and non-face patterns and also locating faces in complex scenes. We also assessed the performance of two age and gender-matched controls with normal vision. Our stimulus set for the 'face/non-face discrimination' task comprised monochrome face images of both genders under different lighting conditions and non-face patterns. The non-face distracters included patterns selected from natural images that had similar power-spectra as the face patterns and also false-alarms from a well-known computational face-detection program developed at the Carnegie Mellon University by Rowley et al (1995). Sample face and non-face images used in our experiments are shown in figures 2a and b, respectively. All of the face images were frontal and showed the face from the middle of the forehead to just below the mouth. Face and non-face patterns were randomly interleaved and, in a 'yes-no' paradigm, the observer was asked to classify them as such. Presentations were self-timed and the images stayed up until the subject had responded verbally. No feedback was provided during the experimental session. The patterns subtended 10 degrees of visual angle, horizontally and vertically.

For the 'face-localization' task, we used natural scenes, containing one, two or three people (a few sample stimuli are shown in figure 2c). Face sizes ranged from 2 to 4 degrees of visual angle. The subjects' task was to indicate the locations of all faces in a scene by touching the display screen with the index-finger. The response was recorded as a 'hit' if the first touch was within a face boundary. Incorrect locations were recorded as 'false-alarms'. Both the number and correctness of responses to each scene were recorded.

Figure 2. The kinds of stimuli we used in our experiments (rows are labeled a-f top to bottom). (a) Images of upright faces (b) Non-face distracters (c) Scenes with front-facing people (d) Blurred upright faces (e) Inverted faces, and (f) Isolated face parts.

As the top row of figure 3 shows, SB and KK exhibited a high hit-rate and a low false-alarm rate on the face/non-face discrimination task, achieving performance similar to that of the age-matched controls. On the face localization task as well, the two groups were comparable. These data suggest that the ability to discriminate between faces and non-faces and also to localize faces in complex scenes can develop despite prolonged visual deprivation. Furthermore, the fact that SB exhibited this performance within six weeks of treatment suggests that face classification abilities develop rapidly after visual onset.

These results bring up the important issue of the nature of information used by SB and KK for accomplishing face-classification tasks. Past work [Le Grand et al, 2001] suggests that individuals with a history of deprivation are impaired at processing faces holistically and instead analyze them in terms of isolated features such as the eyes, nose and mouth. We attempted to determine whether SB and KK's face classification abilities were based on the use of such a piecemeal strategy wherein the presence of a face was indicated by the presence of specific parts. To this end, we performed an additional set of experiments that specifically investigated the use of holistic versus featural information. These experiments used images that were transformed to differentially effect featural versus configural analysis.

We created three stimulus sets. The first comprised low-pass filtered face and non-face patterns. The low-resolution of these images obliterated featural details while preserving the overall facial configuration. The second comprised vertically inverted faces. Vertical inversion is believed to compromise configural processing while leaving featural analysis largely unaffected [Diamond and Carey, 1986]. The third set comprised images of individual features (eye, nose and mouth). These feature images were enlarged so that low-level acuity issues would not confound the recognition results. Sample stimuli from each of these sets are shown in figure 2d-f. The first two sets were used in a face/non-face discrimination task while for the third, the subjects' task was to indicate what the image depicted. A feature-based strategy would predict that performance would be poor with the first set (low-resolution images devoid of featural details), and comparable to controls for the second and third sets.

The results are summarized in the lower row of figure 3. We found that SB and KK performed as well as the age-matched controls on the low-resolution face classification task. However, their performance was significantly poorer with inverted faces and isolated features. Notice that the controls do not exhibit impaired performance with inverted faces. This lack of an 'inversion effect' is not surprising, since the task here is not identification, but simply face/non-face classification. This pattern of results strongly suggests the use of overall configural information by SB and KK. Details of individual face parts appear to be neither necessary nor sufficient for classifying a pattern as a face.

3. Discussion

Taken together, our experimental results suggest that children can rapidly develop face classification abilities even after prolonged visual deprivation, lasting from birth for several years. Furthermore, the face concept used for classification appears to encode configural information rather than piecemeal featural details. This particular encoding strategy may well be a consequence of the relatively poor acuity the children possess following treatment for prolonged blindness. Acuity limitations reduce access to fine featural details and may, thereby, induce the use of holistic face information available in low-resolution images.

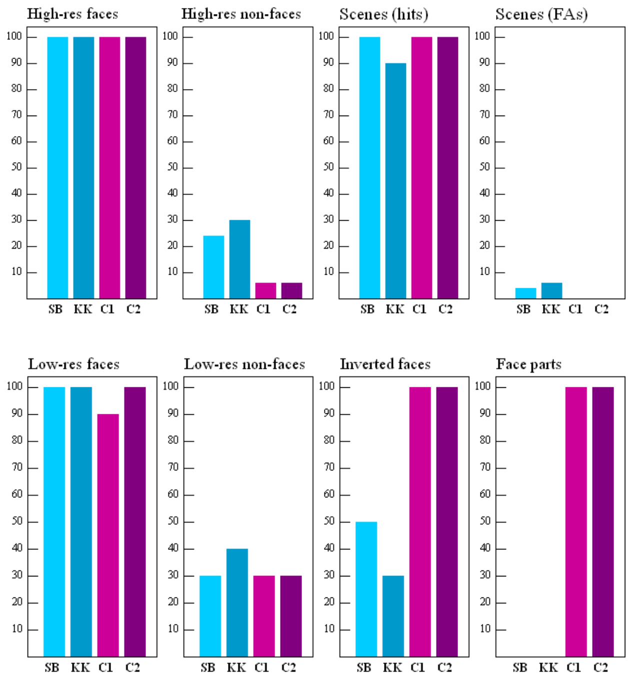


Figure 3. Results from SB, KK and two age-matched controls on various face-perception tasks.

This finding is also interesting in that it may guide the development of computational models of human face detection skills. Most current models implicitly assume that faces are encoded in terms of their parts (Lee and Seung, 1999; Ullman et al., 2002; Heisele et al. 2003). Face concept learning in these models proceeds by first acquiring facial parts which are then, optionally, combined into a larger representation. This emphasis on the use of face-parts as pre-requisites for face classification, is not reflected in our experimental results. A model that proceeded by developing a holistic face representation without need for featural details, which may be added later as higher acuity information becomes available, would be more congruent with these experimental data.

One way of reconciling our results with past reports of piece-meal processing is by assuming that visual deprivation does not compromise the encoding of overall facial configuration per se, but rather, the ability to discern differences between variants of the basic configuration. This has the consequence of increasing reliance on featural differences for distinguishing one face from another, a characteristic of piecemeal processing.

In considering whether these results have any bearing on the development of face perception skills in normal infants, it needs to be remembered that children like SB and KK differ in many ways from neonates. Unlike the newborn, SB and KK have had extensive experience of the environment through sensory modalities other than vision. This experience has likely led to the creation of internal representations that may well interact with the acquisition of visual face concepts. Furthermore, the deprivation may have led to structural changes in neural organization. For instance, projections from other senses may have claimed sections of the cortex that, in normal brains, would be devoted to visual processing. Thus, a priori, we cannot assume that the developmental courses of face perception in a 10 year old recovering from blindness will have much similarity to that in the newborn. However, some interesting parallels deserve further scrutiny. Primary among these is the quality of initial visual input. Both these populations typically commence their visual experience with poor acuity. The compromised images that result may constrain the possible concept learning and encoding strategies in similar ways. Thus, there exists the possibility that normal infants, and children treated for blindness at an advanced age, may develop similar schemes as a consequence of the similarity in their visual experience. However, the validity of this conjecture needs to be tested via further experimentation.

References

[1] Bauer, J. A., and Held, R. (1975). Comparison of visually-guided reaching in normal and deprived infant monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, **1**(4), 298-308.

[2] Diamond, R. and Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General,* **115** 107 – 117.

[3] Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., and MacLeod, D. I. A. (2003). Long term deprivation affects visual perception and cortex. *Nature Neuroscience*, Vol. 6, 915-916.

[4] Goren, C. C., Sarty, M. and Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, **56**(4), 544-549.

[5] Gregory, R. L. and Wallace, J. G. (1963). Recovery from early blindness: A case study. *Quarterly Journal of Psychology*. Monograph no. 2.

[6] Heisele, B., Ho, P., Wu, J. and Poggio, T. (2003). Face recognition: Component based versus global approaches, *Computer Vision and Image Understanding*, Vol. 91, No. 1/2, 6-21.

[7] Hubel DH, Wiesel TN. LeVay S. (1977). Plasticity of ocular dominance columns in monkey striate cortex. *Phil Trans Soc London B.* **278**:377-409.

[8] Jacobson, S. G., Mohindra, I. and Held, R. (1981). Development of visual acuity in infants with congenital cataracts. *British journal of Ophthalmology*, **65**, 727-735.

[9] Johnson, M. H. and Morton, J. (1991). Biology and Gognitive Development: The Case of Face Recognition. Oxford, UK.

[10] Le Grand, R., Mondloch, C. J., Maurer, D. and Brent, H. P. (2001). Early visual experience and face processing. *Nature*, 410, 890. See also, Correction, *Nature*, 412, 786.

[11] Lee, D. and Seung. H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788-791.

[12] LeVay, S., Wiesel, T.N. and Hubel, D.H. (1980). The development of ocular dominance columns in normal and visually deprived monkeys. *J. Comp. Neurol.* **191**:1-53.

[13] Maurer, D. and Barrera, M. (1981). Infants' perception of natural and distorted arrangements of a schematic face. *Child Development*, **52**(1), 196-202.

[14] Nelson, CA, & Ludemann, P. (1989). Past, current, and future trends in infant face perception research. *Canadian Journal of Psychology*, **43**, 183-198.

[15] Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., and Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development*, **18**, 79-85.

[16] Rowley, H. A., Baluja, S. and Kanade, T. (1998). Neural-network based face-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **20**, 23-38.

[17] Sacks, O. (1995). To see and not see. In *An Anthropologist on Mars*, Vintage Books, New York, 108-152.

[18] Ullman, S., Vidal-Naquet, M., and Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience*, **5**, 682-687.

[19] Valvo, A. (1971). *Sight restoration after long-term blindness: The problems and behavior patterns of visual rehabilitation.* New York: American Foundation for the Blind.

[20] Wiesel TN, and Hubel DH. (1963). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J Neurophysiol.* **26**:1003-1017.

Seeing Blobs as Faces or Letters: Modeling Effects on Discrimination

Lingyun Zhang & Garrison W. Cottrell
University of California, San Diego
Computer Science and Engineering
9500 Gilman Dr., La Jolla, CA 92093-0114 USA
{lingyun,gary}@cs.ucsd.edu

Abstract

What is the difference between processing faces and other objects such as letters? What makes humans face experts, and what makes this expertise different from other identification skills? It is well known that people are very sensitive to the configural information in faces. How does the sensitivity to face configuration compare to sensitivity to configurations of other stimuli? To investigate these issues, Nishimura et al. (2004) designed a test to contrast two types of processing using the same stimuli. They primed subjects to see four blobs as either a "Y" or as a face. Then they asked the subjects to discriminate pairs of these stimuli that differed only in small shifts in the blob locations. Although the stimuli were exactly the same, subjects were more accurate in the face condition than the "Y" condition. With Nishimura et al., we assumed that the subjects were relying on their letter recognition networks in the "Y" condition and their face recognition networks in the face condition to perform the task. We therefore trained two networks, a face recognition network and a letter recognition network that were otherwise identical in structure, and show here that the internal representations in the letter network for the blobs were less differentiated than the internal representations for the blobs in the face network. We argue that this is a natural consequence of the requirements of the two tasks.

1. Introduction

We have developed a simple neurocomputational model of face and object recognition that accounts for a number of important phenomena in facial expression processing, holistic processing and visual expertise [9, 7, 11, 19]. Here, we investigate the model's ability to account for a recent experiment that shows differential human sensitivity to configural information based on priming. Nishimura et al. (2004) constructed "blob" stimuli consisting of four gaussian blobs in the same spatial arrangement as the eyes, nose and mouth of human faces. The blob stimuli were also constructed to have about the same variability in location as those features in human faces. They then primed one group of subjects to see these blobs as a "Y" and another group to see them as parts of a face. The first group was less able to discriminate the blob stimuli than the second group. They suggest that this is because the face group is using their face recognition system to discriminate the blobs, and present this as further evidence that face processing utilizes a sensitivity to configuration that other tasks do not.

Why would subjects show these differential sensitivities? We first discuss what is known about face processing. Face processing has long been described as *holistic* or *configural*. Holistic is typically taken to mean that subjects use some kind of whole-face representation when processing faces. This is reflected at least two ways. First, subjects have difficulty recognizing parts of the face in isolation – there is a whole-face superiority effect. Second, subjects have difficulty ignoring parts of a face when making a decision about another part. For example, subjects are slower in making an expression judgement about the top half of a face if the bottom half is displaying an incongruent expression [3]. Our model of face processing is able to account for this kind of data because it uses representations that are global; that is, they are composed of whole-face templates we have called *holons* [7, 8]. Inputs that match part of one of these representational units cause it to fire. Units later in the processing stream take this to be a vote for the whole template, so that the system as a whole responds as if both halves of the face had been of the type matched.

Configural processing means that subjects are sensitive to the relationships between the parts, e.g., the distances between the eyes. Thus, small changes in the spacing of the eyes cause subjects to see the faces as different people. This is presumably due to long experience with many people, and the need to differentiate these faces. This sensitivity to configuration, however, takes a surprisingly long time to develop [23].

What kind of processing is required to recognize objects, and how does it differ from faces? Diamond and Carey [14] were among the first to discriminate between the types of processing involved in face and object recognition. They proposed that first-order relational information, which consists of the coarse spatial relationships between the parts of an object (e.g. eyes are above the nose), is sufficient to recognize most objects at the basic level. By contrast, second-order relational information (e.g. the spacing between individual features such as the eyes and the mouth), is needed for face recognition. They found that inverting images severely disrupted subjects' ability to discriminate between faces, and that this effect was stronger for faces than for dogs and landscapes in naive subjects. However, dog experts also showed an inversion effect for dogs. These results suggest that face processing is a kind of expertise, and that experts become over-tuned to the typical orientation of the stimuli in their domain of expertise. Their ability to discriminate based upon subtle configural differences is overly disrupted by inversion. Diamond and Carey [14] suggest that experience allows people to develop a fine-tuned prototype and to become sensitive to second-order differences between that prototype and new members of that category (e.g. new faces).

One implication of the Diamond and Carey study is that the inversion effect (a large reduction in same/different performance on inverted faces, compared to inverted objects) is based on a relative greater reliance on second-order relational information, and that perhaps this characteristic distinguishes face/expert-level processing from regular object recognition. Farah et al. [15] found that encouraging part-based processing eliminated the inversion effect, whereas allowing/encouraging non-part-based processing resulted in a robust inversion effect. Thus Farah et al. conclude that the inversion effect, in faces and other types of stimuli, is associated with holistic pattern perception. Thus, regular object classification is thought to use a parts-based representation.

However, our model uses the same kind of representation for all stimuli. The only difference is the requirements of the task, between a version of our model that recognizes faces and one that recognizes objects, such as letters. In face identification, the model must take similar looking stimuli, and magnify small differences between them in its internal representation (see Figure 1, left). On the other hand, in order to recognize letters, the model must take similar looking things (the same letter in different fonts, for example) and represent them as the same thing (see Figure 1, right). This point has been made before [17]; here, we construct models that automatically implement those differences via learning the different tasks. Then we may analyze the models in ways that we cannot analyze human subjects.

Our use of separate networks for these two tasks is motivated by fMRI experiments that have shown that face-related tasks and letter-related tasks activate different brain regions [17, 20, 5]. The fusiform face area tends to be in right medial fusiform gyrus, whereas there appears to be a letter form or word form area in the left midfusiform gyrus [5, 6, 22]. We model this by having two networks, each trained to do one of the tasks. We hypothesize that the priming in Nishimura et al.'s experiment causes one of these networks to be primed, and therefore used for the task. Then we show how blobs are represented differentially in the two networks.

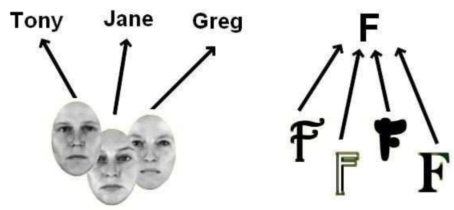


Figure 1. Faces are automatically perceived as different individuals despite the similarity, while the "F"s are perceived as the same letter. Adapted from [17].

In the following, we describe Nishimura et al.'s experiments and our account of their data. We found that different encoding due to different tasks in our model could account for their data, which suggests that the effect in human subjects may come from using the letter system verses the face system to encode the stimuli. Finally, we discuss plans for future work.

1.1. Nishimura et al's Stimuli and Experiments

To compare sensitivity to configural information for faces and non-faces, Nishimura et al. (2004) created a group of stimuli that contain 4 blobs located over the eyes, nose and mouth of configurally different faces (Figure 2). Two groups of subjects were then primed to see the blobs as either a face or as the letter "Y", respectively (Figure 3). The subjects then performed same/different tasks on different pairs of blobs that differed only in their relative locations. The results showed that when people take these blobs as faces, they discriminate them better than when they see them as the letter Y (Figure 4). This suggests that by different priming, ambiguous stimuli might be represented differently. In this work, we concentrate on modelling this using our neurocomputational model of visual object recognition.

2. A Computational Model of Classification

Our model is a three level neural network that has been used in previous work (Figure 5). The model takes manually aligned images as input. The images are first filtered by

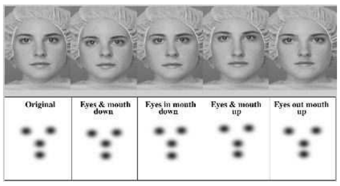


Figure 2. 4 blobs are located over the eyes, nose and mouth of 5 faces used in previous studies (from [24]).

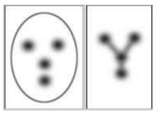


Figure 3. The blobs are primed as either face or letter Y. (from [24]).

2D Gabor wavelet filters, which are a good model of simple cell receptive fields in cat striate cortex [18]. PCA (principal component analysis) is then used to extract a set of features from the high dimensional data. In the last stage, a simple back propagation network is used to assign a class to each image. We now describe each of the components of the model in more detail.

2.1. Perceptual Level

Research suggests that the receptive fields of the striate neurons are restricted to small regions of space, responding to narrow ranges of stimulus orientation and spatial frequency[18]. DeValois et al [13] mapped the receptive fields of V1 cells and found evidence for multiple lobes of excitation and inhibition. 2D Gabor filters [12](Figure 6) have been found to fit the 2D spatial response profile of simple cells quite well[18]. In this processing step the image was filtered with a rigid 23 by 15 grid of overlapping 2-D Gabor filters[12] in quadrature pairs at five scales and eight orientations [11](Figure 7). We thus obtained $23 \times 15 \times 5 \times 8 = 13,800$ filter responses in this level, which is termed the *perceptual* level [11].

2.2. Gestalt Level

In this stage we perform a PCA of the Gabor filter responses. This is a biologically plausible means of dimensionality reduction[11], since it can be learned in a Hebbian manner. PCA learns features that encode correlations between features at the previous level. Thus, for example, if the Gabor filter responses to the left eye are highly correlated with the Gabor filter responses to the right eye, there will be a principal component that corresponds to both of these, capturing the redundancy in the Gabor filter responses. The eigenvectors of the covariance matrix of the patterns are computed, and the patterns are then projected onto the eigenvectors associated with the largest eigenvalues. At this stage, we produce a 50-element PCA representation from the 13,800 Gabor vectors.

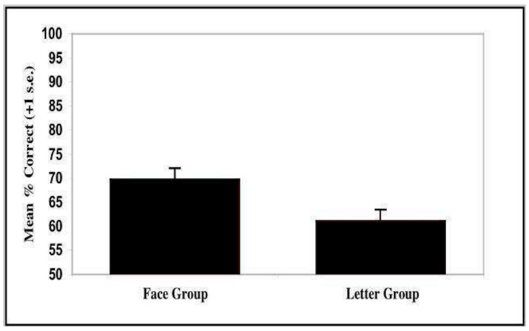


Figure 4. Mean accuracy of the subjects in the face group was higher than that of those in the letter group. (from [24]).

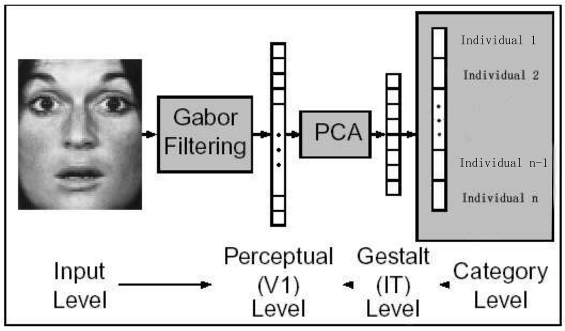


Figure 5. Object recognition model (from [11]).

2.3. Categorization Level

The classification portion of the model is a two-layer back-propagation neural network. 50 hidden units are used. A scaled tanh [21] activation function is used at the hidden layer and the softmax activation function $y_i = e^{a_i} / \sum_k e^{a_k}$ was used at the output level. The network is trained with the cross entropy error function [1] to identify the images using localist outputs. Networks trained in this way learn to produce the conditional probability of the output class given the input.

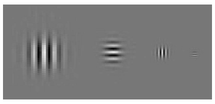


Figure 6. A Gabor function is constructed by multiplying a Gaussian function by sinusoidal function[12]. We use five scales and eight orientations.

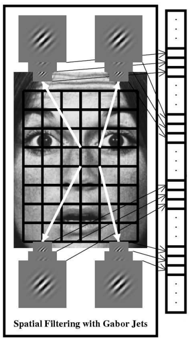


Figure 7. An image filtered with a rigid 23 by 15 grid of overlapping 2-D Gabor filters in quadrature pairs at five scales and eight orientations (from [10]).

