MIT Cog

Brian Scassellati

Contents

Abstract

The Cog Project was an exploration of the idea that human-like intelligence requires human-like interactions with the world [1]. Starting in the summer of 1993, Cog was one of the first humanoid projects in the United States and was a departure from many of the traditional methods promoted by artificial intelligence research at the time. Perhaps the most substantial impact of Cog was an invigorated interest in social and cognitive skills, a legacy that lead to the rapidly developing subfields of social robotics and human-robot interaction that are active today.

B. Scassellati (✉)
Computer Science & Mechanical Engineering & Materials Science, Yale University, New Haven, CT, USA
e-mail: scaz@cs.yale.edu

A. Goswami, P. Vadakkepat (eds.), *Humanoid Robotics: A Reference*,
https://doi.org/10.1007/978-94-007-6046-2_10

1 Introduction

The Cog Project began with a proposition by Rodney Brooks and Lynn Stein that intelligent behavior would be achieved by building initially simple and primitive behaviors and allowing the robot to learn through interactions with the world, much in the same way that human children learned complex skills by scaffolding successively more complex behavior through their interactions [2]. While most humanoid robots of the time focused on emulating the physical capabilities of adults (such as bipedal walking or musical performance), Cog followed a developmental approach to behavior and brought about a focus on social interaction, safe interaction, and developmental learning.

2 Robot Hardware

While the robot underwent some changes over the course of the 10-year project, this section describes the form that the robot maintained through most of the published work during the period 1995–2001. Throughout this period, Cog was an upper-torso robot, with two arms, a head, and a torso, but without legs, feet, or complex hands (Fig. 1). This reflected the focus on interactive and cognitive behavior rather than mobility or manipulation.

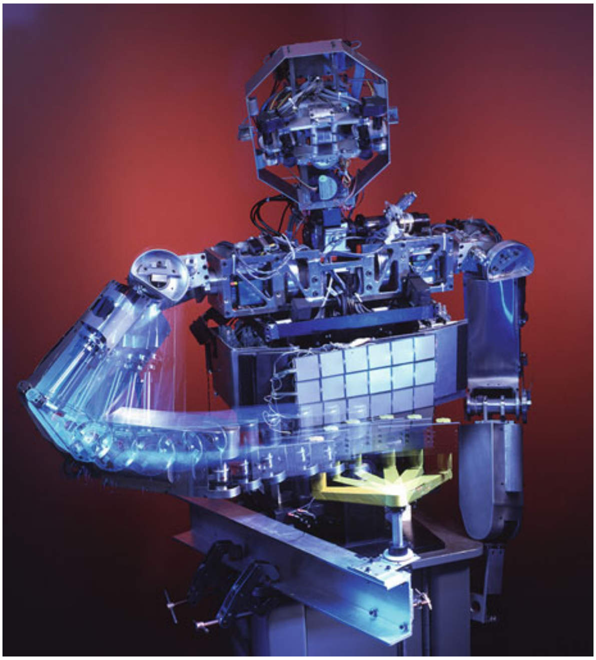

Fig. 1 Cog, an upper-torso humanoid robot, shown in multiple exposures while turning a crank (Photo by Sam Ogden)

Cog was the first humanoid to introduce a number of hardware conventions that are still present in humanoids today, including the use of compliant actuators for physically safe interactions [3] and the use of paired cameras for foveated vision [4].

2.1 Motor Systems

Cog had a total of 21 mechanical degrees-of-freedom (DOF); two six DOF arms, a torso with a two degree-of-freedom (DOF) waist, a one DOF torso twist, a three DOF neck, and three DOF in the eyes.

To allow for untrained users to interact safely with the robot, each of the joints in the robot's arms used a series elastic actuator, which connected a DC electric motor through a series spring to the load [3]. The spring provided torque feedback at each joint, protected the motor gearbox from shock loads, and allowed for torque-based control. The springlike property gave the arms a sensible "natural" behavior: if it were disturbed, or hit an obstacle, the arm simply deflected out of the way. The disturbance was absorbed by the compliant characteristics of the system and needed no explicit sensing or computation. The system also had a low frequency characteristic (large masses and soft springs), which allowed for smooth arm motion at a slower command rate.

2.2 Perceptual Systems

To obtain information about the environment, Cog had a variety of sensory systems that emulated human sensing, including visual, vestibular, auditory, and kinesthetic senses. Although range sensing through non-human-like modalities (LIDAR, SONAR, structured infrared lighting) was available at the time, Cog used only sensing systems that were arguably analogous to human capabilities.

Cog's visual system was designed to mimic some of the capabilities of the human visual system, including binocularity and space-variant sensing [4]. Each eye could rotate about an independent vertical axis (pan) and a coupled horizontal axis (tilt). To allow for both a wide field of view and high-resolution vision, there were two grayscale cameras per eye, one which captured a wide-angle view of the periphery (88.6°(V) × 115.8°(H) field of view) and one which captured a narrow-angle view of the central (foveal) area (18.4°(V) × 24.4°(H) field of view with the same resolution) (see Fig. 2).