3. Modelling Nishimura et al.

We set up two networks, one engaged in face identity classification and the other in letter classification. After training, blobs were fed to these two networks. The mean discriminabilities were computed respectively and then compared between the two classifiers. The results showed that the face network considers the blob stimuli to be more different than the letter network does.

3.1. The Image Sets

We had 182 face images of 26 individuals (7 images for each individual). We also used 182 letter images of the 26 upper case letters (7 images for each letter in 7 fonts). 27 blob stimuli were created by manipulating the eye blob and/or mouth blob's position.

The FERET database is a large database of facial images, which is now standard for face recognition from still images[25]. We used 182 face images of 26 individuals, 7 images each. In [11], where the task was to learn facial expressions, images were aligned so that eyes and mouth went to designated coordinates. This alignment removes the configural information which is crucial for our work, because we are trying to understand how configural processing is better learned in the face recognition task than in the letter recognition task. To avoid this negative effect, we required that the relative spacing between the parts of the face remain the same. We formed a triangle from the eyes and mouth of the original face, and then translated, rotated and scaled this triangle to be as close as possible to a target triangle in terms of the sum squared differences between the final eye and mouth coordinates and the target coordinates.[1] This manipulation preserves the relative distance between the features of the face (Figure 8). Thus, a triangle represented by the eyes and mouth is scaled and moved to fit closely to a reference location, but the relative sizes of the sides of the triangle are not changed. The aligned images were 192 pixels by 128 pixels.

Figure 8. Two examples of face image normalization. The faces were cropped with the eyes and the mouth as close as possible to the target position while keeping the shape of the triangle among these features the same.

We also used 182 192 by 128 pixel letter images of the 26 upper case letters, 7 images each (Figure 9). The letter images were aligned so that the ends of the letter Y were approximately where the eyes and mouth were in the face stimuli.

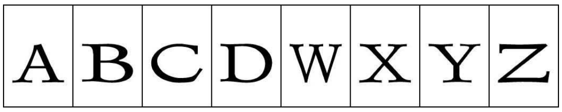


Figure 9. Some letter images.

Blob images were generated by setting gaussian blobs (of width $\sigma = 5$ pixels) at left eye $(80(\pm 3), 36(\pm 3))$, right eye $(80(\pm 3), 92(\pm 3))$, nose $(115, 65)$ and mouth $(150(\pm 3), 65)$ positions. Note the two eye blobs were always symmetric. Thus $3 * 3 * 3 = 27$ blob images were generated.

3.2. Training and Learning

A learning rate of 0.05 and a momentum of 0.5 were used in the results reported here. Two networks were set

[1] The objective function for the minimization was $(||Eye_{right} - targetEye_{right}||^2 + ||Eye_{left} - targetEye_{left}||^2 + ||Mouth - targetMouth||^2)$.

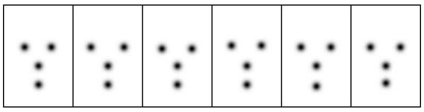


Figure 10. Some "blob" images.

up for faces and letters respectively. In a pilot experiment, 10 percent of the images were selected randomly as a test set and another 10 percent as a validation set [10]. Both networks achieve 80-90 percent accuracy within 50 epochs. This classification rate was good enough to show that our model represented the images well.

For the following experiments, we simply trained both networks on all 182 images, since we are only interested in obtaining a good representation at the hidden layer. Training was stopped at the 50th epoch based on the above pilot experiment. After training, the blob images were presented to the network. Note for the face network, the blob images were projected according to the face image eigenvectors in the PCA level while for the letter network, they were projected onto those of letters.

3.3. Modelling Discrimination

Hidden unit activations were recorded as the network's representation of images. In order to model discriminability between two images, we present an image to the network, and record the hidden unit response vector. We do the same with a second image. We model similarity as the correlation between the two representations, and discriminability as one minus similarity [11]. The pairwise average within the blob image set was taken as the measure of the network's ability to discriminate the blob images. For both the face network and the letter network, the average of the discriminabilities was computed over 50 networks which were all trained in the same way, but used different initial random weights.

The results (Figure 11) showed that the face network better discriminates the difference between blob images than the letter network ($F = 24.72$, $p < 0.001$). I.e. the representations of blob images in the face network were more differentiated (further apart) from one other than those in the letter network.

To visualize this difference, we extracted the principal components of the hidden layer representations and then projected their activations onto the first three principal components (the ones that represent most of the variance of the activations) (Figure 12). Notice the hidden layer representations of the blobs in the face network are better separated, which suggests that the face network is especially sensitive to configural differences.

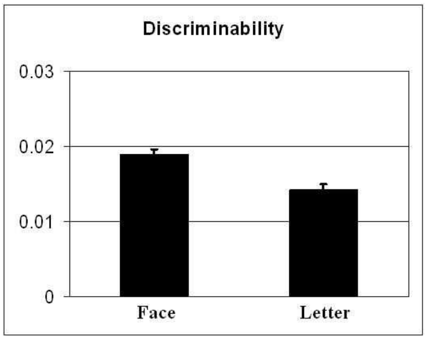


Figure 11. Mean discriminability of the blobs in the face network was higher than that in the letter network ($F = 24.72$,$p < 0.001$).

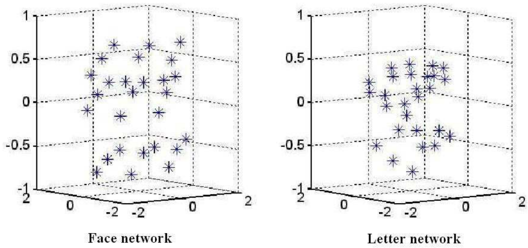


Figure 12. The projection of the hidden activities onto the first 3 principal components.

3.4. Inverted blobs

While Nishimura et al. did not present inverted blobs to their subjects, it is theoretically interesting to see what subjects would do with inverted stimuli if they are perceiving them as letters or as faces. If they see them as faces, then one expects that there should be an "inverted blob effect." This is easy to do with the networks. The same set of the blobs were inverted and then presented to the two types of networks. We expect that the face network should show a greater loss of discrimination than the letter network. The results, plotted along with the original data from Figure 11 for comparison, are shown in Figure 13. As expected, there is an inverted blob effect for the face network, while the letter network shows a slight increase in discriminability for the blobs. As there is no human data for this case, this is a prediction of the model.

4. Discussion

Our results qualitatively match those from human subjects. In the two networks, the difference in the discriminability of the same stimuli comes from the different en-

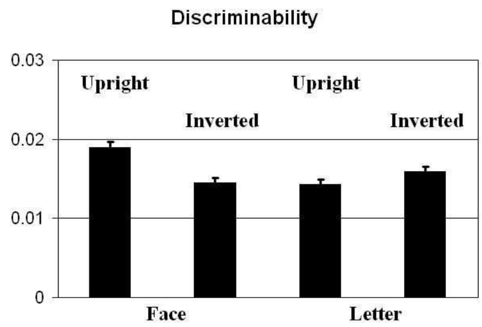


Figure 13. Discriminability results for the letter and face networks when the blobs are presented upright and inverted. There is no difference between the networks on the inverted stimuli ($F = 3.11$,$p = 0.0808$), but there is a significant difference between the face network's ability to discriminate the inverted blobs over the upright blobs ($F = 21.9$,$p < 0.001$). There is a small but significant increased discriminability for the letter networks ($F = 4.3$,$p = 0.0407$).

coding in the two tasks. Note that because all the faces share the same first-order relational features, the categorization need to be carried out at finer level. The face network needs to spread out similar face images to categorize individuals, while the letter network needs to squeeze different fonts of the same letter to a letter prototype. So the face network tends to magnify small differences in face images while the letter network tends to ignore such variability. This is consistent with our previous results with faces, objects and letters [19, 26]. Thus the configural differences between the blob stimuli are better noticed by the face network. Considering our separate face and letter networks to be analogous to the separate face and letter processing systems in the brain, we can apply the reasoning derived from our networks to the brain, as follows. The face processing system learns to pay attention to small differences such as configural changes in faces while the letter processing system ignores them. Thus, when the blobs are perceived as faces, the differences are more present in the inner representation by the face system, while when the blobs are perceived as letters, the differences are less present in the inner representation by the letter system.

Our model makes a prediction concerning an inverted blob effect. We have previously shown that our face networks show an inverted face effect – they are poorer at discriminating upside down faces [27]. We also showed that configural differences take the biggest hit in discrimination when compared with featural changes between the stimuli to be discriminated. The configural effect generalizes to the blobs. Interestingly, the letter network does not show this effect – its discriminability scores increase slightly when the stimuli are inverted. Possibly this is because it is poor at discriminating the upright stimuli that all match a particular letter "Y". In the inverted case, we speculate that the blobs fall in regions where it may have to discriminate slight differences between letters (e.g., "R" and "A").

How might these effects vary with development? As the participants were undergraduates [24], we would expect a lower strength of this effect in young children since sensitivity to configuration is lower in children [16, 23]. Furthermore, if the inverted blob effect is found in adults, as predicted by our model, we would expect that children would not show this effect. This is because developmentally, children show either no inversion effect, or less of an inversion effect, depending on their age [4, 2]. We are currently exploring adding a developmental component to our model in order to account for these developmental changes.

5. Acknowledgements

We thank Carrie Joyce for previous suggestions, Mayu Nishimura, Daphne Maurer and Catherine J. Mondloch for useful comments, and for inspiring this modelling idea. This research was supported by NIMH grant MH57075 to GWC.

References

[1] BISHOP, C. M. *Neural networks for pattern recognition.* Oxford University Press, 1995.

[2] BRACE, N. A., HOLE, G. J., KEMP, R., PIKE, G., DUUREN, M. V., AND NORGATE, L. Developmental changes in the effect of inversion: Using a picture book to investigate face recognition. *Perception 30* (2001), 85–94.

[3] CALDER, A. J., YOUNG, A. W., KEANE, J., AND DEAN, M. Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance 26*, 2 (2000), 526–551.

[4] CAREY, S., AND DIAMOND, R. From piecemeal to configurational representation of faces. *Science 195* (1977), 213–313.

[5] COHEN, L., DEHAENE, S., NACCACHE, L., LEHERICY, S., DEHAENE-LAMBERTZ, G., HENAFF, M.-A., AND MICHEL, F. The visual word form area. Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain 123* (2000), 291–307.

[6] COHEN, L., LEHERICY, S., CHOCHON, F., LEMER, C., RIVARD, S., AND DEHAENE, S. Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain 125* (2002), 1054–1069.

[7] COTTRELL, G. W., BRANSON, K. M., AND CALDER, A. J. Do expression and identity need separate representations? In *Proceedings of the 24th Annual Conference of the Cognitive Science Society* (Mahwah, New Jersey, 2002), The Cognitive Science Society.

[8] COTTRELL, G. W., AND METCALFE, J. Empath: Face, gender and emotion recognition using holons. In *Advances in Neural Information Processing Systems 3* (San Mateo, 1991), R. P. Lippman, J. Moody, and D. S. Touretzky, Eds., Morgan Kaufmann, pp. 564–571.

[9] DAILEY, M. N., AND COTTRELL, G. W. Organization of face and object recognition in modular neural network models. *Neural Networks 12* (1999), 1053–1073.

[10] DAILEY, M. N., COTTRELL, G. W., AND ADOLPHS, R. A six-unit network is all you need to discover happiness. In *TwentySecond Annual Conference of the Cognitive Science Society* (2000).

[11] DAILEY, M. N., COTTRELL, G. W., PADGETT, C., AND ADOLPHS, R. Empath: A neural network that categorizes facial expressions. *Journal of Cognitive Neuroscience 14*, 8 (2002), 1158–1173.

[12] DAUGMAN, J. G. Uncertainty relation for resolution in space, spacial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of American A 2* (1985), 1160–1169.

[13] DEVALOIS, R. L., AND DEVALOIS, K. K. *Spatial Vision*. Oxford University Press, 1988.

[14] DIAMOND, R., AND CAREY, S. Why faces are and are not special: an effect of expertise. *Journal of Experimantal Psychology: General 115*, 2 (1986), 107–117.

[15] FARAH, M., LEVINSON, K., AND KLEIN, K. Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia 33* (1995), 661–674.

[16] FREIRE, A., AND LEE, K. Face recognition in 4- to 7-year-olds: Processing of configural, featural, and paraphernalia information. *Journal of Experimental Child Psychology 80* (2001), 347–371.

[17] GAUTHIER, I., MOYLAN, J., TARR, M. J., ANDREW, A. W., SKUDLARSKI, P., AND GORE, J. C. Automatic subordinate-level processing for faces and letters, 1998. Presentation at *Human Brain Mapping*.

[18] JONES, J. P., AND PALMER, L. A. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology 58*, 6 (1987), 1233–1258.

[19] JOYCE, C., AND COTTRELL, G. W. Solving the visual expertise mystery. In *Proceedings of the Neural Computation and Psychology Workshop 8*, Progress in Neural Processing. World Scientific, London, UK, 2004.

[20] KANWISHER, N., MCDERMOTT, J., AND CHUN, M. M. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience 17* (1997), 4302–4311.

[21] LECUN, Y., BOTTOU, L., ORR, G. B., AND MüLLER, K.-R. Efficient backprop. In *Neural Networks—Tricks of the Trade, Springer Lecture Notes in Computer Sciences* (1998), vol. 1524, pp. 5–50.

[22] MCCANDLISS, B. D., COHEN, L., AND DEHAENE, S. The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences 7*, 7 (2003), 293–299.

[23] MONDLOCH, C. J., GRAND, R. L., AND MAURER, D. Configural face processing develops more slowly than featural face processing. *Perception 31* (2002), 553–566.

[24] NISHIMURA, M., MAURER, D., AND MONDLOCH, C. J. Perceiving changes in spacing among four 'blobs': Are adults more sensitive when primed to see them as facial features?, 2004. Poster presented at *The 2004 Cognitive Neuroscience Society Annual Meeting*.

[25] PHILLIPS, J., WECHSLER, H., HUANG, J., AND RAUSS, P. J. The feret database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing 16*, 5 (1998), 295–306.

[26] TRAN, B. A., JOYCE, C. A., AND COTTRELL, G. W. Visual expertise depends on how you slice the space. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (Mahwah, New Jersey, 2004), The Cognitive Science Society.

[27] ZHANG, L., AND COTTRELL, G. W. When holistic processing is not enough: Local features save the day. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (Mahwah, New Jersey, 2004), The Cognitive Science Society.

Using a Robot to Reexamine Looking Time Experiments

Andrew Lovett
Northwestern University
1890 Maple Avenue, Evanston, IL
Andrew-Lovett@northwestern.edu

Brian Scassellati
Yale University
Social Robotics Lab
51 Prospect Street, New Haven, CT
Brian.Scassellati@yale.edu

Abstract

We present new evidence to support criticisms of infant looking time experiments. One such experiment, in which Baillargeon concluded from looking times that infants understand object permanence, is examined in detail. An alternative model of infant cognition, using the idea that infants look longer at particular scenes based on visual processing at the pre-attentive level, rather than based on any understanding of the objects they are seeing, is suggested. The model is implemented on Nico, a humanoid robot currently being developed in the Yale Social Robotics Lab. The validity of the model is established by running Nico through a reenactment of Baillargeon's initial experiment and achieving comparable results. The argument is made that while these results do not prove the suggested model is correct, they do prove that the suggested model is sufficient for explaining Baillargeon's results. We conclude that the demonstrated validity of the model prevents Baillargeon from claiming that the initial experiment provides proof of an understanding of object permanence in infants. We suggest that the model could be further validated by running Nico through other looking time experiments.

1. Introduction

In the past few years, the validity of infant cognition studies has been under debate. Because infants are unable to communicate their thoughts and, in the case of particularly young infants, unable even to control their bodies in meaningful ways, researchers desiring to determine what infants are thinking have been forced to rely exclusively on indirect cues such as looking times. In one popular looking time paradigm, the violation of expectation study, the infant watches a scene that is either physically possible or physically impossible. If the infant looks longer at the impossible scene, the researcher concludes that the infant was surprised by the scene because it is impossible, and that the infant must therefore possess some understanding of the principles which make the scene impossible. This paradigm was used famously by Karen Wynn to suggest that infants understand basic arithmetic [8]. A similar paradigm, in which infants habituate to a basic scene before being shown either an impossible or a possible variation of it, is also quite popular. It has been used in a number of studies to suggest infants understand various concepts about the laws of physics, such as object continuity [3]. The problem with these studies, critics such as Haith claim, is that they assume too much about the conceptual understanding of infants based on perceptual evidence [4]. According to Haith, infants may behave the way they do without any knowledge of such concepts as arithmetic, object continuity, or impossibility. Infants may follow much simpler rules for tracking objects and may respond to novel stimuli, stimuli that the rules did not predict, by looking longer at these stimuli, without giving any thought to what is possible or impossible.

Because of the difficulty of finding new ways to study infant cognition, the debate remains unresolved. However, the field of robotics may be key to determining whether a less generous model of infant cognition is sufficient to explain the results of infant cognition studies [1]. If robots, following basic rules without any knowledge of higher-level concepts, can replicate infant behavior, then it is possible that infant cognition follows similar basic rules. Schlesinger [6] created a very simple block animation based on Baillargeon's study [2] in which a cart moves down a track, briefly passing behind an occluder. Schlesinger evolved a neural network which received as input all the pixels in a given frame of the animation, as well as all the pixels in the foveal area, the part of the animation at which the agent was currently "looking." The neural network produced as output directives on where in the scene the agent should look next. After training the net with the basic animation, Schlesinger tested it on animations corresponding to the possible and impossible scenes used by Baillargeon.

Schlesinger was able to draw parallels between the performance of his neural network and the performance of infants in Baillargeon's study. However, because Schlesinger used a basic simulation with a simplified animation to represent what his agent was seeing, his interpretations are open to criticism. The agent did not have to deal with real world constraints such as noise in an image, nor could it benefit from real world cues such as depth. Furthermore, because the simulation involved evolving a net for a particular animation, using the same agent to mimic other infant cognition studies would be a nontrivial matter. At the very least, animations would have to be built for each study, and a separate neural network might need to be evolved for interacting with each one. Even then, it is unclear whether the same types of results would be achieved. Thus, it is hard to justify generalizing his results to the performance of infants in the real world.

In the present paper, a new model of infant cognition is proposed as an alternative explanation for the results achieved in Baillargeon's looking time experiment. This alternative model is a pre-attentive model, meaning the model presupposes that the results in looking time experiments can be explained in terms of visual processing that goes on in the infant's brain before the infant is consciously aware of any visual stimuli. If the model is correct, then infants do not look longer at a particular stimulus because they make a conscious decision to look at what they find particularly surprising. Instead, they look longer at a stimulus because of basic, automatic mechanisms.

The model is implemented on the humanoid robot Nico. Because Nico is an embodied entity existing in the physical world, the robot must deal with all the constraints and cues facing actual infants, rather than merely those factors which a programmer has thought to include in a simulation. The model is tested by running Nico through an experiment similar to Baillargeon's. The purpose of this experiment is not to determine whether the model accurately describes what is occurring within an infant's brain. The purpose is rather to demonstrate that the model could potentially result in infants behaving the way Baillargeon found them to behave.

2. Baillargeon's Experiment

In Baillargeon's object permanence experiment [2], 6-month-old and 8-month-old infants were first shown two randomly ordered familiarization trials. In both trials, a yellow cart stood immobile on a downward-sloping track. A red screen, which would otherwise have occluded the cart, was held up above the track. A green box was placed on the track next to the yellow cart in one trial and behind the track in the other trial. After seeing both familiarization trials, infants were exposed to a series of habituation trials. In each trial, they were shown an eight-second scene that repeated until they lost interest in it. They were considered to have lost interest when they looked away for at least two seconds. Because the scene was divided into two-second segments, the infants had to lose interest for at least one entire segment for the trial to end. As the scene began, the red screen stood in front of the track, occluding a small part of it. In the first two-second segment, the red screen was lifted up into the air and then lowered. In the second segment, nothing happened. In the third segment, the yellow cart appeared at the top of track, took approximately two seconds to travel down the track, briefly disappearing behind the red occluder, and then moved out of the infant's view at the bottom of the track. In the fourth segment, nothing happened again. Experimenters recorded the total time it took infants to lose interest in this scene and then repeated the habituation trials until the time in three consecutive trials was half of what it had been in the first three trials, at which point the infant was considered to have habituated to the scene. On average this required about eight habituation trials.

After habituation was achieved, each infant was given the two familiarization trials again. Then, the infant was given three more trials. For half the infants, these trials followed a possible-impossible-possible pattern, and for half the infants the trials followed an impossible-possible-impossible pattern. The possible trials were the same as the habituation trials, except that when the red screen rose into the air, the infants saw the green box sitting behind the track. The impossible trials were the same except that when the red screen rose, the infants saw the green box sitting on the track. This scene was considered impossible because the presence of the green box on the track suggested that as the cart travelled behind the occluder, the cart was actually moving through the green box. According to Baillargeon, the infants looked at the impossible scene longer because they were surprised that one object could move through another object. Since the impossible event apparently occured behind the red screen, Baillargeon concluded that infants represent the locations of objects located behind occluders, represent the velocity trajectories of objects moving behind occluders, and understand object permanence, i.e. the idea that one object cannot move through another.

3. An Alternative Explanatory Model

Our model begins with one of the most basic human perceptual abilities: feature detection. Every human has neurons in the brain that are activated for certain types of visual stimuli at certain locations in the visual field. These neurons respond at a pre-attentive level. They include neurons that respond to stimuli of particular colors, neurons that respond to stimuli at particular depths, and neurons that respond to stimuli with particular motion vectors [5]. When an infant sees a red ball moving through its visual field, the appropriate neurons for a red object at a particular depth moving with a particular velocity are activated.

Another basic cognitive ability that does not require conscious thought is the ability to create associations in the mind between two stimuli. Classical conditioning experiments show us that even rodents can do this [7]. Suppose, then, that human infants can build and remember associations between groups of stimuli. These associations can be seen as very basic mental constructs, which will be called elements. Imagine that an element is represented by an array of neurons, with each neuron in the array corresponding to a location in the visual field. In the case of the red ball, the infant might associate one such array with the redness, the motion vector, the location in the visual field, and the depth in the visual field. If the infant later saw a red ball moving in the same direction very close to where the red ball element was first formed, the associations between all these features and the red ball element would cause the neuron for the new red ball's location to be excited. If, on the other hand, the red ball vanished and a blue ball appeared in its place, the neuron would be excited to a lesser degree, as the element would share only the location, velocity, and depth features of the new stimulus.

Of course, remembering an element's location in the visual field is only of limited use. A human's eyes are constantly moving as the person focuses on different stimuli. Once an element disappears from view, perhaps by moving behind an occluder, its last known location in the visual field quickly becomes unreliable. Therefore, it may be more useful to remember an element's location relative to the locations of other elements, particularly if they are near each other. One can imagine that there are associations between elements in the infant's memory. If one element was last seen moving behind a second, then there will be a strong association between the elements, and simply looking at the second will excite the neuron for the first at the same location.

The model can be examined in greater detail, although details beyond the basic framework are little more than speculation. When an object appears in the visual field, if it shares any features with any elements in the infant's memory, those elements will be excited for the object's location in the visual space. If an element is sufficiently excited for a particular location, then it will become activated for that location. This has several effects. First, an element's activation at one location in space inhibits its activation at other locations, meaning an element cannot be activated for more than one location at the same time. Second, an element's activation inhibits the creation of a new element at the same location, so the infant will not make the mistake of assuming an object is both an old element and a new element. Thirdly, the element's activation causes the infant to habituate to the element, meaning the infant will find the element slightly less exciting than it has in the past. Finally, the element's activation stimulates the infant to look at the object's location in visual space. The element's activation does not inhibit the activation of another pre-existing element at the same location in space, meaning that two elements may be activated for a single object. This quirk in the infant's cognitive ability will prove important in explaining infant behavior in Baillargeon's study.

When an object disappears from the visual field, the activation of any elements associated with it will persevere for a short time and then cease. At this point, the element's last known location in the visual field will decay very quickly. Unless the object reappears in a short time, the element will lose all association with any location in visual space. The other associations, however, will remain for a longer time. As the infant gradually forgets the element, the element's level of habituation will decrease. However, the more time the infant has spent looking at the element, the harder it is for the infant to forget the element.

4. Interpreting Baillargeon's Results

Now, suppose we apply the model to Baillargeon's experiment. Instead of imagining that infants understand and are able to habituate to an entire scene, the model suggests that infants are merely habituating to a set of elements. In the habituation trials, the red screen is always in the infant's field of view, while the cart is only in the field of view for about two out of every eight seconds. Thus, one might assume that infants would habituate to the red screen much more quickly than they would habituate to the yellow cart. However, it seems plausible to suggest that when the stimuli associated with an object change significantly, such as when the red screen moves up and down, the degree of habituation decreases. Since the red screen remains stationary most of the time, it may become much more interesting to the infant during the two seconds when it is moving. Since the yellow cart is always moving, its motion would be less exciting. Suppose that the infant will never look away during the two seconds when the red screen is moving. Once the infant has habituated sufficiently to the red screen, the infant will look away during the two seconds before and after the red screen moves, as nothing happens during these intervals. However, the infant must look away for a total of at least two seconds for the trial to end. This means that unless the infant looks away immediately at the beginning of one of the intervals where nothing happens, the infant will need to look for some time beyond the length of these intervals for the trial to end. The only remaining interval is the two seconds when the yellow cart is visible. Thus, the infant must completely habituate to the yellow cart before the infant will look away for a total of two or more seconds.

During the intervals between trials, none of the objects are visible to the infant. Presumably, during these intervals the infant partially forgets the elements associated with the red and yellow objects, losing some of the habituation to these elements. After each trial, the infant has seen the objects for more time and thus has more trouble forgetting their elements, so after each trial the infant dehabituates to a lesser degree, causing each successive trial to take less time.

In the possible test trial, the green box is introduced. The green box was visible during the familiarization trials but not during any of the habituation trials, and thus it is much more interesting to the infant. However, the box is only visible while the red screen is moving up and down, and this is not the period when the infant looks away. Thus, the green box has little effect on the total time of the trial.

The impossible event is similar to the possible event, but with one important distinction: the green box is placed on the cart's track. The infant may not realize this, or even know what a track is. However, this distinction also means that the green box's depth, i.e. its distance from the eyes of the infant, is the same as the depth of the cart. When the red screen moves up and down, the infant creates a green box element and associates this element with its depth, with its color, and with the red screen element, since it was last seen being occluded by the red screen. During the first cycle of the impossible trial, as the yellow cart appears, it catches the infant's attention, and the infant's eyes move to follow it. As the cart moves behind the red screen, it is briefly located directly next to the screen. The infant's green box element is excited for the position of the red screen because of the association between the red screen and the green box. The green box element is also excited because the yellow cart is located at the same depth as the green box. Some of the time, this may be sufficient to activate the green box for this location. As the yellow cart comes out from behind the red screen on the other side, it again may excite the infant's green box element for that location. Since the cart moves quickly on, the activation is very brief. However, because elements that have disappeared from view persevere for a short time, the infant continues attending to the imagined green box for a little while longer, while the yellow cart moves on. Because the infant has had less time to habituate to the green box in previous trials, the infant finds the green box more interesting than the yellow cart. Thus, the infant pays less attention to the yellow cart and so habituates to the cart at a slower rate. It takes the infant more time to habituate to the yellow cart, and so the impossible trial takes more time to complete.

5. Methodology

The cognitive model was implemented by building a *memory* module for the robot Nico. Nico is a humanoid robot currently being developed at the Yale Social Robotics Lab. Nico is designed to both look and behave like a nine-month-old infant. Nico's head contains six motors, three of which control Nico's "eyes." One pan motor rotates each eye from side to side, while a third tilt motor rotates both eyes up and down. The eyes each consist of two small cameras. One camera possesses a wide field of view, and the other camera, which represents the fovea, possesses a narrower field of view. Nico also possesses a torso and an arm, although only software for controlling the head was used in the present experiment.

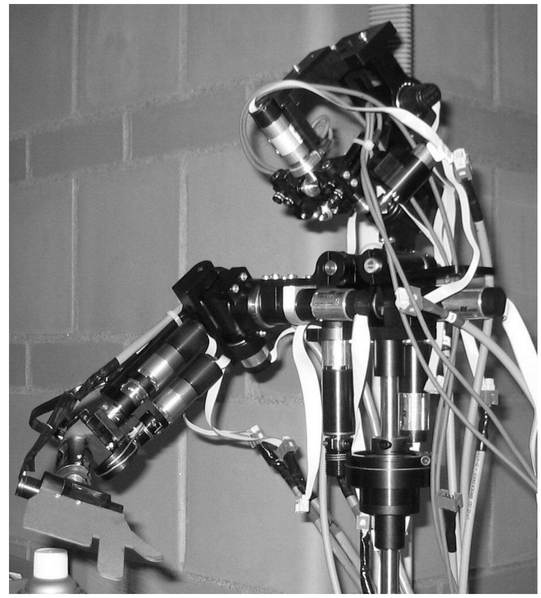


Figure 1. Nico, an upper-torso humanoid robot used in the implementation of our alternative model

5.1. Nico's Software

Nico is controlled by a set of software modules running in parallel on 16 networked computers. Some of these modules have been ported from code written for Cog, a humanoid robot at MIT, while other modules have been developed by members of the lab. The modules pass information to each other through the "port system." Each module performs a basic cognitive operation similar to a function that might be performed by a particular area in the human brain. The modules can be divided into perceptual processing modules and behavior control modules.

The basic purpose of the perceptual processing modules is to extract information from the video cameras that make up Nico's eyes. The lowest-level modules detect single-pixel features in the image. For example, the color module detects bright colors in an image. The skin module detects colors that are likely to be skin tones. The motion module detects motion in the form of changes in a single

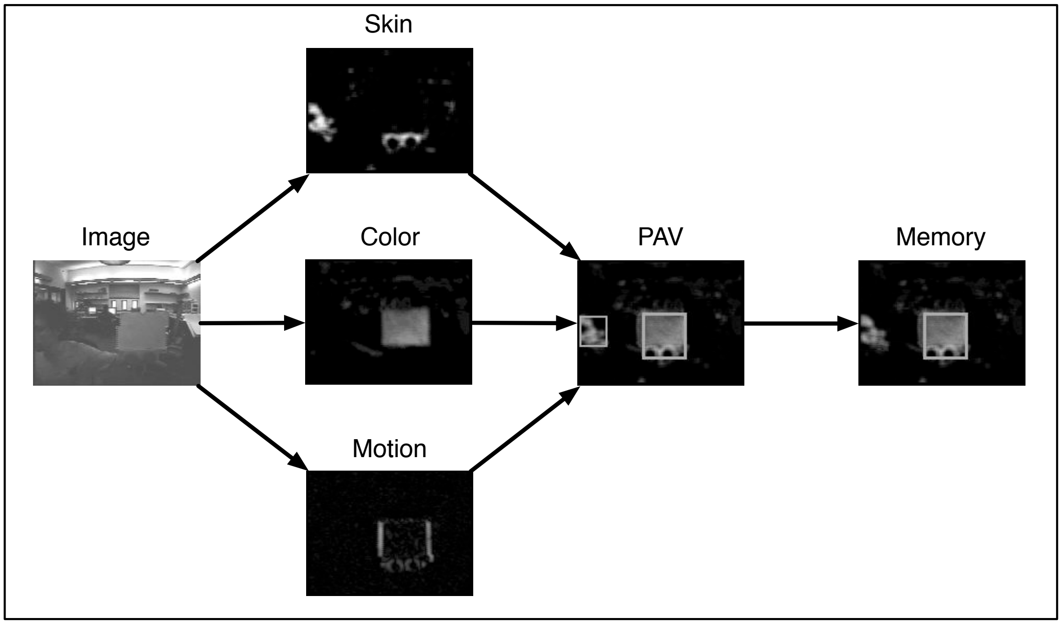


Figure 2. The Visual Processing Pipeline: the image is passed to the Skin, Color, and Motion modules, each of which produces a saliency map; the PAV module performs a weighted summation of their output and boxes interesting objects; the *memory* module performs habituation and chooses the most salient object (the images shown above represent typical output from each module)

pixel's intensity over time. Each of these modules produces a saliency map, a grayscale image in which the value for each pixel represents the degree to which the module found that pixel to be salient.

The pre-attentive vision module, or PAV, receives any combination of saliency maps and performs a weighted summation of their values, using weights determined either by the user or by some other module. Because the weights are variable, PAV can be adjusted to cause Nico to pay more attention to particular types of stimuli. After computing the weighted sum, PAV groups neighboring highly salient pixels together to form boxes. These boxes generally correspond to actual objects in the physical world. For example, if PAV is attending to color and a red ball is placed in front of it, it would box all of the adjacent red pixels together to form a box representing the location of that ball in the visual field. PAV can also be adjusted to change what types of pixels will be grouped together in the same box. For example, it can be set to box skin and color separately, so that a green toy and the hand holding the toy are treated as distinct objects. It can even be set to box different colors separately, so that a blue object and a yellow object that are adjacent in the visual field are kept distinct. Once PAV has found its boxes, it sends the most salient ones on to the next module.

There are several modules which may use the output from PAV as their input. Two such modules which are used for behavioral control are saccade and smooth pursuit. These modules represent two different types of eye movements found in humans. A saccade is a movement in which, after a person first becomes interested in an object, the person's eyes quickly move to a position in which they are fixated on that object. Smooth pursuit is used after a saccade has placed an object in the center of a person's field of view. As the object moves around in space, the person's eyes follow that object, so that it remains in the center of the person's field of view. In Nico, both of these modules receive boxes from PAV telling them where in the visual space objects of interest are located. They perform basic transformations to determine how far or quickly the eye rotation motors should move to compensate for any discrepancy between the most salient object's location and the center of the visual field. They produce output in the form of motor commands that are sent to an arbiter, which keeps track of whether Nico is currently engaging in a saccade or smooth pursuit and passes on the appropriate command to the motor module, which actually communicates with the eye motors.

Because Baillargeon conducted her experiments with a red object, a yellow object, and a green object, the color module was used as the sole input to PAV in the present experiment. PAV was set to box differently colored pixels separately, based on the assumption that infants have no trouble distinguishing between two adjacent objects of different colors. A separate instance of the color and PAV modules was run for each of Nico's eyes. Only the wide field of view cameras were used, as they provided sufficient detail for a foveal view to be unnecessary. The two PAV modules each passed their output boxes to depth, a module that matches up the PAV boxes from the left and right eyes and outputs the horizontal disparity between them. Disparities values give a rough idea of the relative distance from the eyes to an object, i.e. the object's depth. However, on their own they are only reliable if the eye cameras are parallel and stationary, so they could not be used as an indicator of absolute depth in the present experiment. The output from both the left eye's PAV module and the depth module were connected to the *memory* module.

5.2. The *memory* Module

The *memory* module receives PAV boxes as input and produces its own modified boxes as output. The module uses an array of elements to "remember" what it has seen over time. When it first receives a box from PAV, representing an object in the visual field it has not seen before, it matches the box up with a disparity from depth, representing the object's relative depth. It then associates one of its elements with the features of the new box. It stores the element's color and location in the visual field. If there is currently another visible element that has been present for some time, it associates the new element with that old element, storing the new element's location and depth, relative to the old element. This is a somewhat simplistic treatment of the infant cognition model, since there are presumably many more features that may be associated with an object. However, the *memory* module is limited by Nico's current perceptual abilities. The features mentioned are sufficient for replicating Baillargeon's experiment.

Every timestep, *memory* receives a set of boxes from PAV. It first checks to see whether those elements that were visible the previous timestep are still visible. Elements that are currently visible are associated with a location in the visual field and with a color. *Memory* then checks to see whether elements that were not visible the last timestep are now visible. If an element was last seen near its associated relative element, it will be strongly associated both with its last known depth relative to that element and with that element's location. Suppose element A disappears while directly adjacent to element B. If a box with the same color as A later appears near B, it will activate element A. However, if a box later appears adjacent to B and that box does not have the same color as A but the box is at the same relative depth, it can still result in the activation of element A.

When Nico is attending to a particular PAV box, i.e. when that box is the most salient, Nico habituates to the element associated with the box at an especially high rate. As the element's level of habituation increases, the box's saliency decreases. However, if the box to which Nico is attending begins moving after remaining stationary for some time, the associated element's level of habituation jumps down, and the box immediately becomes more interesting. After the box stops moving, the level of habituation returns to its previous level.

Because the details of how habituation works in an infant are unimportant for the present experiment, the *memory* module does not not use a realistic implementation of habituation. The implementation is simply designed to be sufficient for fitting the data from Baillargeon's experiment. While an element is visible, Nico habituates to the element at a steady rate. This continues until the element leaves Nico's sight or it reaches a saliency of 0, at which point Nico ignores the object with which the element is associated completely. When an element disappears from Nico's view, Nico remains habituated to the element for several seconds and then slowly begins to lose habituation. However, Nico does not lose all habituation to the element. The minimum level of habituation to an element, which begins at 0 when the element is first created, gradually increases as Nico habituates to an element. After the element leaves Nico's sight, Nico's level of habituation can drop no lower than this value. Thus, every time Nico sees the element, it takes Nico less time to habituate to the element. This allows Nico to habituate more quickly on each successive trial run, just as the infants did. The minimum level of habituate cannot rise higher than a little above 1/2, so trial times will not drop significantly below half of the initial trial time.

5.3. The Experiment

Baillargeon's experiment was replicated by building a short metal ramp. A toy train with a yellow piece of poster board affixed to it represented the yellow cart. It traveled down the ramp behind a thin red screen, the occluder, which could be lifted and lowered. A thin green folder represented

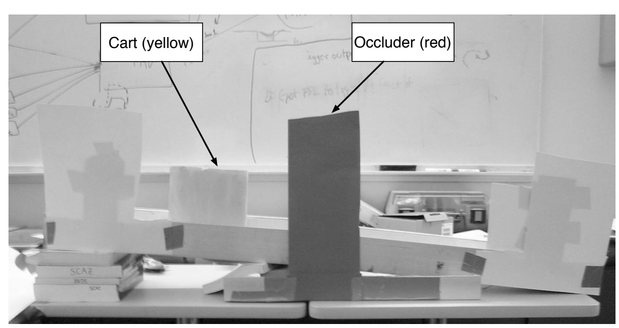


Figure 3. The Setup: the cart moves down the ramp, behind the occluder

the green box. It could either be clipped to the front of the track immediately behind the red occluder or be held up in a vice a good distance behind the occluder. Because Nico's depth perception is poor, the green folder appeared to be at the exact same depth as the yellow cart when it was clipped to the ramp. The fact that it was actually a thin folder rather than being a box placed across the ramp is irrelevant since Nico has no understanding of three-dimensional objects. The trials were run in a dark room with a light shining directly onto the ramp. No other brightly colored objects in the room received enough light to distract Nico.

Because Nico's habituation to objects is free of the myriad arbitrary factors that may affect infants, a simpler criterion was used for ending the habituation trials. Instead of ending when Nico concluded three consecutive trials in at most half the time of the first three trials, the trials simply ended when Nico concluded a single trial in half the time of the first trial. After each trial, the red occluder was covered so that there were no salient objects in Nico's field of view, and Nico was given time to lose habituation to the elements in its memory. If a shortcut hadn't been programmed in to speed up this process, it would have required a delay of at most 80 seconds between each trial, which does not seem to be an unreasonable amount of time. Because Nico has no understanding of the objects it sees beyond their association with elements in Nico's memory, the familiarization trials were deemed unnecessary and were not used.

Finally, because there simply was not time to run the experiment in two-second segments, the segments were lengthened to three seconds, meaning that the entire scene repeated at 12-second intervals rather than 8-second intervals. This also meant that Nico had to lose interest in the scene for three seconds before a habituation trial ended. The criterion for ending a trial was further increased to four seconds so that occasional mistakes by the experimenter that might have caused one of the boring segments to last for slightly longer than three seconds would not result in a trial ending prematurely. These changes should not have affected on the generalizability of the experiment, as Baillargeon did not claim there was anything significant about using 8-second cycles.

6. Results

Nico's state was saved after each trial so that if a component of the system should crash, the trials could be continued from the same point with the same habituation values used for each element. The data that will be reported was obtained from a single series of trials. However, because there is no random factor in Nico's system, there is no reason to suspect there would be any variation in multiple runs of the experiment. On occasions when trials were repeated, comparable results were achieved, with any variation resulting from differences in the way the scene was presented by the experimenter. In the initial trial run, Nico met the criterion for losing interest in the scene after 61 seconds. On the sixth trial, Nico lost interest after 25 seconds, meeting the criterion for ending the habituation trials. These results are similar to those achieved by Baillargeon, who found that infants on average showed a 50% decrease in trial times in three consecutive trials after the eighth trial.

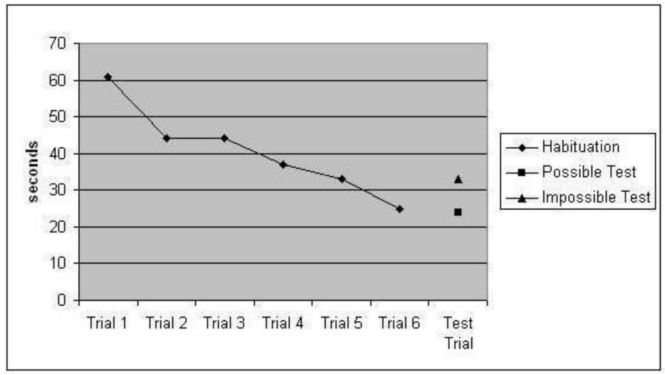


Figure 4. Time required to habituate in the habituation and test trials

After the sixth trial, Nico's state was saved. The saved state was tested on the "possible" and "impossible" test trials. On the "possible" test trial, in which the green folder was placed far behind the red occluder, the green folder failed to interfere with Nico's performance because its associated element was only activated while the red occluder was up. This trial took 24 seconds, about the same as the sixth habituation trial. On the "impossible" trial, in which the green folder was located directly behind the red occluder, the green folder's associated element was activated immediately before and immediately after the yellow cart moved behind the red occluder. This activation did not interfere significantly with the habituation to the yellow cart simply because Nico had already been exposed to the cart on six previous trials, and so it took Nico very little time to habituate to the cart. However, the activation did prevent Nico from losing interest in the scene when it otherwise would have because the activation of the green folder persevered for several timesteps after Nico had fully habituated to the yellow cart. As a result, the "impossible" trial lasted 33 seconds, 9 seconds longer than the "possible" trial. This difference was similar to the difference that Baillargeon found between the possible and impossible trials.

7. Discussion

The result from the "impossible" test trial is particularly interesting because while the increase in total time for the trial was expected, the reason for the increase was unexpected. The prediction was that the ”impossible” trial would take longer because Nico would habituate to the yellow cart more slowly, whereas what actually happened was that the trial took longer because Niko took an interest in the imagined appearance of the green element. This result demonstrates one of the great advantages of robot studies: no matter how well thought-out a model may be, it is impossible to say for certain how the model will work until the model is tested in the real world. There are simply too many factors to consider all of them in theory or test all of them in simulation. With a robot, a scientist can test a theory in the real world while at the same time being able to look into the robot's head and see exactly how its cognitive operations are interacting with the feedback from the world.

Unfortunately, the results achieved in this experiment are not perfect. While they do match those predicted by the cognitive model, they do not entirely match the results reported by Baillargeon. Although Baillargeon does not report exactly how long the habituation trials took on average, she does say that trials were automatically ended if they took more than 60 seconds, implying that the first habituation trial lasted slightly under a minute, while the last habituation trial lasted well under 30 seconds. However, she reports that the ”possible” and ”impossible” test trials took about 48 seconds and 61 seconds, respectively. These results seem to suggest that even in the ”possible” trial, the mere presence of the green block in the scene caused the infants to find the scene as a whole more interesting, resulting in a longer looking time. If this is true, then infants are habituating to the entire scene, rather than to individual objects within the scene, meaning the theory proposed in this paper is incorrect.