To mimic the human vestibular system, Cog had three rate gyroscopes mounted on orthogonal axes (corresponding to the semicircular canals) and two linear accelerometers (corresponding to the otolith organs). Each of these devices was mounted in the head of the robot, slightly below eye level. Analog signals from each of these sensors were amplified on-board the robot and processed off-board by a commercial A/D converter.

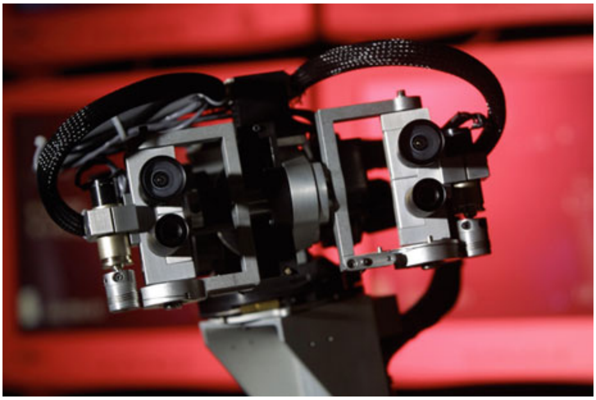

Fig. 2 Close-up of the robot's head. The vision system featured a joint tilt axis with independent pan axes for vergence. Each "eye" used two cameras, a wide-angle peripheral camera, and a narrow-angle foveal camera

To provide auditory information, two omnidirectional microphones were mounted on the head of the robot. To facilitate localization, crude pinnae were constructed around the microphones.

Feedback concerning the state of Cog's motor system was provided by a variety of sensors located at each joint. The eye axes utilized only the simplest form of feedback; each actuator had a single digital encoder, which gave position information. The neck and torso joints had encoders, as well as motor current sensing (for crude torque feedback), temperature sensors on the motors and driver chips, and limit switches at the extremes of joint movement. The arms joints had the most involved kinesthetic sensing. In addition to all the previous sensors, each of the 12 arm joints also had strain gauges for torque sensing and potentiometers for absolute position feedback.

2.3 Computational System

The computational system driving the robot underwent many revisions during the 10-year span of the project. As new microcontrollers and new computer architectures were introduced, changes were made to allow for faster and denser processing of the robot's sensory data. As the robot was in continuous development, there were often multiple generations of processing systems that maintained partial backward compatibility. During most of the early years, a heterogeneous network of different processors types operated at different levels in the control hierarchy, ranging from small microcontrollers for joint-level control to faster and more general processing systems for behavior selection and sensory processing.

The original design of the robot used a network of 16 MHz Motorola 68,332 microcontrollers on custom-built boards, connected through dual-port RAM. Each of these nodes ran L, a multithreading subset of Common Lisp. Each joint on the robot had a dedicated local motor controller, a custom-built board with a Motorola

HC11 microcontroller, which processed encoder and analog inputs, performed servo calculations, and drove the motor via pulse-width modulation. For the arms, the microcontroller generated a virtual spring behavior at 1 kHz, based on torque feedback from strain gauges in the joints.

Starting around 1997, the core network of Motorola microcontrollers was supplanted by a network of 200 MHz industrial PC computers running the QNX real-time operating system and connected by 100VG ethernet. The network initially contained four nodes but allowed for easy expansion via the network. QNX, commonly used at the time in embedded controllers for automobiles, provided transparent and fault-tolerant interprocess communication over the network. A custom-built shared memory ISA interface card allowed for communication between the QNX-based PC nodes and the original Motorola microcontrollers.

At times, the robot also utilized a separate video and audio preprocessing network. This network of Texas Instruments C40 digital signal processors communicated via the proprietary C40 communications port interface. The network included C40-based acquisition boards, display boards, and audio I/O ports. This subsystem relayed data to the core processor network via ISA and PCI interface cards.

3 Research Directions

The Cog robot was a platform that hosted a range of research efforts between 1994 and 2003, but the most notable work revolved around four central themes: building behaviors by following a developmental progression, reliance on social interaction to structure learning, leveraging physical embodiment to aid computation, and using robots to explore theories of human intelligence. (See Refs. [5, 6] for more consideration of these points.)

3.1 Building Behaviors Developmentally

While other contemporary projects focused on the deployment of adultlike competencies, the behaviors for Cog were constructed to follow a gradual developmental process similar to that of a human infant. Simple behaviors were constructed to mimic the primitive, innate abilities that the robot should possess with the expectation that more complex behavior would arise through the interactions of these behaviors with each other, with the environment, and through active learning processes. The hope of this approach was that adult-level competencies could be achieved by following this bottom-up constructivist approach to cognitive engineering.

Following this approach, research projects on Cog created flexible behaviors for visually guided pointing [7], for visual attention [8], and for primitive forms of theory of mind and social awareness [9]. All of these projects leveraged simple, easy-to-construct behaviors to enable the construction of higher-level behaviors that required fewer training examples, were simpler and less burdensome to create, or that had never been demonstrated before using any other approach. While these projects have had impact as proof-of-concept, the most lasting aspect of this approach was the application of bottom-up techniques from traditional behavior-based robotics to the more mechanically complex and behaviorally rich humanoid robots.