However, there are alternative explanations. Baillargeon did not attempt to standardize the delays of time between trials. It is possible that the delays before or after the familiarization trials were longer, giving the infants more time to forget the objects before viewing the test trials. It is even possible that the infants lost habituation to the elements associated with the objects during the familiarization events themselves, perhaps because the infants were used to seeing the yellow cart moving rather than stationary. A more accurate reenactment of Baillargeon's experiment, perhaps using her own apparatus, might help to clear up this quandary.

One could easily argue that the findings from this experiment are not generalizable to most looking time studies. The *memory* module in its current form uses only a small number of features. Even one of these features, depth, is not entirely dependable, meaning that *memory* had to be modified to throw out noisy depth values, values that in another experiment might actually be valid. It seems as though the *memory* module has been designed specifically for the purpose of replicating Baillargeon's results. However, it is important to remember that Nico is still in the process of being developed. As new modules that extract new features are built for Nico, Nico will be able to more closely approximate the suggested model of infant cognition. That model was not merely designed to match Baillargeon's results. It was also meant to be applicable to other looking time studies. As individual studies are tested out using Nico, the model can change to accommodate those studies. Eventually, the model may have to be thrown out entirely. After all, it is designed to be a possible model, not a correct model. As long as it or some other model that can be tested with a robot remains a possibility, experimenters who use looking times to study infants will have to concede that their interpretations of cognitive abilities may be a little too generous.

References

[1] B. Adams et al., "Humanoid Robots: A New Kind of Tool," *Intelligent Systems and Their Applications: Special Issue on Humanoid Robots*, vol. 15, no. 4, IEEE Computer Soc. Press, Los Alamitos, CA, July/Aug. 2000, pp. 25-31.

[2] R. Baillargeon, "Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month-old infants," *Cognition*, vol. 23, no. 1, Elsevier Science, Netherlands, June 1986, pp. 21-41.

[3] R. Baillargeon and J. DeVos, "Object permanence in young infants: Further evidence," *Child Development*, vol. 62, no. 6, Blackwell Publishers, US, pp. 1227-1246.

[4] M. Haith, "Who put the cog in infant cognition? Is rich interpretation too costly?", *Infant Behavior & Development*, vol. 21, no. 2, Ablex Publishing Corp., US, April-June 1998, pp. 167-179.

[5] Kandell, E. R., J. H. Schwartz, and T.M. Jessel, *Principles of Neural Science: 4th edition*, McGraw-Hill Professional Publishing, New York, 2000.

[6] M. Schlesinger, "A lesson from robotics: Modeling infants as autonomous agents," *Adaptive Behavior*, vol. 11, no. 2, Sage Publications, US, Spr. 2003, pp. 97-107.

[7] P. Sharp, , J. James, and A. Wagner, "Habituation of a 'blocked' stimulus during Pavlovian conditioning," *Bulletin of the Psychonomic Society*, vol. 15, no. 4, Psychonomic Society, Inc., US, Mar. 1980, pp. 139-142.

[8] K. Wynn, "Addition and subtraction by human infants," *Nature*, vol. 358, no. 6389, Nature Publishing Group, United Kingdom, Aug. 1992, pp. 749-750.

Colors and sensorimotor theory, or what do we really perceive?

D. Philipona, O. J.-M. D. Coenen
Sony CSL, 6 rue Amyot
75005 Paris, France

J. Kevin O'Regan
Lab. de Psychologie Exprimentale, CNRS
92774 Boulogne-Billancourt Cedex

Abstract

Color perception is a scathing expositor of the very different points of view that exist about perception. Besides the philosophical debate about phenomenal experience, colors raise fundamental questions about the contribution of innate and acquired knowledge at the perceptual level, and about the respective weight of neuronal and environmental constraints in learning. Colors provide a specially incisive testbed because they show, in addition to these difficult questions, that the identification of the object of perception is not always a simple matter: psychophysical experiments indeed support the idea that color perception is concerned with the reflecting properties of surfaces[1], rather than with light per se as it is often assumed.

The case of color perception recalls an obvious point: before addressing the way neuronal adaptation takes place in living organisms (be it at the phylogenetical or ontogenetical scale), it is mandatory to question what they adapt to. This is not a simple question because the nervous system can rely to such an extent on indirect cues (spectral composition of light) to estimate the object of perception (surface reflectance), that the presentation of these cues elicits a sensation without the presence of the actual object of perception. It is thus difficult to distinguish, practically and conceptually, what is cue and what is object of perception. Further, to make things even more complicated, it is problematic to satisfy oneself with an understanding of the physical object of perception without taking into account the under-determinacy of the sensorimotor system. For instance, it is obvious that reflecting properties of surfaces concerned with lights outside of the visible spectrum are irrelevant for color perception. But just as important is the fact that there are much less obvious aspects within *the so-called visible spectrum that are irrelevant as a result of the few photopigments that biological organisms possess.*

We will show how to interpret the physical notion of reflectance in a biological way, so that it involves only those aspects of the light/surface interaction that are relevant with respect to a given set of molecular photopigments. From this it will become apparent that surfaces exhibit categorical differences in the way they modify those aspects of light's spectral composition that impact the photopigments of that set. With the three photopigment kinds usually assumed for the human visual system, this biological interpretation predicts eight special colors of three different kinds, corresponding to white/black, red/green/blue and yellow/purple/cyan. These differences will finally be shown to correspond to differences in the sensorimotor contingency[2] that the organism engages in when visually exploring colored surfaces.

References

[1] D. H. Brainard. Color constancy. *The Visual Neurosciences*, 2003.

[2] J. K. O'Regan and A. Noë. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), 2001.

Cross-anchoring for binding tactile and visual sensations via unique association through self-perception

Yuichiro Yoshikawa, Koh Hosoda*, and Minoru Asada*
* HANDAI Frontier Research Center,
Graduate School of Engineering,
Osaka University
Yamada-Oka 2-1, Suita, Osaka Japan
yoshikawa@er.ams.eng.osaka-u.ac.jp, {hosoda, asada}@ams.eng.osaka-u.ac.jp

Abstract

Binding is one of the most fundamental cognitive functions, how to find the correspondence of sensations between different modalities such as vision and touch. Without a priori knowledge on this correspondence, binding is regarded to be a formidable issue for a robot since it often perceives multiple physical phenomena in its different modal sensors, therefore it should correctly match the foci of attention in different modalities that may have multiple correspondences each other. Learning the multimodal representation of the body is supposed to be the first step toward binding since the morphological constraints in self-body-observation would make the binding problem tractable. The multimodal sensations are expected to be constrained in perceiving own body so as to configurate the unique parts in the multiple correspondence reflecting its morphology. In this paper, we propose a method to match the foci of attention in vision and touch through the unique association by cross-anchoring different modalities. Simple experiments show the validity of the proposed method, and future issues are discussed.

1 Introduction

Binding is one of the most fundamental cognitive functions, how to find the correspondence of sensations between different modalities such as vision and touch, both of which are major sources of perception not only for the external world but also for the agent's body itself. The latter is closely related to the body representation which is often given by the designer and fixed but has much influence on the adaptability to the changes in the environment and the robot body itself. Assuming that the designer does not give any explicit knowledge on the body representation, a robot should construct its body representation only from its uninterpreted multimodal sensory data. In this process, *binding* has a significant role.

Recently, researchers in other fields focus on the binding problem, which concerns the capability to integrate information of different attributes [1]. Although there are already some models of binding, for example based on attention [2], firing in synchrony [3, 4], and so on, it has not been still clear how to bind different sensor modalities since they focused on the binding problem between visual attributes. To propose the model for the cross-modal binding mechanism of humans based on a constructivist approach, we should start with an assumption that the designer does not give any *a priori* knowledge on what the robot's sensors receive, but the robot can discriminate the different sensor modalities such as vision and touch. Generally, receptive fields for touch and vision are simultaneously stimulated, but often respond to different physical phenomena since the foci of attention in these modalities are often different. In other words, the robot does not always watch its touching region. Therefore, to bind different modalities, the robot should correctly match the foci of attention in different modalities that may have multiple correspondences each other.

We suppose that learning the multimodal representation of body should be the first step toward binding since the morphological constraints in self-body-observation would make the binding problem tractable. The multimodal sensations are expected to be constrained in perceiving own body so as to configurate the unique parts in the multiple correspondence reflecting its morphology. Therefore, building a robot that can acquire the representation from multimodal sensory data is an interesting issue from a viewpoint of a constructivist approach towards both establishing the design principle for an intelligent robot and understanding the process how humans acquire their body representation [7]. In this study, as an example of the binding problem, we focus how it can learn to watch its body part when it detects the collision on it. The previous work on the issue of acquiring body representation escaped from this kind of problem by assuming that it can observe only matched sensations in different modalities (ex. [5, 6]).

Yoshikawa et al. have proposed the method to learn the multimodal representation of the body surface through *double-touching*, that is touching its body with its own body part [8]. It is based on the fact that double-touching co-occurs with *self-occlusion*, that is the occlusion caused by covering its body with its own body part in its view. Although they did not take multiple self-occlusions caused by the physical volume of the body into account, which makes the binding problem remain formidable, it seems still reasonable to utilize the fact that double-touching co-occurs with self-occlusions. In this paper, we presents a method to match the foci of attention to its own body in vision and touch by virtue of the morphological constraint in the relationship between double-touching and self-occlusion. In the proposed method, the mismatched responses in these modalities can be discarded through the process of *unique association* where corresponding pairs of subsets in different attributes are exclusively connected with each other by what we call *cross-anchoring*.

In the rest of the paper, first what kind of problem is to be solved for *binding* in different modalities is explained. Then, a possible developmental course towards binding and a basic idea utilizing the morphological constraint of the human-like robot's body to perform binding are argued. After introducing an anchoring Hebbian learning rule to perform unique association, we show preliminary computer simulations of the robot which has 1-DoF or 2-DoFs arm and a camera head to test the proposed learning rule works. Finally, discussion with future issues is given.

2 The binding problem in different modalities

In order to propose the model for the binding mechanism of humans based on a constructivist approach, we should start with an assumption that the designer does not give any *a priori* knowledge on the robot body representation, but a robot can discriminate the different sensor modalities such as vision and touch. Therefore, the problem for the robot is how to associate these different sensations to build its body representation needed to accomplish various tasks such as collision avoidance and object manipulation. In this study, as an example of the binding problem, we focus how it can learn to watch its body part in which it detects the collision. So what is the problem?

There is a series of studies on modeling the brain mechanism of binding different visual attributes that are processed in segregated areas, such as form, shape and color [3, 4]. They built a system capable of binding these attributes through the dynamic process called reentry with topographic connections between the segregated groups of neurons and made several suggestions to understand the mechanism of binding in vision. However, since the topographical relationship between different modal sensors depends how they are embedded in the body, we need to release the assumption of *a priori* topographic connection in order to cope with the binding problem between different modalities.

This kind of binding problem has not been focused so far although it is an important issue for acquiring the multimodal representation of the body. In some previous studies on own body representation with both tactile and visual sensors, the designer provided the agent with the competence to detect the position of the touch sensors in its view [6] or assumed that there is only one object which can collide with its body [5]. In other words, the binding problem is solved by the designer instead of the robot itself by assuming that it can observe only matched sensations in different modalities. However, there are often multiple visual responses to both the body and nonbody that co-occur with the tactile one on the body since the agent watches multiple objects at a moment. Therefore, the robot must determine which visual ones should be bound.

On the issue how to find the matched sensations in these modalities, Yoshikawa et al. have proposed the cross-modal map among vision, touch, and proprioception to learn the representation of the body surface [8]. It is based on the idea that the tactile sensors which collide with each other also coincide with each other in its vision. In other words, touching its body with its own body part (i.e., *double-touching*) always co-occurs the occlusion caused by covering its body with its own body part in its view (i.e., *self-occlusion*). They assumed that there is only one self-occlusion at a moment. However, there can be multiple self-occlusions since the body occupy a certain volume in the physical space. For example, when the agent touches its body trunk with its hand, not only the hand but also its arm cover its body trunk from its sight, therefore, multiple self-occlusions occur. Therefore, there still remains the binding problem where it must determine which self-occlusion should be bound to the double-touching and vice versa.

3 A basic idea

As suggested from the previous work [8], it seems reasonable to utilize the fact that double-touching co-occurs with self-occlusion although they did not take the physical volume of the body into account, which makes the binding problem remain formidable. We will explain a basic idea how the robot can correctly match double-touching and self-occlusion based on this fact. In the following, first we introduce the assumptions what kinds of cognitive competences it should possess and argue a possible developmental course to acquire them. Then, we show a basic idea of cross-anchoring to solve the binding problem by virtue of the morphological constraint. In the following argument, we suppose that it has a human-like configuration in which it has a trunk with a camera and an end-effector connected through serial links, that is, the robot consists of its trunk, a camera head and an arm.

3.1 A possible developmental course of prerequisite competences for binding

We assume that the robot has acquired the competences to detect double-touching and self-occlusion. However, it is worthy to argue how such assumptions are justified from a viewpoint of robotics and analogy to the human development. We propose a possible developmental course of the prerequisite competences, which consists of three stages: 1) learning to detect double-touching, 2) finding the major components of visual changes caused by its own camera head motions, and 3) learning to detect self-occlusion.

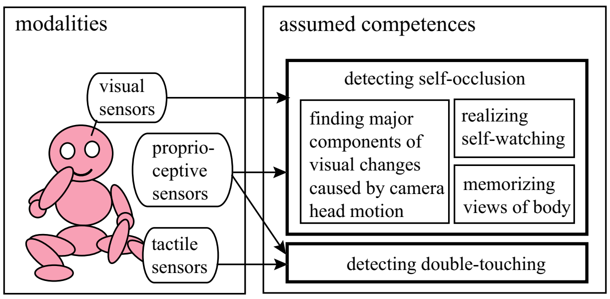
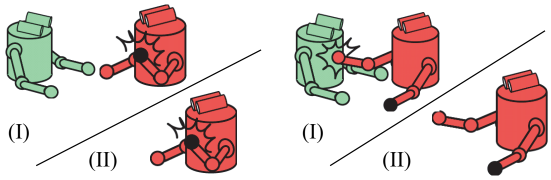


Figure 1. The modalities of the robot and the competences supposed to possess

Learning to detect double-touching At the first stage, a robot learns to detect double-touching through the iterations of double-touching. According to the assumption in the previous work [9] that the sensation of its body is invariant with its posture (see Fig. 2), it can judge whether the tactile sensation is caused by double-touching after learning the tactile-proprioceptive map that represents how invariant with its postures the tactile sensations are. Therefore, it can detect double-touching when it occurs after this stage.

Since it is observed that a human fetus touches its body with its hands in the womb [10] and it is reported that a human neonate can distinguish double-touching from being touched by the other in the study on rooting reflex [11], it seems biological plausible to assume that an infant has already possessed the competence to detect double-touching before the following stages with the visual sensation.



(a) a posture with the invariant touch

(b) a posture with the variant touch

Figure 2. Examples of the invariance (a) and variance (b) of the tactile sensation with certain postures in the different environment (I) and (II)

Finding major components of visual changes caused by camera head motion This is prerequisite for detecting self-occlusion. When the robot moves, an optical flow is induced in its vision. It is considered that the motion of the camera mainly contribute to this optical flow due to the following two facts: (1) the motion of the camera usually induces larger optical flow components in wider region than one of the arm and (2) what a human-like robot observes is not usually its arm but the environment due to its configuration. In other words, the motion of the camera head would often predict the changes in the optic flow of a robot with human-like configuration.

Therefore, the robot can find that the major components of visual changes by finding the principle component of the motion to predict the optic flow as performed in the previous work [12]. Human infants can not hold their heads up in their first several month [13], and therefore, usually lie on the bed. It can be conjectured that this kind of immaturities constraints infant's motion so that the neck motion is easily found since it usually watches external world rather than its arm in such a situation.

Learning to detect self-occlusion According to the same idea used in the case of double-touching, a robot learns to detect the occurrence of self-watching, that is watching its own body, by judging whether the visual sensations is invariant with its posture. If the robot learned the invariance with respect to the posture of the camera head, it is expected that it can detect the occurrence of self-watching of its trunk since the visual sensations of the trunk is invariant with the posture of the camera head. Once it learned to detect the occurrence of self-watching and memorized the invariant visual sensations with each posture, it can detect the occurrence of self-occlusion by comparing the current sensations and the memorized one.

3.2 A learning rule for the binding problem

Since the robot does not have *a priori* knowledge how to bind, we suppose that it keeps changing the posture both of its arm and its camera head at random to explore for binding. It perceives its posture and the view in the center region of its camera. Fig. 3 illustrates an example of self-watching view of a human-like robot (imagine that the robot watches its body in sitting). Fig. 4 illustrates the simplified situations of the robot's exploration for binding in Fig. 3, where DoFs for the arm is simplified to one to slide its end-effector while DoFs for the camera head is simplified to one to displace the camera. Note that the notations in Fig. 4 correspond to those in Fig. 3.

Concerning a point on the trunk, in this case B, there are following five types of experience of the robot (See Fig. 4). In each experience, the robot detects the posture of its arm, namely, Θ_B in Fig. 4(a)-(c), Θ_C in Fig. 4(d), and Θ_A in Fig. 4(e). In the case where the robot double-touches with B and tries to self-watch B (Fig. 4(a)), it detects the occurrence of self-occlusion. In the other cases where the robot double-touches B, it sometimes detects self-occlusion (Fig. 4(b)) but sometimes does not (Fig. 4(c)) depending on the posture of its camera head. In the other cases where it self-watches B, it sometimes detects self-occlusion (Fig. 4(d)) but sometimes does not (Fig. 4(e)) depending on the posture of its arm. The robot can not distinguish the self-occlusion by its end-effector (Fig. 4(a)) from self-occlusions by the link (Fig. 4(b) and 4(d)). Note that these kinds of experiences uniformly occur for each part on the trunk since the robot explores at random.

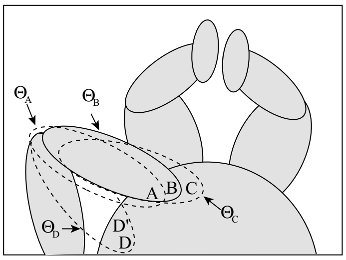
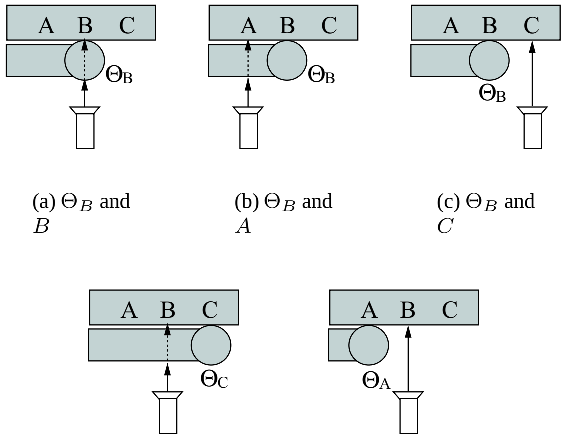


Figure 3. An example of self-watching view of the robot in double-touching: three ellipsoids with solid and broken lines are links in the postures (Θ_A, Θ_B, and Θ_C) by which it touches with the parts labeled A, B, and C, respectively. The rest ellipsoid with the symbols D and D' is one in the posture (Θ_D) by which it touches the part labeled D with occlusion at D and D'.



(a) Θ_B and B

(b) Θ_B and A

(c) Θ_B and C

(d) Θ_C and B

(e) Θ_A and B

Figure 4. Five simplified situations in the robot exploration: A top rectangle indicates the robot's trunk while a rectangle with a circle in the middle indicates the robot's arm that has a sliding DoF in the horizontal axis. The bottom object shows the posture of the robot's camera head while the arrow indicates its focus of attention. Notations correspond to those in Fig. 3.

A problem in the statistical approach As mentioned in section 2, the fact that the body occupies a certain volume in the physical space remains binding problem formidable. For example, a double-touching posture causes self-occlusions at multiple parts (see Figs. 4 (a) and (b)) while a self-occlusion at a part is caused by several double-touching postures (see figs. 4 (a) and (d)).

In the explorations, the robot sometimes experiences the matched responses in the different modalities which are caused by focusing on the same region, in this case detecting the self-occlusion at the double-touching point (see Fig. 4 (a)). However, such experiences of the matched response are not significantly frequent compared to mismatched responses (see Fig. 4 (b) or (d)) since it explores at random instead of utilizing *a priori* knowledge. In other words, the correctly matched responses are not significantly major in the obtained data. Therefore, it is difficult to associate them by considering all obtained data through the exploration. Then, we need a mechanism to narrow down the influence of the mismatched data on learning while augmenting the influence of the matched one.

Cross-anchoring association Due to the morphological constraints on the human-like configuration, we can utilize the following two morphological constraints: 1) how many double-touching postures occlude a certain part on the trunk depends on the location of the part to be occluded, and 2) how many parts the robot occlude by a double-touching posture depends on the location of the contact part.

These facts indicate that there exist *cue* nodes which have fewer candidates for matched response in other modalities to be bound. In this case, the self-occlusion at D can be the cue of double-touching in Θ_D while a double-touching in Θ_A can be the cue of the occlusion at A. Note that the matched response of a cue node is also experienced with mismatched response. For example, the robot sometimes detects a self-occlusion also at D' during the cue double-touching in Θ_D while it sometimes detects a double touching in Θ_B during detecting the cue self-occlusion at A. Since the desired correspondence between touch and vision can be found by unique association in this case, we can utilize such cue nodes as anchors of the unique association. Therefore, we introduce a learning rule with an anchoring mechanism which can adapt the learning rate according how much the responses simultaneously observed are regarded as unique to each other.

4 Cross-anchoring Hebbian learning rule

In this section, we introduce an cross-anchoring Hebbian learning rule as an implementation of the learning rule with the anchoring mechanism. The architecture consists of two layers called the double-touching layer and the self-occlusion one (see Fig. 5). In the double-touching layer, there are N_t nodes each of which is responsible for a set of certain posture of the arm $\Theta_i, (i = 1, \cdots, N_t)$ which is assumed to be quantized in advance. When the posture of the arm is $\boldsymbol{\theta} \in \Re^m$, the activation of the i-th node is calculated by

$$ {}^{t}a_i(\boldsymbol{\theta}) = \begin{cases} 1 & \boldsymbol{\theta} \in \boldsymbol{\Theta}_i \\ 0 & \text{otherwise} \end{cases}. \quad (1) $$

On the other hands, in the self-occlusion layer, there are N_o nodes each of which is responsible for the self-occlusion in a set of certain posture of the camera head $\Phi_j, (j = 1, \cdots, N_o)$ which is assumed to be quantized in advance. When the posture of the camera head is $\boldsymbol{\phi} \in \Re^n$, the activation of the j-th node is calculated by

$$ {}^{o}a_j(\boldsymbol{\phi}) = \begin{cases} 1 & \boldsymbol{\phi} \in \boldsymbol{\Phi}_j, O \\ 0 & \text{otherwise,} \end{cases} \quad (2) $$

where O is the phenomenon of detecting occlusion.

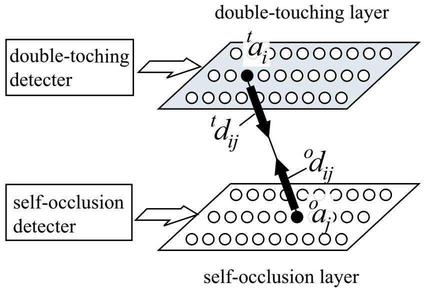


Figure 5. The architecture

Let a connection weight between the i-th neuron in the double-touching layer and the j-th neuron in the self-occlusion layer be w_{ij}. By the cross-anchoring Hebbian learning rule, $w_{i^*j^*}$ is updated as following:

$$ \Delta w_{i^*j^*} = \eta({}^{t}d_{i^*j^*}\,{}^{t}a_{i^*} \cdot {}^{o}d_{i^*j^*}\,{}^{o}a_{j^*} - w_{i^*j^*}), \quad (3) $$

where i^* and j^* are the most activated units in the double-touching and the self-occlusion layer, η is a constant learning rate. The dynamic anchoring rates, ${}^{t}d_{ij}$ and ${}^{o}d_{ij}$, determine the degrees of anchoring on the j-th node in the self-occlusion layer from the i-th nodes in the double-touching layer and on the i-th node in the double-touching layer from the j-th nodes in the self-occlusion layer, respectively. They are calculated by

$$ \begin{aligned} {}^{t}d_{ij} &= \exp\left(-\frac{\sum_{k,k\neq j} w_{ik}}{{}^{t}\sigma^2}\right), \\ {}^{o}d_{ij} &= \exp\left(-\frac{\sum_{k,k\neq i} w_{kj}}{{}^{o}\sigma^2}\right), \end{aligned} \quad (4) $$

where ${}^{t}\sigma$ and ${}^{o}\sigma$ are parameters that determine the degree of anchoring. Meanwhile, the remaining connection weights are decreased because they lost the competition;

$$ \begin{aligned} w_{ij^*}(t+1) &= w_{ij^*}(t) - \eta_t(1 - {}^{t}d_{ij^*})\Delta w_{i^*j^*}, \\ w_{i^*j}(t+1) &= w_{i^*j}(t) - \eta_o(1 - {}^{o}d_{i^*j})\Delta w_{i^*j^*}, \end{aligned} \quad (5) $$

where η_t and η_o are constant coefficients of the competition.

In such an anchoring process, more unique combinations of double-touching and self-occluded are bound earlier. Meanwhile, some of the rest combinations become more unique since the other responses decrease the number of candidates to be bound by losing the responses that are already bound to others. Therefore, the process of binding proceeds step by step. This process is expected to converge since it is considered that there exist anchoring sensations in each modality due to the constraint in its human-like configuration.

5 Simulation results

As preliminary experiments, we tested the cross-anchoring Hebbian learning rule works so that the robot solves the binding problem by using the computer simulation. First, we examined a robot with a single DoF to show how learning proceeds. Then, we examined a robot with more DoFs.

5.1 Simulation with 1-D robot

In the first experiment, we simulated a robot with a 1-DoF sliding arm, a 1-DoF rotating camera head. During the exploration for binding, it moves its arm and camera head at random and detects self-occlusion and double-touching. Fig. 6 shows an example of self-watching view of the simulated robot in double-touching. For the reader's understanding, we quantized the posture space both of the arm and the camera head so that the nodes in both layers were matched with each other in one-to-one manner. The robot was trained for binding in 4,000 double-touching trials with the following network parameters: $N_t = N_o = 10$, $\eta = 0.1, \eta_t = \eta_o = 0.5$, and ${}^t\sigma = {}^o\sigma = 1.0$.

Figs. 7 (a):(I) $\sim$ (V) show the process of learning connection between double-touching layer and self-occlusion one. It can be seen that it starts with multiple connections and finally succeeded in binding since it obtained the correct one-to-one mapping at the 4,000-th step. Furthermore, we can see that the connections grew up both from the right and left ends to the center. It seems to show the process that cross-anchoring between a pair of nodes seems to make neighbor pairs of nodes more unique to each other and therefore guides cross-anchoring between the neighbor pairs. Such propagation of cross-anchoring starts from the pairs of nodes, either of which is a cue node. Consistently with the analysis of the learning procedure, the left end node in the bottom layer and the right end node in the top layer were cue nodes due to the morphological constraints. In this case, since the camera and the end-effector were connected through a serial link, how to double-touch and how to self-occlude were constrained. For example, the double-touching at the left end of the trunk could guide the self-occlusion only at the same part while the self-occlusion at the right end could be caused by the double-touching only at the same part.

Figs. 7 (b):(I) $\sim$ (IV) show the process of the learning connections in the case that the posture spaces of the camera head and the arm were quantized in different resolutions. In this case, the resolution of the double-touching was twice in the case of self-occlusion. The parameters were $N_t = 12$, $N_o = 6$, $\eta = 0.1, \eta_t = 0.5, \eta_o = 0.25$, and ${}^t\sigma = {}^o\sigma = 1.0$. Since it finally obtains the desired one-to-many mapping, we may conclude that it succeeds in binding despite the different resolutions.

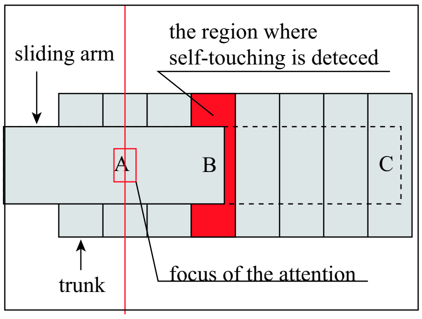


Figure 6. An example of self-watching view of the simulated 1-D robot in double-touching: The biggest gray box is its trunk while the other gray box is its sliding arm. Although each part labeled by a symbol is corresponding to Fig. 3, it is supposed that its trunk is a plane and the DoFs for the motion of its arm is one for the simplicity. The small box on the horizontal line indicates the focus of the attention in vision. The black box indicates the region where double-touching is detected.

After these processes, when the robot double-touches its body trunk, it can use the acquired mapping to know how to shift the focus of attention in the vision to the double-touching part by propagating the activation of the nodes responsible for the double-touching through the learned connection. Shortly, it can watch its touching part on its body.

5.2 Simulation of 2-D robot

In the second experiment, we simulated a more realistic robot with a 2-DoF rotating arm, a 2-DoF rotating camera head. Taking a posture of the camera head was emulated by changing the focus of the attention in vision. Fig. 8 shows an example of self-watching view of the simulated robot in double-touching. For the reader's understanding, we quantized the posture space both of the arm and the camera head so that the nodes in both layers were matched with each other in one-to-one manner. We let the robot learn the connections 100 times in each of 100,000 double-touching trials with the following network parameters: $N_t = 25$, $N_o = 25$, $\eta = 0.1, \eta_t = \eta_o = 2.0$, and ${}^t\sigma = {}^o\sigma = 3.0$.

Fig. 9 shows the process of learning connections between double-touching layer (top) and self-occlusion one (bottom). It can be seen that it started with multiple connections and finally succeeded in binding since it obtained the correct one-to-one mapping at the 100,000-th step. Fig. 10 shows the averaged time course of the matching errors of

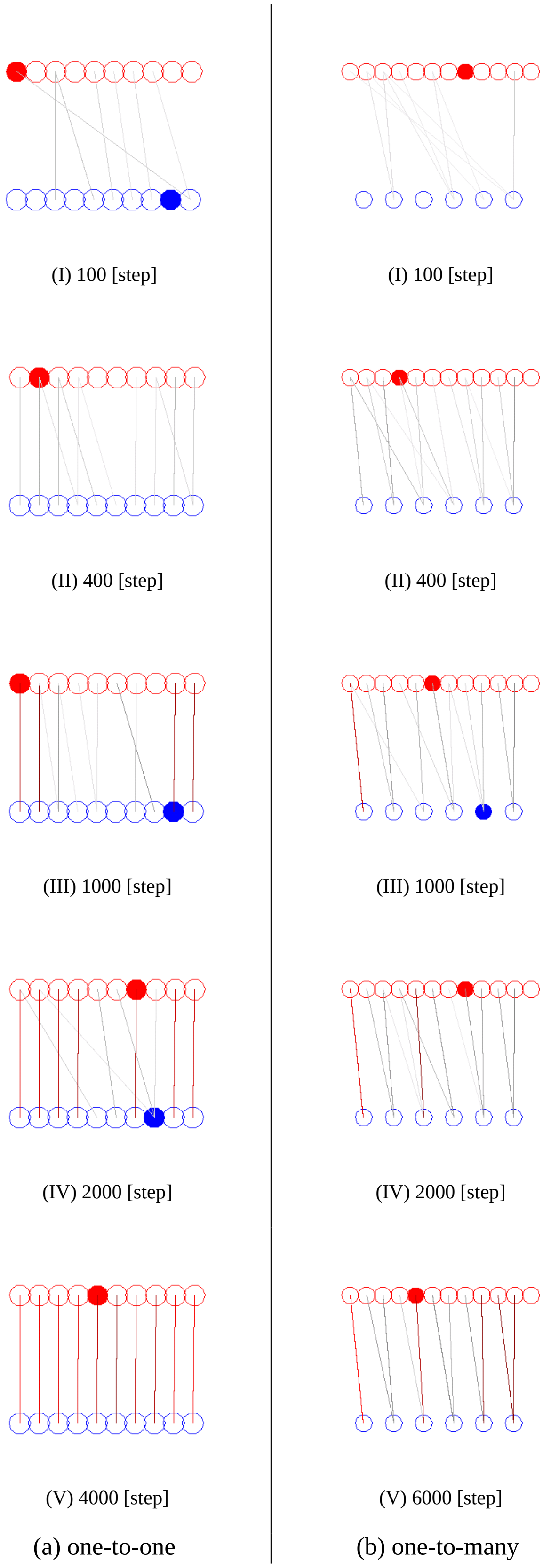


Figure 7. The process of learning connection between the layers ((I) $\sim$ (V)) with the same resolution (a) and with the different one (b).

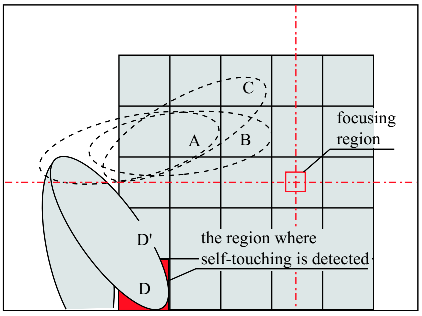


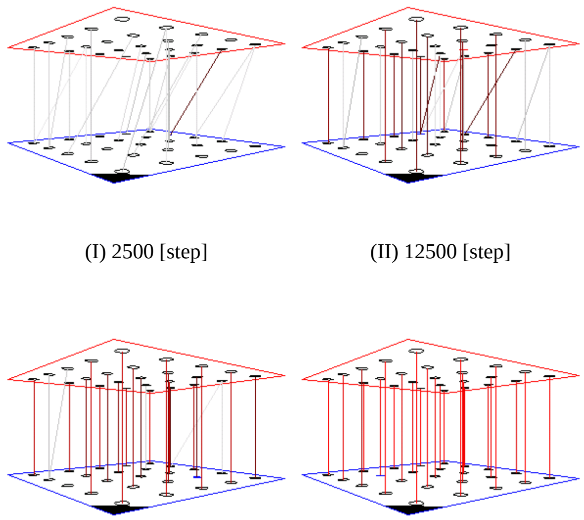
Figure 8. An example of self-watching view of the simulated robot in double-touching: Three ellipsoids with broken curves and one with the symbols D and D' are the postures of the robot arm by which it touches with its trunk. Although each part labeled by a symbol is corresponding to Fig. 3, the trunk is supposed to be a plane for the simplicity. The cross point of the vertical and the horizontal chain lines indicates the focus of the attention in vision. The black box indicates the region where double-touching is detected.

100 times in which the processes of exploration are different. Note that only the experimenter knows the desired connection and can determine the matching error. We can see that learning for binding converged to almost completely correct one-to-one mapping with small variance. Therefore, we may conclude that the robot can robustly find the matched response in different modalities by the anchoring Hebbian learning rule in a more realistic embodiment.

6 Conclusion

In this paper, we address the issue how to solve the binding problem in different modalities for body representation. We propose a method called cross-anchoring Hebbian learning rule to perform binding by virtue of the morphological constraints of its human-like configuration in perceiving the self. In the preliminary computer simulations, we showed that the robot can bind its tactile and visual sensations through the exploration by the iterations of self-watching and double-touching.

There are parameters in the proposed learning rule that determine how much the degree of anchoring is. Since it should be well selected to obtain the unique association, we should put a mechanism to adapt it when the system fail to bind. Topographical constraint caused by the the receptive fields with continuity that reflects the physical continuity could be a criteria for adaptation. Furthermore, the robot



(I) 2500 [step] (II) 12500 [step]

(III) 25000 [step] (IV) 100000 [step]

Figure 9. The process of learning connection between the layers with the same resolution in the 2-D robot simulation

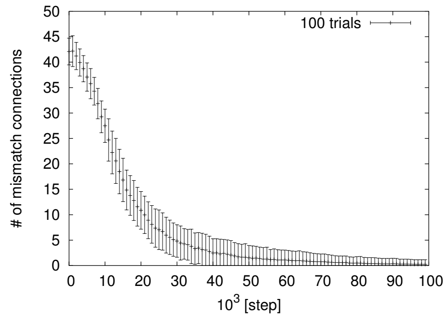


Figure 10. The number of the mismatched connections: A connection is counted as mismatched if it is to a mismatched neuron with more than 0.1 strength and if it is to the matched one with less than 0.8 strength.

needs the competence of binding in the case where it learns multimodal representation of the external objects. Although we concentrated on the binding problem concerning the self body in this paper, extending the proposed method for the binding problem involving tactile sensations of being touched by others is one of our future work.

Acknowledgment The Advanced and Innovational Research program in Life Sciences of the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government and a Research Fellowship for Young Scientists from Japan Society for the Promotion of Science supported this research.

References

[1] A. Treisman. Solutions to the binding problem: Progress through controversy and convergence. *Neuron*, Vol. 24, pp. 105–110, 1999.

[2] J. H. Reynolds and R. Desimone. The role of neural mechanism of attention in solving the binding problem. *Neuron*, Vol. 24, pp. 19–29, 1999.

[3] G. Tononi, O. Sporns, and G.M. Edelman. Reentry and the problem of integrating multiple cortical areas: Simulation of dynamic integration in the visual system. *Cerebral Cortex*, Vol. 2, pp. 310–335, 1992.

[4] A.K. Seth, J.L. McKinstry, G.M. Edelman, and J.L. Krichmar. Visual binding, reentry, and neuronal synchrony in a physically situated brain-based device. In *Proc. of the 3rd Intl. Workshop on Epigenetic Robotics*, pp. 177–178, 2003.

[5] K.F. MacDorman, K. Tatani, Y. Miyazaki, and M. Koeda. Proto-symbol emergence. In *Proc. of the Intl. Conf. on Intelligent Robot and Systems*, pp. 1619–1625, 2000.

[6] A. Sotytchev. Development and extension of the robot body schema. In *Proc. of the Intl. Workshop on Epigenetic Robotics*, pp. 179–180, 2003.

[7] M. Asada, K.-F. MacDorman, H. Ishiguro, and Y. Kuniyoshi. Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Robotics and Autonomous System*, Vol. 37, pp. 185–193, 2001.

[8] Y. Yoshikawa, H. Kawanishi, M. Asada, and K. Hosoda. Body scheme acquisition by cross modal map learning among tactile, image, and proprioceptive spaces. In *Proc. of Intl. Workshop on Epigenetic Robotics*, pp. 181–184, 2002.

[9] Y. Yoshikawa, K. Hosoda, and M. Asada. Does the invariance in multi-modalities represent the body scheme? - a case study with vision and proprioception -. In *Proc. of Intl. Symp. on Adaptive Motion of Animals and Machines*, 2003.

[10] J. W. Sparling, J. V. Tol, and N. C. Chescheir. Fetal and neonatal hand movement. *Physical Therapy*, Vol. 79, No. 1, pp. 24–39, 1999.

[11] P. Rochat and S.J. Hespos. Differential rooting response by neonates: Evidence for an early sense of self. *Early Development and Pareting*, Vol. 6, pp. 105–112, 1997.

[12] T. Nakamura and M. Asada. Motion sketch: Acquisition of visual motion guided behavior. In *Proc. of Intl. Joint Conf. on Artificial Intelligence*, pp. 126–132, 1995.

[13] G. Bremner. *Infancy*. Blackwell, 1994.

Object recognition, Adaptive Behavior and Learning in Brain-Based Devices

Jeffrey L. Krichmar, Douglas A. Nitz, Gerald M. Edelman
The Neurosciences Institute, 10640 John J. Hopkins Drive, San Diego, CA, USA
krichmar@nsi.edu, nitz@nsi.edu, edelman@nsi.edu

Abstract

An open question in neuroscience is how animals combine the various attributes of stimuli in their environments into coherent perceptual categories and how they discriminate among objects in a scene. Testing a theory of visual binding would require the simultaneous study of brain function at many levels of organization. Present day electrophysiology only allows the recording of at most hundreds of neurons while an animal is performing a behavioral task. Because of this limitation and the sheer complexity of the nervous system, computational modeling has become essential for investigating theories of brain function. Accordingly, our group has constructed a series of brain-based devices (BBDs); i.e. physical devices with simulated nervous systems that guide behavior, to serve as heuristic bases for testing theories of brain function. Unlike animal models, BBDs permit analysis of activity at all levels of its nervous system as the device behaves in its real environment. We present a possible solution to the binding problem based on synchronous activity across neuronal groups brought about by reentrant connectivity. We first show the sufficiency of this theory in a laboratory setting and then demonstrate that these principles can be transferred to a more real-world application: robots capable of playing a game of soccer with humans.

1. Introduction

Animals effortlessly combine the various attributes of visual stimuli to form coherent perceptual categories and to discriminate among multiple objects in a scene. Yet the visual brain is functionally segregated: Separate cortical regions are specialized to respond to features such as shape, color, and object motion, and no single region has superordinate control. This poses the so-called binding problem [1]: How do these functionally segregated regions coordinate their activities to link various features of individual objects while distinguishing among different objects? Most proposed mechanisms for solving the binding problem fall into one of two general classes: 1) binding through the influence of attentional processes, executive mechanisms, or superordinate maps [2, 3]. 2) binding through the selective synchronization of dynamically formed neuronal groups [4-6]. Advocates of neural synchrony suggest that binding is an automatic, dynamic, and pre-attentive process arising from low-level neural dynamics. For example, the linkage of neuronal groups by reentry, the recursive exchange of signals across multiple, parallel and reciprocal connections [7], can lead to selective synchronization [8-11]. Synchronization of activity among neuronal groups can form coherent circuits corresponding to perceptual categories [8]. A fundamental challenge for proponents of neural synchrony is to show how such emergent functional circuits contribute to an organism's adaptive behavior, especially in situations that require preferential behavior towards one object among many in a scene.

Elucidation of brain mechanisms underlying behavior, such as visual binding followed by discriminatory action, requires simultaneous measurements across multiple levels. The heuristic value of synthetic modeling using brain-based devices (BBDs), which are described here, is supported by the fact that such types of measurements are difficult to obtain and compare in living animals. Given the successful construction of BBDs [12-14], we observe their overall behavior while simultaneously recording the state of all components of their simulated nervous systems. Since our purpose is to test theories of real nervous systems in order to arrive at a better understanding of brain function, we base the BBD's organization on real neuroanatomy and physiology.

We argue that a BBD should be constrained by the following design principles: 1) The device needs to engage in a behavioral task. 2) The device's behavior must be controlled by a simulated nervous system having a design that reflects the brain's architecture and dynamics. 3) The device needs to be situated in the real-world [15, 16]. 4) The behavior of the device and the activity of its simulated nervous system must allow comparisons with empirical data. Because of these constraints, BBD simulations tend to require large-scale networks of neuronal elements that reflect vertebrate brain architecture and dynamics, high performance computing to run the network in real-time, and the engineering of specialized physical devices to embody the network.

BBDs are not programmed by instructions like computers, but instead, like biological systems, they operate according to selectional principles that allow them to adapt to the environment [7]. Their design, which possesses neuroanatomical structure and large-scale neural dynamics, differs fundamentally from that of robots. Robotic approaches using classical artificial intelligence are based on data representation, rule-driven algorithms, and the manipulation of formal symbol systems.

BBDs must have a morphology or body plan that allows for active exploration in a real environment with a brain simulation controlling the BBD's behavior. Changes in the nervous system that result in lasting modifications of the device's behavior are realized through a neuromodulatory value system that signals the salience of environmental cues triggering broad changes to the BBD's nervous system. These features yield a system that generalizes signals from the environment into perceptual categories and adapts its behavior so that it becomes increasingly successful in coping with its environment. The BBDs have been designated the Darwin series of automata. Over the last 12 years, various Darwin automata have been shown to develop perceptual categorization, invariant visual object recognition, integration of scenes containing multiple visual shapes with overlapping features, fusion of different sensory modalities, and learning in the form of operant conditioning [12-14, 17, 18].

In this paper we describe two recent BBDs that address the problem of visual binding, scene segmentation, and motor behavior. First, we describe Darwin VIII, a BBD that demonstrated visual binding through synchronous activity across cortical areas brought about by reentrant signalling. Second, we apply the principles of Darwin VIII to a novel platform, the Segway Robotic Mobility Platform (RMP), in a dynamic environment, namely a soccer game.

2. Visual binding in a laboratory setting.

Darwin VIII, a BBD incorporating an extensive visual cortical neural simulation, demonstrated the ability to parse a scene composed of ambiguous visual shapes into separate and coherent perceptual categories. It solved the so-called binding problem, that is, it linked responses in different brain areas and modalities to yield selective responses to percepts in the absence of any superordinate control from a master or executive brain area [1]. The behavior of Darwin VIII exploited interaction between neural areas, and revealed that reentrant activity (i.e. ongoing reciprocal excitatory activity brought about by connections between neuronal units in different neural areas) is sufficient for recognizing and distinguishing among multiple objects in a scene.

Darwin VIII was designed to demonstrate visual categorization and selective conditioning in a rich environment [18]. This BBD had a camera for vision, microphones to pick up auditory cues from the environment, and infrared (IR) sensors to detect the boundaries of the environment.

2.1. Experimental Setup

Figure 1 shows Darwin VIII in its environment, which consists of an enclosed area with various shapes hung on two walls. Near the walls with visual shapes, infrared beams were set up that control speakers (see Figure 1A). When Darwin VIII's movement broke an IR beam, a tone was emitted. Darwin VIII reflexively oriented towards the sound source and gradually came to associate the sound with the object it saw near the sound source. After conditioning, the sound is no longer necessary; Darwin VIII approaches visual objects that have become associated with preferred sounds.

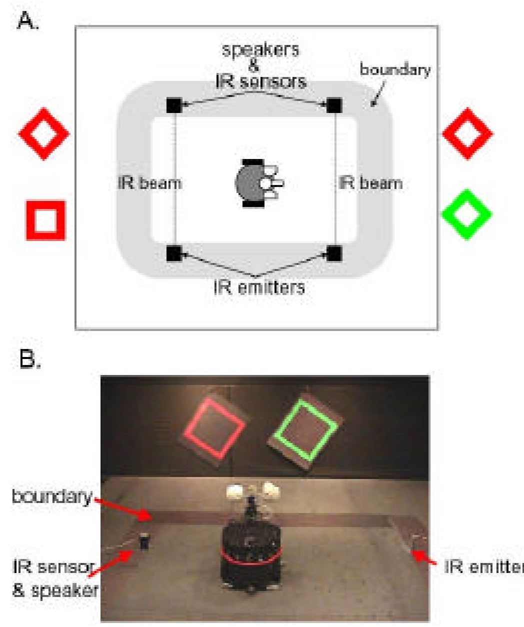


Figure 1. Experimental setup for Darwin VIII. A. Darwin VIII views objects on two of the walls of an arena. The area Darwin VIII explores is constrained by a boundary of reflective construction paper. Detection of this boundary by infrared sensor triggers a reflexive turn. When Darwin VIII breaks the beam from the IR emitter to the IR sensor, a tone is emitted from the speaker. B. Photograph of the experimental environment.

2.2. Simulated Nervous System

Darwin VIII's simulated nervous system contains areas corresponding to cortical and sub-cortical areas in the vertebrate nervous system (see Figure 2). Specifically, Darwin VIII's brain includes simulated cortical areas of the visual system that respond to shape and color (*V1*→*V2*→*V4*→*IT*), a motor system (*C*), an auditory system (*Mic-left*→*Aleft*, *Mic-right*→*Aright*), and a value system (*S*). Activity in *S* is analogous to that of ascending neuromodulatory systems in that it is triggered by salient events, influences large regions of the simulated nervous system, and persists for several cycles [19]. Due to its projection to the tracking area *C*, area *S* has a direct influence on behavior.

Neuronal units in Darwin VIII roughly correspond to the activity of 100 real neurons over 100 ms. The neuronal units have a firing phase parameter, which specifies the relative timing of this activity within each simulation cycle (for details, see [8, 18]). This modeling feature provides temporal specificity without incurring the computational costs associated with modeling spiking neurons in real time. Simulated synaptic connections follow known vertebrate neuroanatomical projections (arrows in Figure 2) and include extensive reentrant connectivity within and among neural areas. In Darwin VIII, reentrant connections among neuronal units encourage phase coherence and therefore lead to the emergence of neural synchrony.

Synaptic strengths are subject to modification according to a synaptic rule that depends on the phase and activities of the pre- and postsynaptic neuronal units. Plastic synaptic connections are either value-independent (see *IT*→*IT* in Figure 2) or value-dependent (see *IT*→*S*, *IT*→*C* in Figure 2). Both of these rules are based on a modified BCM learning rule [20] in which thresholds defining the regions of depression and potentiation are a function of the phase difference between the presynaptic and postsynaptic neuronal units (for details, see [18]). Synapses between neuronal units with strongly correlated firing phases are potentiated and synapses between neuronal units with weakly correlated phases are depressed; the magnitude of change is determined as well by pre- and postsynaptic activities. This learning rule is similar to a spike-time dependent plasticity rule [21] applied to jittered spike trains where the region of potentiation has a high peak and a thin tail, and the region of depression has a comparatively small peak and fat tail [22].

Value-dependent synaptic plasticity differs from the value-independent rule in that an additional term, based on the activity and phase of the value system, modulates the synaptic strength changes. Synaptic connections terminating on neuronal units that are in phase with the value system are potentiated, and connections terminating on units out of phase with the value system are depressed.

2.3. Image Processing and Computation

The CCD camera sent 320x240 pixel RGB video images, via an RF transmitter, to a frame grabber attached to one of the workstations running the neural simulation. The image was spatially averaged to an 80x60 pixel image. Gabor filters were used to detect edges of different orientations (45, 90, 135, 180 degrees). The output of the Gabor function mapped directly onto the neuronal units of the corresponding *V1* sub-area. Color filters (red positive center with a green negative surround, red negative center with a green positive surround) were applied to the image. The outputs of the color filters were mapped directly onto the neuronal units of *V1-Red* and *V1-Green*. *V1* neuronal units projected retinotopically to neuronal units in *V2* (see Figure 2).

Computation in the Darwin VIII simulation was carried out on a Beowulf cluster with 12 1.4 GHz Pentium IV processors using MPI. A simulation cycle, in which all the neuronal units and plastic synaptic connections were updated, took approximately 100 ms.

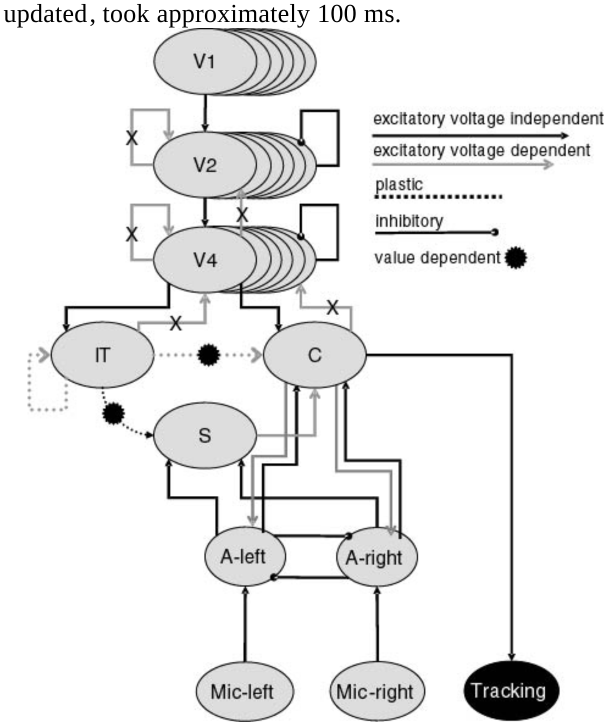


Figure 2. Global schematic of the regional and functional neuroanatomy of Darwin VIII. In the version used in the present experiments, the simulated nervous system contained 28 neuronal areas, 53,450 neuronal units, and approximately 1.7 million synaptic connections. The gray ellipses denote different neural areas. Arrows between the areas denote projections from one area to another. Projections marked with an 'X' are removed during lesion experiments. Tracking commands were issued to NOMAD's wheels based on activity in area *C*.

When the BBD triggers a speaker as it approaches a visual object, the tone emitted by the speaker activates its value system. At this time all of the value-dependent connections between neural areas (see value-dependent projections in Figure 2) are subject to value-dependent modification. Specifically, the changes dictated by the BCM synaptic change rule are further modulated by the average activity of the value system (area *S* in Figure 2).

As a consequence of these anatomical and dynamical characteristics, Darwin VIII autonomously approaches and views multiple visual shapes containing overlapping features (e.g. red squares, red diamonds, green squares and green diamonds) and can be trained to prefer one of these shapes by associating that shape with a positive-value tone (see Figure 1). It demonstrates this preference by orienting toward the preferred object.

2.4. Experimental Results

When confronted by a pair of these shapes, Darwin VIII learns successfully to track towards the preferred object, designated the target, and to avoid the other objects, designated the distracters. At first, this orientation is in response to the tone, but after approximately 10-15 minutes of viewing pairs of objects, the visual pattern alone is enough to elicit this preference.

All subjects successfully track the four different targets over 80% of the time (Figure 3A). Successful performance on this task is not trivial. Targets and distracters appear in the visual field at many different scales and at many different positions as Darwin VIII explores its environment. Moreover, because of shared properties, targets cannot be reliably distinguished from distracters on the basis of color or shape alone. Thus, the behavior of Darwin VIII demonstrates visual categorization and selective conditioning in a rich visual environment.

To investigate the importance of the presence of reentrant connections in the model, certain interareal reentrant connections were lesioned at different stages of the experimental paradigm. In one case, previously trained subjects were retested after lesioning. In a second, reentrant connections were lesioned in both training and testing stages. Lesions were applied to a subset of interareal excitatory reentrant connections (projections marked with an 'X' in Figure 2), which had the effect of transforming the simulated nervous system into a 'feed-forward' model of visual processing. To compensate for the reduction in activity due to these lesions, neuronal unit outputs in areas *V2* and *V4* were amplified. Figure 3B shows that subjects with intact reentrant connections performed significantly better than either lesioned group. The decrease in performance observed in the absence of reentry indicates that reentrant connections are essential for behavior above chance in the discrimination task.

These observations indicate that reentrant connectivity is necessary for the reliable discrimination of targets from visually similar distracters. In contrast to previous models of target selection, which required external intervention or an artificial environment [23, 24], Darwin VIII autonomously solved the binding problem in a rich environment even in the face of self-movement that generated changes in the size and location of visual stimuli.

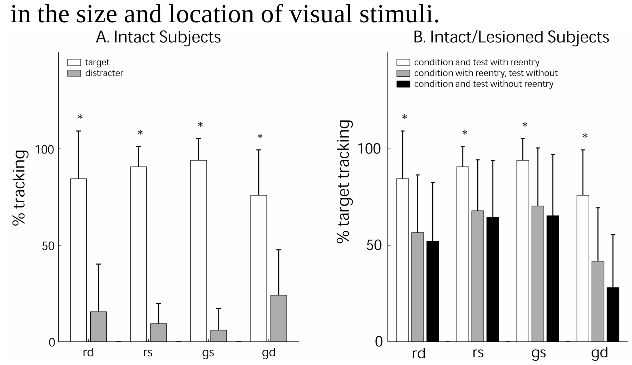


Figure 3. Darwin VIII behavior following conditioning. Three separate Darwin VIII subjects were conditioned to prefer one of 4 target shapes ('rd' = red diamond, 'rs' = red square, 'gs' = green square, 'gd' = green diamond). Activity in *V2* areas was used to assess the percentage of time for which NOMAD's visual field was centered on a particular visual shape. Bars in both graphs represent the mean percentage tracking time with error bars denoting the standard deviation. A. Darwin VIII subjects with intact reentrant connections tracked the targets (white bars) significantly more than the distracters (gray bars) for each target shape, averaging over all approaches (asterisks denote p < 0.01 using a paired sample nonparametric sign test). B. Comparison of intact subjects with lesioned subjects. White bars indicate target tracking performance of subjects with reentrant connections intact, light gray bars indicate subjects with lesions only during testing, and black bars indicate subjects with lesions during both training and testing. Intact subjects tracked significantly better than both lesion groups (Asterisks denote p < 0.01 using the Wilcoxon ranksum test).

During the behavior of an intact Darwin VIII subject, we observed circuits comprising synchronously active neuronal groups distributed throughout different areas in the simulated nervous system. Multiple objects were distinguishable by the differences in phase between the corresponding active circuits. A snapshot of Darwin VIII's neural responses is given in Figure 4, in which the device is approaching a red diamond target and a green diamond distracter towards the end of a training session. Each pixel in each neural area represents the activity (brightness) and phase (color) of a single neuronal unit. The figure shows two dynamic neural circuits differentiated by their distinct phases which were elicited respectively by the red diamond and the green diamond. As shown in the figure, Darwin VIII had not yet reached the beam that triggers the speaker to emit a tone. The activity of area *S* was nonetheless in phase with the activity in areas *V2* and *V4* corresponding to the target, and is therefore predictive of the target's saliency or value. Area *IT* has two patterns of activity, indicated by the two different phase colors, which reflect two perceptual categories. The increased activity in area *C* on the side of the target is causing Darwin VIII to orient towards the target (i.e. the red diamond).

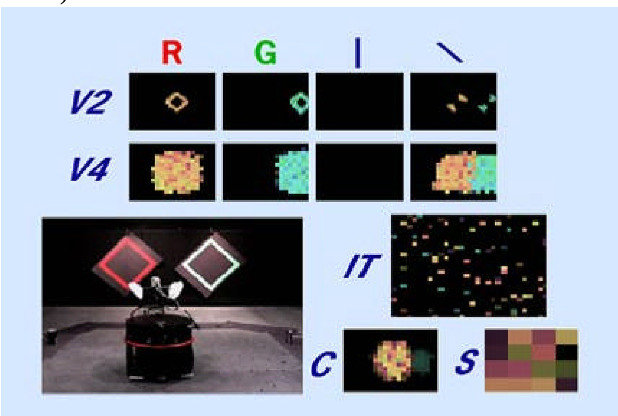


Figure 4. Snapshot of Darwin VIII's neuronal unit activity after approximately 10 minutes of conditioning. Darwin VIII is approaching a red diamond target (left) and a green diamond distracter (right) towards the end of a training session. Darwin VIII has not yet broken the beam that triggers the sound from the speakers located on the left side of the floor. The panels next to Darwin VIII show the activity and phase of selected neural areas (top row; *V2-red*, *V2-green*, *V2-vertical*, *V2-diagonal*, second row; *V4-red*, *V4-green*, *V4-vertical*, *V4-diagonal*, third row (to the right of Darwin VIII); *IT*, fourth row (to the right of Darwin VIII); *C* and *S*). Each pixel in the selected neural area represents a neuronal unit; the activity is normalized from no activity (black) to maximum activity (bright colors), and the phase is indicated by the color of the pixel (colors were chosen from a pseudocolor map, there is no connection between the color of the stimulus object and the color representing the phases of neuronal responses). The neuronal units responding to the attributes of the red diamond share a common phase (red-orange color), whereas the neuronal units responding to the green diamond share a different phase (blue-green color).

Object recognition and perceptual categorization were unsupervised. The simulated nervous system of a given subject developed distinct patterns of activity for each object it observed based on its own experience. Because images of the visual objects varied considerably in size and position as Darwin VIII explored its enclosure, successful discrimination required invariant object recognition. Darwin VIII's response to objects was position and scale invariant; it responded reliably to target images which appeared within ±20-degrees of the center field of view (the range of the visual field was approximately ±35-degrees) and as the apparent target size ranged from 8-degrees to 27-degrees of visual angle. This invariance was achieved due to generalization of a continuous stream of input due to self-movement.

The key mechanisms incorporated into Darwin VIII are reentrant connections within and among areas, neuronal units with a mean firing rate and a relative firing phase, and a value system modulating synaptic plasticity. The operation of these mechanisms, in conjunction with the sensorimotor correlations generated by self-motion, enable Darwin VIII to categorize visual objects, bind the features of visual objects, segment a scene, and demonstrate selective behavior in a rich real-world environment.

Our results are consistent with the hypothesis that visual binding results from the dynamic synchronization of neural activity mediated by reentrant connections among many dispersed neural areas. The performance of Darwin VIII suggests that specific timing relations and firing rates can act in a complementary mode to regulate behavior, and that synchrony among groups of neurons, as distinct from synchrony between pairs of individual neurons, may play a significant role in adaptive neural function.

3. Brain-Based Device Playing Soccer

Recently, we applied the visual binding and scene segmentation model of Darwin VIII to a BBD that can play soccer in both indoor and outdoor environments under varying lighting conditions and surfaces. The platform for this device is based on a modification of the Segway balancing technology and allows people on Segway Human Transporters (HT) to interact with Segway RMP robots (http://www.segway.com/segway/rmp/). The rules for this game are currently under development (see [25]) and a new league based on the Segways will be proposed for RoboCup 2005 (http://www.robocup.org). The rules will dictate that the BBD have the ability to both catch and kick a ball and also that that all sensing, actuating, and computing are local to the device (see Figure 5).

The neural simulation constructed for segmenting a soccer scene used the same principles of image processing, visual categorization, reentrant signalling and value-dependent learning as our previous Darwin automata. A high-level schematic of the neural architecture is given in Figure 6.

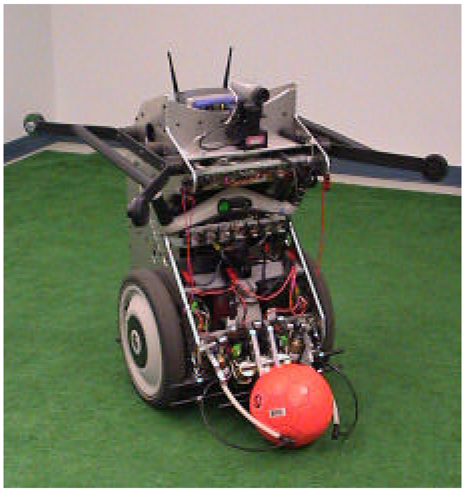


Figure 5. Brain-based soccer playing device based on the Segway RMP balancing platform. The BBD neural simulation receives sensory input from a CCD camera, IR sensors used for ball detection and obstacle avoidance, and odometry from the RMP. The simulation outputs to a camera pan-tilt unit, solenoids to capture the ball, solenoids to kick the ball, and motor commands to the RMP wheels.

Since many of the key elements of the soccer scene were pre-specified by the rules, certain neural areas were dedicated to responding to objects such as the ball, goal post, and teammate. Each of these areas had neuronal units that responded selectively to attributes of the object (e.g. *Goal* neuronal units were active when pre-synaptic neuronal units with similar receptive fields in areas *V2-Green*, *V2-Yellow,* and *V2-Horizontal* neuronal units were active). Activity in these object detection areas triggered motor actions (e.g. activity in Goal triggered kicking the soccer ball).

Because the mapping from visual ball recognition to the Segway RMP wheel motions was non-linear and complex, value-dependent plasticity was used to learn this mapping (see value-dependent learning projections in Figure 6). A form of temporal difference learning was developed in which value was increased when the number of active neuronal units in the *Ball* neural area increased and the activity of *Ball* neuronal units that respond to the center of device's field of view increased. This learning rule had the effects of potentiating *Ball*→*M* and *MEC*→*M* connections when a motor movement brought the ball closer to the device or of depressing connections due to erroneous movements away from the ball (see Figure 6).

Training the BBD to effectively track a soccer ball through value-dependent plasticity took approximately three minutes. In the first twenty seconds, movements were slow and the device's camera did not stay centered on the ball. By the final twenty seconds, ball tracking was fast and the device's camera stayed fixated on the ball.

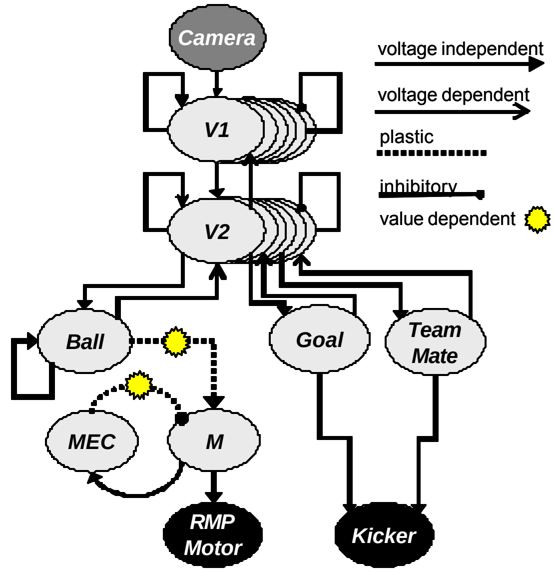


Figure 6. Schematic of the neural simulation architecture used for the Brain-based soccer playing device. Visual areas and their connectivity are similar to that of Darwin VIII. Neural areas *Ball*, *Goal*, and *TeamMate* respond specifically to those key objects on a soccer field and cause reflexive motor actions such as moving the Segway RMP or capturing and kicking the soccer ball. Ball tracking was achieved via plastic connections from the *Ball* and motor efference copy (*MEC*) areas to the motorneurons (*M*).

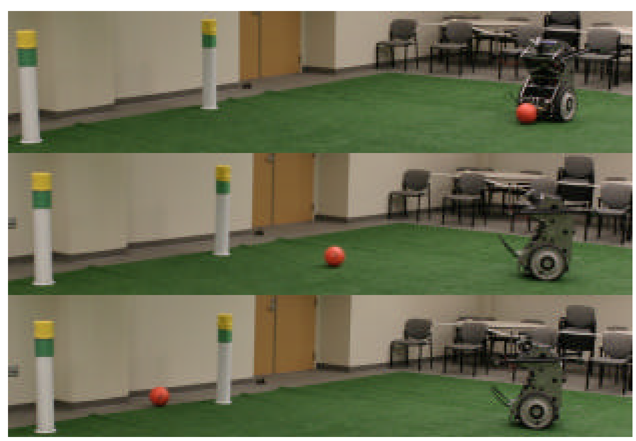


Figure 7. Goal shooting sequence. In the top panel, the BBD recognizes and acquires the soccer ball and then centers its gaze between the two goal posts. In the bottom two panels, the BBD kicks the ball between the goal posts.

Kicking to a teammate or the goal was achieved by the device recognizing the appropriate object, centering the object on its camera, and then kicking the ball (see Figure 7). Reentrant connections between neural areas facilitated dynamic synchronization, and color constancy [26]; allowing the BBD to recognize objects in a noisy environment under non-uniform lighting conditions (see Figure 8).

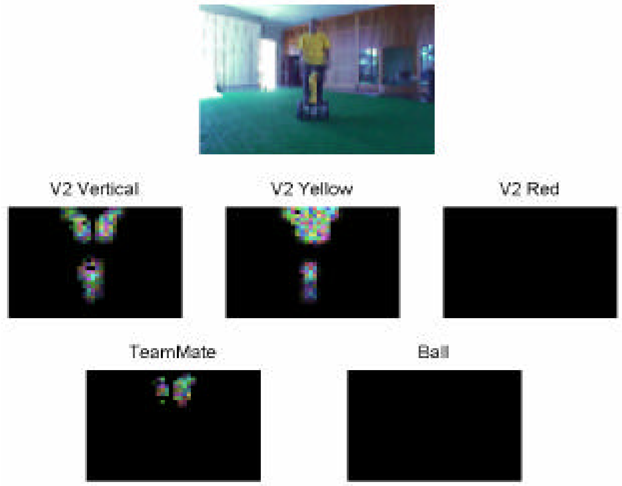


Figure 8. Snapshot of the BBD's camera input and selected areas in the neural simulation just prior to passing to a teammate. The camera input is sub-sampled to 60x80 pixels before processing. The activity and phase representation in the neural areas is the same as depicted in Figure 4.

The soccer playing BBD has successfully performed key elements of soccer playing; ball chasing, passing back and forth between itself and a human teammate on a Segway HT, and goal kicking. Video clips of the devices soccer playing capabilities can be found at: http://www.nsi.edu/nomad/segway.

4. DISCUSSION

Higher brain functions depend on the cooperative activity of an entire nervous system, reflecting its morphology, its dynamics, and its interaction with the phenotype and the environment. BBDs are designed to incorporate these attributes to allow tests of the self-sufficiency of such theories of brain function. We have demonstrated that BBDs can address many difficult tasks, without instruction or intervention, such as invariant object recognition [14], visual binding of objects in a scene [18], and texture discrimination using whiskers [17]. Like the brain, these BBDs operate according to selectional principles through which they form categorical memory, associate categories with innate value, and adapt to novel environments. These devices may provide the groundwork for the development of intelligent machines that follow neurobiological rather than computational principles in their construction.

We designed BBDs to simultaneously test parallel brain functions that could not presently be examined in any single animal in the laboratory. The BBDs were designed to yield data, in the form of neuronal activities in different brain regions that could be directly compared with experimental data. Equally important in the design is that a BBD must demonstrate adaptive behavior and this behavior must be measurable by an observer. The BBD's neural model, by necessity, is developed at a systems level, in which the structure of the brain and its different regions gives rise to adaptive behavior. Although the devices are still too simple to make direct comparisons to neurophysiology, they can make predictions about the neuroanatomical and dynamical constraints that subserve adaptive behavior.

Any model of brain function must not only take into consideration the structure of different brain regions, but must also pay attention to the connectivity within and between these brain areas. Brain function is more than the activity of disparate regions; it is the interaction between these areas that is crucial, as we have shown in Darwin VIII and the soccer playing BBD using the Segway. Thus, brains are defined by a distinct neuroanatomy in which there are areas of special function, which are defined by their connectivity to sensory input, motor output, and to each other.

Brains do not function in isolation; rather are tightly coupled to the organism's morphology, history, and environment. Therefore, our brain models are embodied in a physical device and explore a real as opposed to a simulated environment. The real environment is required for two reasons. First, simulating an environment can introduce unwanted and unintentional biases in a model. For example, a computer generated object presented to a vision model already has its shape and segmentation defined by the modeler and is directly presented to the model, whereas a device that views an object hanging on a wall has to discern the shape and figure from ground by segmentation based on its active vision. Second, real environments are rich, multimodal, and noisy. An artificial design of such an environment would be computationally intensive and difficult to simulate. All these interesting features of the environment come for "free" when we place the BBD in the real world. The modeler is freed from simulating an environment and can concentrate on the development of a device that can actively explore the real world.

The advantage of a synthetic model is that these measurements can be carried out in every neuron and synapse of the BBD's nervous system during the acquisition and recall of a behavior. To be effective, researchers using synthetic models need to analyze their data in such a way that they can compare their results to empirical data. By analyzing the neural dynamics of the model (i.e. spike rates, correlations between areas, neural dynamics and prediction), and choosing a behavioral paradigm similar to those used when studying behaving animals (i.e. mazes, conditioning paradigms, decision-making tasks, etc.), the modeler can directly compare the BBD's behavior with the results of psychological and neurophysiological experiments. This places the burden on modelers to include sufficient complexity in their models so that these psychological and physiological metrics can be compared.

We have obtained a number of insights and made several predictions based on the results of experiments with BBDs. In Darwin VIII, the model suggests that synchrony between widely separated neural areas may play a key role in solving the binding problem and demonstrates the importance of reentrant connections in facilitating binding through synchrony. The observed behavior demonstrates that binding through synchrony is feasible in an unlabeled real-world environment in which objects are constantly changing in size and position.

The development of adaptive and autonomous behavior by BBDs is novel in its neurally based approach and has implications for the construction of intelligent machines. The design and construction of such behaving devices based on principles of nervous systems may have much to offer to basic understanding and practical applications.

Acknowledgements. This work was supported by the Neurosciences Research Foundation. We thank J. Snook, D. Moore, and D. Hutson for their contribution to the design of the BBDs.

References

[1] A. Treisman, "The binding problem," *Curr Opin Neurobiol*, vol. 6, pp. 171-8, 1996.
[2] K. M. Shafritz, J. C. Gore, and R. Marois, "The role of the parietal cortex in feature binding.," *Proc Natl Acad Sci USA*, vol. 99, pp. 10917-10922, 2002.
[3] M. N. Shadlen and J. A. Movshon, "Synchrony unbound: a critical evaluation of the temporal binding hypothesis.," *Neuron*, vol. 24, pp. 67-77, 1999.
[4] G. M. Edelman, *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books, Inc., 1987.
[5] C. M. Gray, "The temporal correlation hypothesis of visual feature integration: still alive and well.," *Neuron*, vol. 24, pp. 31-47, 1999.
[6] W. Singer, "Neuronal synchrony: a versatile code for the definition of relations?," *Neuron*, vol. 24, pp. 49-65, 1999.
[7] G. M. Edelman, "Neural Darwinism: selection and reentrant signaling in higher brain function," *Neuron*, vol. 10, pp. 115-25., 1993.
[8] G. Tononi, O. Sporns, and G. M. Edelman, "Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system," *Cereb Cortex*, vol. 2, pp. 310-35, 1992.
[9] C. von der Malsburg and J. Buhmann, "Sensory segmentation with coupled neural oscillators," *Biol Cybern*, vol. 67, pp. 233-42., 1992.
[10] A. Knoblauch and G. Palm, "Scene segmentation by spike synchronization in reciprocally connected visual areas. II. Global assemblies and synchronization on larger space and time scales," *Biol Cybern*, vol. 87, pp. 168-84., 2002.
[11] A. Knoblauch and G. Palm, "Scene segmentation by spike synchronization in reciprocally connected visual areas. I. Local effects of cortical feedback," *Biol Cybern*, vol. 87, pp. 151-67., 2002.
[12] G. M. Edelman, G. N. Reeke, W. E. Gall, G. Tononi, D. Williams, and O. Sporns, "Synthetic neural modeling applied to a real-world artifact," *Proc Natl Acad Sci U S A*, vol. 89, pp. 7267-71, 1992.
[13] N. Almassy, G. M. Edelman, and O. Sporns, "Behavioral constraints in the development of neuronal properties: a cortical model embedded in a real-world device," *Cereb Cortex*, vol. 8, pp. 346-61, 1998.
[14] J. L. Krichmar and G. M. Edelman, "Machine psychology: autonomous behavior, perceptual categorization and conditioning in a brain-based device," *Cereb Cortex*, vol. 12, pp. 818-30., 2002.
[15] H. J. Chiel and R. D. Beer, "The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment," *Trends Neurosci*, vol. 20, pp. 553-7, 1997.
[16] A. Clark, *Being there. Putting brain, body, and world together again*. Cambridge, MA: MIT Press, 1997.
[17] A. K. Seth, J. L. McKinstry, G. M. Edelman, and J. L. Krichmar, "Spatiotemporal processing of whisker input supports texture discrimination by a brain-based device," in *Animals to Animats 8: Proceedings of the Eighth International Conference on the Simulation of Adaptive Behavior*, S. Schaal, A. Ijspeert, A. Billard, S. Vijayakumar, J. Hallam, and J. A. Meyer, Eds. Cambridge, MA: The MIT Press, 2004.
[18] A. K. Seth, J. L. McKinstry, G. M. Edelman, and J. L. Krichmar, "Visual binding through reentrant connectivity and dynamic synchronization in a brain-based device," *Cereb Cortex*, 2004, In press (http://www.nsi.edu/nomad/aks_jlk_cc2004_preprint.pdf)
[19] W. Schultz, P. Dayan, and P. R. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, pp. 1593-9., 1997.
[20] E. L. Bienenstock, L. N. Cooper, and P. W. Munro, "Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex," *J Neurosci*, vol. 2, pp. 32-48., 1982.
[21] W. Senn, H. Markram, and M. Tsodyks, "An algorithm for modifying neurotransmitter release probability based on pre- and postsynaptic spike timing," *Neural Comput*, vol. 13, pp. 35-67, 2001.
[22] E. M. Izhikevich and N. S. Desai, "Relating STDP to BCM," *Neural Comput*, vol. 15, pp. 1511-23, 2003.
[23] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts in attention.," *Vision Res*, vol. 40, 2000.
[24] B. A. Olshausen, C. H. Anderson, and D. C. Van Essen, "A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information.," *Journal of Neuroscience*, vol. 13, pp. 4700-4719, 1993.
[25] B. Browning, P. E. Rybski, J. Searock, and M. M. Veloso, "Development of a Soccer-Playing Dynamically-Balancing Mobile Robot," presented at IEEE International Conference on Robotics & Automation, New Orleans, 2004.
[26] J. Wray and G. M. Edelman, "A model of color vision based on cortical reentry," *Cereb Cortex*, vol. 6, pp. 701-716, 1996.

Modularity and Specialized Learning: Reexamining Behavior-Based Artificial Intelligence

Joanna J. Bryson
Artificial models of natural Intelligence
University of Bath, BA2 7AY, United Kingdom
J.J.Bryson@bath.ac.uk

Abstract

Learning, like any search, is only tractable if it is tightly focused. Modularity can provide the information a learning system needs by supporting specialized representation. Behavior-based artificial intelligence is a well-known modular theory of intelligent design, but has not been used systematically in this way. This paper describes a new design methodology, behavior-oriented design (BOD), which does. Examples drawn from mobile robotics and models of learning in non-human primates show the diversity of information this approach supports, from implicit and perceptual learning to tasks, maps and relationships.

1. Introduction

Behavior-based artificial intelligence (BBAI) is one of the best-known modular theories of intelligent design. Historically, however, although researchers have sometimes incorporated learning modules [e.g. 24, 28], there has been no systematic incorporation of learning into pure behavior-based design [though see 17]. Some hybrid systems have been developed which incorporate both BBAI and traditional planning and learning, but these lose the full advantages of modularity. Contemporary multi-agent systems (MAS) are fully modular, but overlook the BBAI advances in organizing distributed systems of complicated modules.

In this paper I describe how BBAI can be adapted to fully support modular learning. I begin by reviewing the history of BBAI. Then I discuss my own methodology, Behavior-Oriented Design (BOD), and explain how it exploits specialized learning, using examples from both robotics and ALife models of non-human primates. BOD allows a system with preprogrammed reactive control to behave in an adaptive manner, because its control relies on modules containing variable state. These modules combine current sensor readings with predictions based on learning.

2 Behavior-Based Artificial Intelligence

Behavior-Based Artificial Intelligence (BBAI) is a methodology for constructing intelligent agents which specifies that the attributes of their intelligence should be decomposed into semi-autonomous modules. The expressed behavior of these modules is made coherent through some system of arbitration between these modules. Both the arbitration system and the individual modules are intended to require relatively little processing power or time, so that the agent can respond quickly and appropriately to challenges and opportunities in complex dynamic environments.

When BBAI was introduced by Brooks [3, 5], its primary purpose was to provide a means to create these responsive (or *reactive*) agents. Creating such agents is difficult because a rich environment provides so many things to react to. Any living agent in a complex environment must choose between a large number of possible actions, where each action is itself dependent on a large number of environmental contingencies, and is motivated by competing, mutually exclusive goals.Choosing an optimal next action is impossible [15]. Even choosing a pretty good one requires searching an enormous space of possibilities.

Because an individual agent does not have time for such a search in real time, most of its decisions must be made in advance of its active life. However, this does not remove the complexity of the decision nor the amount of search necessary for a pretty-good choice. For animals, most of this search has been performed by evolution over a period of billions of years. For animats, the analogous role to evolution's is further split between the search conducted by the individual animat designer and that performed by the designer's culture. Designers must anticipate the behaviorally-salient contingencies that their agent may encounter, and provide rapid ways to recognize and select the appropriate response. We do this both through our own analysis and experimentation, and through exploiting the scaffolding of design knowledge we have previously learned.

BBAI is a piece of design knowledge that significantly advanced the state of agent design, particularly in the areas of mobile robotics and virtual reality. I believe that the primary reasons for this success are:

- the increased emphasis on providing engineered knowledge, which is a side effect of the emphasis on bottom-up control (sensing not representing), and
- the modular decomposition around individual expressed behaviors. This exploits the designers' existing skills and talents for writing simple programs.

After these significant advances, the complexity of the agents being built seems to have plateaued before the development of animal-level intelligence [22, 26]. Again, I believe there were two primary causes:

1. the fact that at least *some* expertise is best developed by the agent through experience, particularly of the local variations of its own physical plant ('body'), and its own local environment, and
2. the complexity of programming the behavior-arbitration systems increases exponentially as the complexity and number of behavior modules increases.

The first point is key to the thesis of this paper: modularity presents BBAI with the opportunity to maximally facilitate individual adaptation through providing specialized representations and processes. The second point, although important in the history of BBAI, is really a special case of the first. Modularizing the process of behavior arbitration and providing it with appropriate representations can greatly simplify the design process for a behavior-based agent.

3 A Brief History of Modular AI

This is a brief history of the critical attributes of BBAI systems which will support the claims I outlined above. More extensive reviews of the BBAI literature are available [6, 8], as are more thorough comparisons to neural and psychological theories [8, 10].

I will begin with Fodor's "The Modularity of Mind" [19], both because it introduces many of the concepts familiar to BBAI, and because it presents a theory of intelligence decomposition which is still actively researched in the natural sciences today [e.g. 14].

Fodor introduces the terms 'horizontal' vs. 'vertical' to describe two different sorts of decomposition of intelligence. *Horizontal* decompositions for Fodor are those which identify processes (e.g. memory, attention, perception, judgment) which underlie all of cognition. *Vertical* decompositions identify particular skills or faculties (e.g. mathematics, language, metaphysics) which each have their own characteristic processes of memory, attention and so forth. Roughly speaking, evidence for horizontal decomposition is the extent to which performance across domains is correlated for a particular individual; evidence for vertical decomposition is the extent to which it is not.

Fodor believes that *part* of human intelligence is decomposed in this vertical sense; those parts being perception and (separately) action. In Fodor's system, a number of semi-autonomous perceptual modules run simultaneously giving quick, automatic analysis of the perceptual scene. Each module recognizes its own best input, and effectively trumps the other modules when it is best utilized. The output of these modules is in the language of thought, which is operated on by a horizontal reasoning system. He presumes that the reasoning system's output is interpreted into action in a similar way, but theorizes less about this process.

Another precursor of BBAI is the "Society of Mind" [18, 29]. Minsky's proposal is more substantially vertical than Fodor's, although it still has some horizontal elements. An individual's actions are determined by simpler agencies, which are effectively specialists in particular domains. Minsky's agencies exploit hierarchy for organization, so for example the agency of play is composed of agencies such as block-play and doll-play. Arbitration between agencies is also hierarchical, so the play agency competes with the food agency for the individual's attention. Once play establishes control, the block and doll agencies compete.

Minsky's agents have both perception and action, but not memory, which is managed by another network of agencies of a different sort. Memory (K) agencies are interconnected both with each other and with the other, actor (S) agents; each can activate the other. Keeping the whole system working requires another horizontal faculty: the B brain, which monitors the main (A) brain for internally obvious problems such as redundancy or feedback cycles.

The term 'behavior-based artificial intelligence' was invented to describe a simplified but fully-implemented system used to control multiple, robotic agents. This was the subsumption architecture [3, 5]. The subsumption architecture is purely vertical. The modules were originally finite state machines, and arbitration between them was conducted exclusively by wires connecting the modules — originally literally, eventually as encoded in software. Each wire could connect one module to another's input or output wires, the signal of which the first module could then either monitor, suppress or overwrite.

Brooks initially asserted that most apparent horizontal faculties (e.g. memory, judgment, attention, reasoning) were actually abstractions emergent from an agent's expressed behavior, but had no place in the agent's actual control [5, p. 146–147]. However, his system was rapidly extended to have learning systems either inside modules or local to layers of modules [e.g. 4, 28]. Unfortunately, this promising approach was apparently smothered by the attractive simplicity and radicalism of his deemphasis on representation and centralized control.

Of the researchers who did *not* immediately adopt "no representation" as a mantra, most attributed the impressive success of Brooks approach to the fact that he had created abstracted primitives — the action/perception modules. Because these primitive units could sort out many of the details of a problem themselves, they made the composition of intelligence under *any* approach relatively easy [27]. Thus behavior systems were incorporated as a component into a large variety of AI architectures which still maintained centralized, logic-based planning and learning systems [e.g. 2, 21]. Unfortunately, the difficulty of reasoning about relatively autonomous components motivates the trivialization of behavior modules, e.g. to "fuzzy rules" [25] or vector spaces [1] which can be easily composed. Despite the lack of commonality of such approaches to Brooks' original ideal, they are still often called either behavior-based or hybrid behavior-based systems. Further, by the late nineties, the work of these researchers had so far outstripped that of the 'pure' BBAI researchers that two significant publications declared this hybrid approach to have been demonstrated superior to non-hybrid ones [22, 26].

Given the attributes of BBAI outlined earlier, in some senses multi-agent systems (MAS) are closer to BBAI than hybrid behavior-based systems. Each agent performs a particular task, and may have its own private knowledge store and representations which are presumably well suited to its function. However, to date there are fundamental differences between a MAS and a single, modular agent. These differences can be seen in the emphasis on communication and negotiation between modules / agents [35]. The MAS community is concerned with interoperability between unspecified numbers and types of agents, and with distribution across multiple platforms. This creates an administrative overhead not necessary for a single, modular agent[1].

In summary, BBAI was originally conceived and implemented as a clean, simple version of modular hypotheses that were already influential in psychology and AI. It lead to substantial improvements in real-time AI, and still has a great deal of influence not only in robotics [1, 26] but in virtual reality [33]. However, it is famously difficult to program [33, 35]. This difficulty has supported the widespread acceptance of hybridization between behavior-based and traditional AI. Unfortunately, these hybrids lose many of the advantages of modularity. The next section suggests ways to reclaim these advantages.

[1]Where MAS are in fact limited to a single platform and a relatively fixed architecture, I suspect their engineers should consider them to be modular single agents [9].

4 Modularity and Learning

In the previous section I explained Fodor's use of the terms "horizontal" and "vertical" to describe modular decompositions along generic function vs. task specific lines (respectively.) I also showed that the original behavior-based AI, the subsumption architecture, used the most strictly vertical modular decomposition. In this section I describe my own approach to BBAI and modular decomposition — that is, the problem of deciding how many modules an agent needs and what should be their capacities.

I believe modular decomposition should be determined by the requirements of variable state needed for learning. This idea is not entirely original; it is inspired by object-oriented design [e.g. 16, 30]. Consequently, I call it Behavior-Oriented Design (BOD). Under BOD, modular decomposition is done along the lines of specialized representations underlying adaptive requirements for the agent to be implemented. Most of these representations will support vertical abilities, for example representations underlying navigation or language, but some of them reliably support horizontal abilities, such as behavior arbitration or smoothing motor control.

Although this suggestion is simple, I think it brings a great deal both to BBAI and to the understanding of learning in intelligent systems, including animals. Compared to the original BBAI, BOD provides for learning while simplifying behavior arbitration. Compared to hybrid BBAI, BOD provides both a return to full modularity and a reemphasis on facilitating hand design.

In terms of understanding learning in intelligent systems, BOD makes explicit the continuum of adaptivity underlying intelligent behavior. The BOD development process [see 8] emphasizes two things:

- increasing the probability of success in learning (or any other type of search) by providing the agent with as much information (bias) as possible, and
- maintaining the simplicity of the agent by trading off complexity between various representations.

4.1 A Module for Behavior Arbitration

BOD particularly emphasizes the tradeoffs to be made between adaptive state for specialized perception and that for action selection through behavior arbitration [8]. This goes back to the notion of whether a module can, on its own, recognize a situation in which it should operate. I believe it is more reasonable for a module to recognize when it *can* operate. To recognize when it *should* operate requires more information than a largely encapsulated, semi-autonomous module ought to have access to.

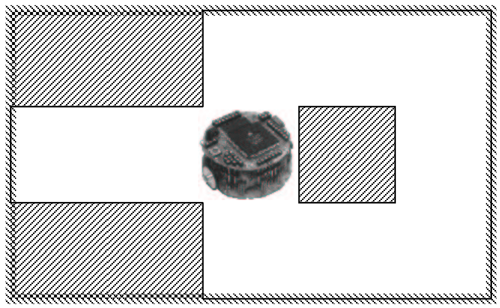


Figure 1. A patrolling robot cannot base its steering decisions entirely on external context and cover the entire maze.

In any particular context, there may well be more than one module that could or even should operate. This is the familiar problem of *perceptual aliasing*, which was originally seen as a problem of perception, but is in fact just a characteristic of the world. For example, consider a watch-robot intended to patrol an office space composed of corridors and junctions (Figure 1). For some junctions, the direction to go is entirely determined by either the robot's history (where it has most recently been) or its intentions (where it needs to go next.) We could try to read the robot's history or intentions off of its physical states (such as the direction it is pointing) but these can be perturbed by other subtasks such as avoiding people in the hallway. It is far simpler to keep a brief history of decisions or intentions in the specialized state that supports arbitration.

The strategy of making behavior arbitration into a special, horizontal module allows for a tradeoff between the complexity of action selection and the complexity of perception. I have argued at length elsewhere that ideally there should be a structured hierarchical representation underlying behavior arbitration, which represents behavior ordering and prioritization given a particular context [6, 7]. The advantage of such a decomposition is that it simplifies knowledge acquisition by separating acquisition tasks that have minimal correlation between them. The behavior-arbitration module doesn't need to know how task modules recognize context or perform their tasks; task modules don't need to know what other tasks might be performed in the same location at the same time, or what their relative priorities are.

4.2 From Perception to Knowledge

I will use the domain of mobile-robot navigation in order to demonstrate the variety of adaptation usefully modeled in behaviors in the BOD system. Although the robot work described here is old [12], the problems of robot perception and action provide clear and intuitive explanations for the different requirements for variable state. The robot, a radially symmetric, 16 sided Nomad 200, navigated in a smooth, continuous fashion around an office environment, negotiating doorways barely wider than itself and avoiding obstacles. It also learned simple paths from instruction.

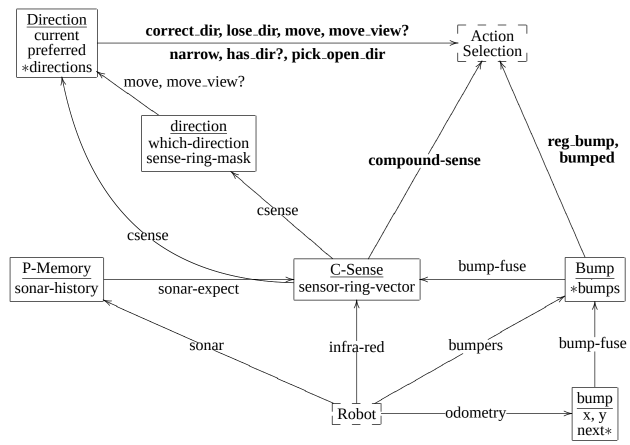


Figure 2. Behaviors for moving a robot.

Figure 2 shows behaviors that allow the robot to choose its speed and precise direction given that it has already determined an approximate goal heading. The vertical modules have solid boxes, the horizontal ones (including the robot's body) are dashed. Beginning at the bottom of the figure, the robot provides four types of sensory information relevant to picking a safe path. A **direction** behavior will determine the speed and direction for the robot, based on a 16 value array representing the approximate distance from each of the robot's faces to the next obstacle. This array is maintained by **C-sense** (compound sense).

Sonar, infra-red and bumpers all give information about the location of obstacles. Sonar operates by emitting sound then listening for it to bounce off obstacles. It can be accurate from about 20cm to 6m, but is subject to a variety of deflections and interference which can make objects appear suddenly closer or further away. Perceptual memory, **P-Memory** processes this information with a simple 6 item memory buffer. Each time a new sonar reading is received (about 7 times a second) the reading for each sensor is compared with those of the previous half minute. If a major discontinuity is perceived in one reading, it is ignored, and a new one computed based on the previous average value. However, if the new reading persists for 2 more readings, it is then 'believed' and becomes the new value for that sonar.

Infra-red sensors do not have the non-linearities of sonar, but have a far more limited range (approximately 0-24cm), and are also influenced by the color of the reflected surface. Infra-red sensors must be used for delicate maneuvers such as passing through doorways which require obstacle detection within the blind zone of the sonars. However, some things will not be detected by either long-range system, and are instead detected by the robots bumpers. The **bump** behaviors each represent one such past event. Since a bump *is* only detectable at the time and location of the event, the robot must compute the bump's approximate location after having disengaged from the obstacle in order to avoid it. This computation is based on odometric data. However, odometry accumulates errors rapidly, so bump events are forgotten after the robot has moved a few yards.

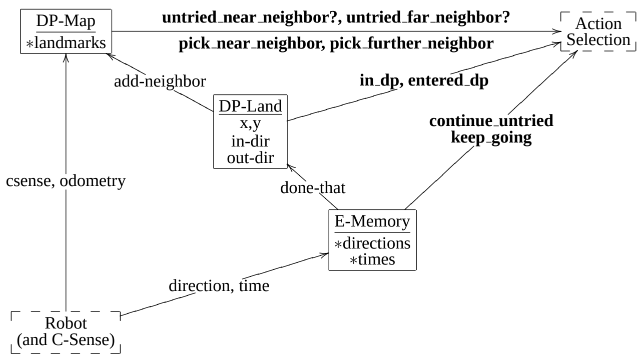


Figure 3. Behaviors added for map learning.

The robot thus brings a diverse array of "knowledge" to the continuous task of choosing a new speed and direction at any given instant. **Direction** and **Action Selection** work in concert for determining which **direction** controls these variables. **Direction** stores the current intended direction, while **Action Selection** determines the behavioral context (e.g. going forward normally toward a goal direction, or backing up after a collision). Each **direction** contains a template for determining discounts on the importance of the values of the array in **C-Sense** pertaining to whether the particular array value is directly in front, mostly to the side, or behind the direction of motion before that **direction**'s face. The value of the discount templates in the **direction** behaviors was learned off-line by the developer. The values in the **C-Sense** array are determined at any time, based on the most recent infra-read reading, the last half second of sonar readings, and perhaps a few minutes of bumper readings.

None of this adaptation would be considered "learning" in the common usage of the term, because it does not change state permanently for the lifetime of the agent. Nevertheless, all this knowledge may be considered predictions which lead to adaptive behavior. For example, the state recording the last direction of motion is used to predict the next one, which in turn determines what values are used in computing the robot's velocities. Similarly, the historic sonar readings are treated as more predictive of the true distance to obstacles than any one current sensor reading. The only reason to have adaptive state in the robot is because the past can be used to predict the present, and can do so more reliably than sensors on their own.

This argument extends to the modules that do map learning (see Figure 3, described further in [8, Section 7.6.3]). Here *decision points* — locations where the robot suddenly has a choice of direction (e.g. when it enters a room or encounters a doorway in a hall) are stored along with the decisions that were made at them, possibly after soliciting advice. Thus the robot can create a map (or learn a plan) from an instructor. This particular robot does not learn a complete, connected 2-D representation of the world, but rather a set of cues that can be read from the environment in order to make future decisions autonomously. Nevertheless, it behaves as if it has learned its way around. Now the common usage of 'learning' does apply, but the knowledge system is fundamentally the same.

5 Generic Types of Specialized State

The key observation about the robot example above is that BOD has been used to produce a reactive system which can operate well in a dynamic environment. It does this by exploiting a variety of types of information:

- *Engineering*, provided by the developer (or evolution), which does not change over the lifetime of the agent. This includes both fixed program code and parameters set by off-line tweaking and experimentation.

- *Reactive plans*, which keep track of the robots current decision context and focus its attention on particular behaviors. These are the basic representation underlying the Action Selection module.

- *Learned values of variable state*. Variable state is at the heart of the vertical / task modules. The 'learning' may persist only as very-short-term perceptual memory, as medium-term working memory, or for the lifetime of the agent.

This decomposition can also be found in real animals [10, for more details]. The engineered information is roughly equivalent to genetic predispositions, though in real animals, it is more difficult to separate development from learning, since development has evolved to rely on ubiquitous features of the environment as an information source. Reactive plans play a similar role to the behavior of the vertebrate forebrain [31], which, when working correctly, selects, sequences and inhibits behavior expression [13], though again note that in animals this can be more plastic than it is in BOD. Finally, the vertical behaviors I would equate with various sorts of cortical activation and plasticity. BOD does not currently discriminate between plasticity from activation levels and plasticity through long-term changes in connectivity.

These three types of information are not entirely disjoint: the reactive plans are hand coded, and are run in a special action-selection module. Reactive plans are themselves an elaborate form of specialized variable state. They encode both engineered information in the form of contingencies the designer anticipates the agent will encounter, and variable state indicating recent decision-making context, which constrains choices in the immediate future in order to provide persistence and reduce search.

In fact, all modules mix engineering with variable state. What makes the reactive plans special is that both their representation and the code that exploits it are used in all BOD agents. Extensive research has lead me to believe the BOD reactive plans are simply the best way to do behavior arbitration in a modular single agent [7, 8]. Obviously it would be useful to find other such generically useful representations, since reusing solutions reduces development time. In the rest of this section, I will discuss three other biologically-inspired types of learning or plasticity, two of which I am currently developing under BOD.

5.1 Drives and Emotions

Because the reactive plans underlying BOD action selection are relatively fixed, they do not represent well the sorts of variation that the brain represents chemically such as drives for food or sleep, or emotional states such as anger or fear. Drives and emotions represent and intermediate time course for intelligent state between the electrical / neural firing rate (which BOD represents in reactive plans) and long-term potentiation (which BOD stores in modules.) Reactive agents without this sort of state can seem erratic [33]. We are currently exploring how including this sort of durative decision state influences action selection, both from the perspective of believability (for VR agents) and for evolving social behavior (in Artificial Life agents.

The way to encode variable state in BOD is as behavior modules. However, these behaviors are so stereotyped, and have such simple state (essentially a single drive level) that they are effectively their own type. We are consequently developing a standardized representation for modeling of emotions and drives. Although the drive level itself is simple variable, each drive or emotion has its own onset and decay characteristics [34]. Further, the interactions between these states — with each other and with standard action selection — varies. For example, there must either be a latching or a blending mechanism to decide which of two conflicting drives or emotions is expressed.

To date we have used this type of behaviors both to create realistic real-time facial animation [34] and to create a simulation of a primate colony. The primate colony is the first exploration of combining all three types of intelligent state together. Its members have two drives: one for grooming (a stand in for general social behavior) and one for wandering alone (a stand in for foraging.) I have been using this model to explore the impact of adding simple social behaviors (such as tolerance of grooming) on the time spent by the group as a whole pursuing their goals [9]. We are currently extending the social model to include emotions or drives such as anger and affinity in an effort to model differences in different species of primates social structures.

5.2 Task Learning

The fact that BOD reactive plans are engineered bars BOD agents from doing something else real animals do: learn new tasks or new vertical modules. Again though, the extent to which animals have this capacity tends to be exaggerated in folk psychology. For example, pigeons can't learn to flap their wings for food or to peck to avoid shock, although they *can* learn to flap their wings to avoid shock or to peck to get food [see further 20, 32].

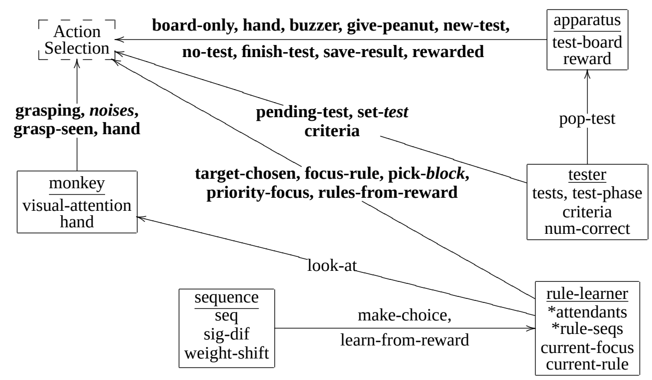


Figure 4. Behaviors used for an artificial life model of transitive inference learning.

I have built a model that learns what is effectively one component of a reactive plan within a particular context. The context is a model of transitive inference learning as performed by animals and children [8, 11]. The model shows simultaneous learning of both context / action pairs, and a set of prioritizations between the different contexts. These prioritizations determine when more than one context applies, which action should be taken. This amounts to a reactive plan — a prioritized set of context / action pairs.

To date we have demonstrated that this models both human and non-human primate learning of transitive inference. I am currently working to extend this model. True task learning should include not only context / action pairs and their priorities, but also when new contexts or actions need to be discriminated, and how this impacts the task representation as a whole. The performance context the agent believes itself to be in will determine the set of things it might learn as well as the things it might do.

This task-learning mechanism also has a biological correlate: the hippocampal learning system [11]. Nevertheless, such a general-purpose horizontal task-learning module should probably not become an expected component of all BOD agents. Such open-ended learning takes a great deal of time even with heavy bias, so defies the BOD principle of guaranteeing successful and timely learning. However, it is necessary for true mammalian intelligence.

5.3 Phenotype Swaps

Finally, I'd like to describe a very different form of natural plasticity. Hofmann and Fernald [23] have shown that both physical characteristics and expressed behavior can change extremely rapidly (within minutes) following a single traumatic (whether positive or negative) social event. The representations underlying these changes seem to be *phenotypic* in nature, with concurrent changes of gene expression in large numbers of neural synapses. The phenotypes in question determine whether a male Cichlid fish follows a behavior pattern of simple schooling, feeding and growth, or one of aggressive mating and territory defense which does not allow much time for growth. Male cichlid apparently alternate between these phenotypes. Not only behavior, but coloration change immediately after a decisive social event (a fight outcome), while gonad and overall size and shape gradually shift during the following weeks.

I have no immediate plans to model this sort of behavior, but it could be fairly easily done by implementing more than one action-selection plan hierarchy per agent, plus a special arbitration mechanism dedicated to swapping between these two plans. Since top-down expectations influence which behaviors are actively utilized by a BOD agent, this would effectively (though not actually) excite or inhibit other relevant behavior modules.

Is this learning? The representation involves no mental structure, and could not be used or manipulated in any other way. Yet an event (the result of a fight) selects a set of behavior which is only cost effective if that outcome serves to predict a reasonable period of success in near-future events of the same kind. The fish will only receive payoff for the hard work of defending a territory if it does so long enough to reproduce and protect its progeny. Again, adaptation is a continuum, and this perhaps the ultimate vertical module.

6 Conclusions

In this paper, I have described how modularity can be used to facilitate specialized learning, and shown how this is central to intelligent behavior. I have concentrated on how this perspective illuminates the history of modular AI, but I believe it also informs the modularity debate in psychology, and provides some explanation for the modularization we see in the brain. In the future I hope that modular AI will be able to do for psychology and systems neuroscience what neural networks research has done for neuroscience — provide testbeds and intuition pumps to help the natural sciences form and refine their models.

Acknowledgments

Thanks to Will Lowe, Marc Hauser (NSF Grant EIA-0132707), Martin Butz, Olivier Sigaud and Samarth Swarup. The robot in Figure 1 is a Khepera, ©K-Team SA Switzerland 2002. The Nomad robot work was in the University of Edinburgh Cognitive Neuroscience Laboratory.

References

[1] Arkin, R. C. (1998). *Behavior-Based Robotics*. MIT Press, Cambridge, MA.

[2] Bonasso, R. P., Firby, R. J., Gat, E., Kortenkamp, D., Miller, D. P., and Slack, M. G. (1997). Experiences with an architecture for intelligent, reactive agents. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2/3):237–256.

[3] Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2:14–23.

[4] Brooks, R. A. (1991a). Intelligence without reason. In *Proceedings of the 1991 International Joint Conference on Artificial Intelligence*, pages 569–595, Sydney.

[5] Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence*, 47:139–159.

[6] Bryson, J. J. (2000a). Cross-paradigm analysis of autonomous agent architecture. *Journal of Experimental and Theoretical Artificial Intelligence*, 12(2):165–190.

[7] Bryson, J. J. (2000b). Hierarchy and sequence vs. full parallelism in reactive action selection architectures. In *From Animals to Animats 6 (SAB00)*, pages 147–156, Cambridge, MA. MIT Press.

[8] Bryson, J. J. (2001). *Intelligence by Design: Principles of Modularity and Coordination for Engineering Complex Adaptive Agents*. PhD thesis, MIT, Department of EECS, Cambridge, MA. AI Technical Report 2001-003.

[9] Bryson, J. J. (2003). Where should complexity go? Cooperation in complex agents with minimal communication. In Truszkowski, W., Rouff, C., and Hinchey, M., editors, *Innovative Concepts for Agent-Based Systems*, pages 298–313. Springer.

[10] Bryson, J. J. (2004). Modular representations of cognitive phenomena in AI, psychology and neuroscience. In Davis, D., editor, *Visions of Mind*. Idea Group.

[11] Bryson, J. J. and Leong, J. C. S. (2004). Representations underlying transitive choice in humans and other primates. submitted.

[12] Bryson, J. J. and McGonigle, B. (1998). Agent architecture as object oriented design. In Singh, M. P., Rao, A. S., and Wooldridge, M. J., editors, *The 4^{th} International Workshop on Agent Theories, Architectures, and Languages (ATAL97)*, pages 15–30. Springer-Verlag.

[13] Carlson, N. R. (2000). *Physiology of Behavior*. Allyn and Bacon, Boston.

[14] Carruthers, P. (2003). The cognitive functions of language. *Brain and Behavioral Sciences*, 25(6).

[15] Chapman, D. (1987). Planning for conjunctive goals. *Artificial Intelligence*, 32:333–378.

[16] Coad, P., North, D., and Mayfield, M. (1997). *Object Models: Strategies, Patterns and Applications*. Prentice Hall, 2nd edition.

[17] Dahl, T. S. and Giraud-Carrier, C. (2001). PLANCS: Classes for programming adaptive behaviour based robots. In Cleeremans, A. and Lewicki, P., editors, *AISB'01 Symposium on Nonconscious Intelligence: From Natural to Artificial*.

[18] Doyle, J. (1983). A society of mind. Technical Report 127, CMU Department of Computer Science.

[19] Fodor, J. A. (1983). *The Modularity of Mind*. Bradford Books. MIT Press, Cambridge, MA.

[20] Gallistel, C., Brown, A. L., Carey, S., Gelman, R., and Keil, F. C. (1991). Lessons from animal learning for the study of cognitive development. In Carey, S. and Gelman, R., editors, *The Epigenesis of Mind*, pages 3–36. Lawrence Erlbaum, Hillsdale, NJ.

[21] Gat, E. (1991). *Reliable Goal-Directed Reactive Control of Autonomous Mobile Robots*. PhD thesis, Virginia Polytechnic Institute and State University.

[22] Hexmoor, H., Horswill, I., and Kortenkamp, D. (1997). Special issue: Software architectures for hardware agents. *Journal of Experimental and Theoretical Artificial Intelligence*, 9(2/3).

[23] Hofmann, H. A. and Fernald, R. D. (2000). Social status controls somatostatin-neuron size and growth. *Journal of Neuroscience*, 20:1248–1252.

[24] Horswill, I. D. (1993). *Specialization of Perceptual Processes*. PhD thesis, MIT, Department of EECS, Cambridge, MA.

[25] Konolige, K. and Myers, K. (1998). The Saphira architecture for autonomous mobile robots. In [26], pages 211–242. MIT Press, Cambridge, MA.

[26] Kortenkamp, D., Bonasso, R. P., and Murphy, R., editors (1998). *Artificial Intelligence and Mobile Robots: Case Studies of Successful Robot Systems*. MIT Press, Cambridge, MA.

[27] Malcolm, C., Smithers, T., and Hallam, J. (1989). An emerging paradigm in robot architecture. In *Proceedings of the International Conference on Intelligent Autonomous Systems (IAS)*, volume 2, pages 545–564, Amsterdam. Elsevier.

[28] Matarić, M. J. (1990). A distributed model for mobile robot environment-learning and navigation. Technical Report 1228, Massachusetts Institute of TechnologyArtificial Intelligence Lab, Cambridge, Massachusetts.

[29] Minsky, M. (1985). *The Society of Mind*. Simon and Schuster Inc., New York, NY.

[30] Parnas, D. L., Clements, P. C., and Weiss, D. M. (1985). The modular structure of complex systems. *IEEE Transactions on Software Engineering*, SE-11(3):259–266.

[31] Prescott, T. J., Redgrave, P., and Gurney, K. (1999). Layered control architectures in robots and vertebrates. *Adaptive Behavior*, (7):99–127.

[32] Roper, T. J. (1983). Learning as a biological phenomena. In Halliday, T. R. and Slater, P. J. B., editors, *Genes, Development and Learning*, volume 3 of *Animal Behaviour*, chapter 6, pages 178–212. Blackwell Scientific Publications, Oxford.

[33] Sengers, P. (1998). Do the thing right: An architecture for action expression. In Sycara, K. P. and Wooldridge, M., editors, *Proceedings of the 2^{nd} International Conference on Autonomous Agents*, pages 24–31. ACM Press.

[34] Tanguy, E. A. R., Willis, P. J., and Bryson, J. J. (2003). A layered Dynamic Emotion Representation for the creation of complex facial animation. In Rist, T., Aylett, R., Ballin, D., and Rickel, J., editors, *Intelligent Virtual Agents*, pages 101–105. Springer.

[35] Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152.