3.2 Social Interaction Structures Learning

The Cog Project also played a role in focusing research effort on the development of interactive, social behaviors for humanoid robots (Fig. 3). While many of the systems of that time concentrated on developing skills in isolation, the Cog Project brought a focus on developing simple, yet effective interactive skills that allowed the robot to engage with untrained bystanders. These behaviors often were subtle and not as visibly impressive as other robotic tasks (such as manipulation, walking, or playing musical instruments). However, the presence of these behaviors often encouraged people to treat the robot as an agent, with thoughts, desires, and intentions; this classification led to more pronounced instructional behavior, more natural interactions, and a readiness to engage in more frequent teaching behavior.

The interaction between social behavior and machine learning was explored under a variety of contexts during the Cog Project, including direct studies of imitation learning [10], explorations of how to enhance social feedback for learning systems [11], and how early social skills can shape the nature of problems that need to be learned [12]. While most of these studies demonstrated how computational simplification could be achieved by using interactive social skills, they often relied on relatively simple computational learning techniques. The integration of complex social behavior with state-of-the-art computational learning techniques would not be addressed until years after the project ended; however, these studies laid the foundation for how this integration could be beneficial.

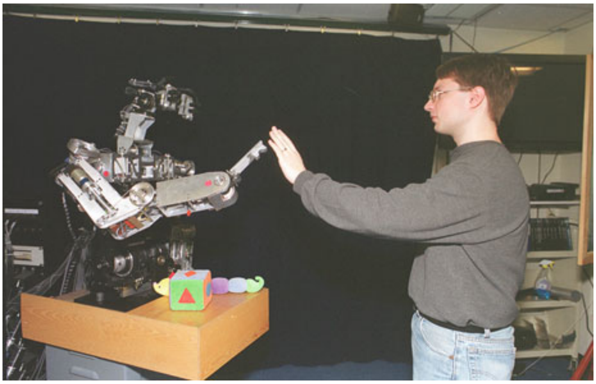

Fig. 3 Cog engaged in imitation behavior. Most of the robot's social skills were developed by mimicking models of infant development

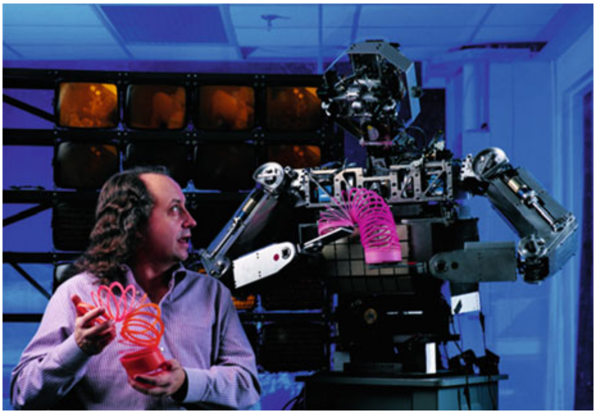

Fig. 4 Rodney Brooks and Cog. Rhythmic oscillations were linked to the physical embodiment of the robot rather than to complex control architectures. Photo by Peter Menzel

3.3 Physical Embodiment Aids Computation

One of the fundamental premises of the Cog Project was that computationally complex tasks could be simplified by the physical embodiment and presence of the robot [2] (see Fig. 4). This idea took two different forms during the progression of the project. In some cases, having the correct physical structure replaced a computationally expensive operation with a different, but simpler, computational problem. One of the most lasting impacts of this idea was the use of series elastic actuators to drive rhythmic control [13]. By replacing standard direct-drive position controlled actuators with a compliant, force-controlled actuator, complex repetitive maneuvers could be achieved with minimal computation [14].

In other cases, the physical presence and extent of the robot structured the kinds of input, the rate of presentation, or the complexity of the challenges being presented to the robot. For example, while the impact of a moving camera on machine vision techniques was studied by the active vision movement, the impact of moving cameras during social interactions introduced an additional set of constraints that simplified some of the computational vision problems that the robot needed to address [15]. Cog was also used to demonstrate that active exploration and manipulation of objects could simplify the computation needed for object identification and other vision processes [16].

3.4 Using Robots to Explore Human Intelligence

While many humanoid projects draw inspiration in their architecture and behavioral structure from what we know of human intelligence, the Cog Project also had a strong emphasis on using robots as a mechanism to further our understanding of human intelligence [17]. While direct evidence about biological mechanisms cannot be inferred from a computational model, this research furthered our exploration of biological mechanisms of human intelligence through three different mechanisms. First, the emphasis on building complete, embodied systems could uncover questions or aspects of a task that would need implementation but that were often ignored in both the computational and biological literature. For example, studies of imitation often ignore fundamental questions about what should be imitated, who should be imitated, and what constitutes success in an imitated behavior [18]. Second, observations on successful robot implementations can lead to new proposed models of biological intelligence. The behavior-driven, interactive architectures that developed from this project have generated a set of proposed models for cognitive science and biological intelligence [6, 19]. Finally, direct implementations of proposed models of biological intelligence and development on the robotic system can uncover unintended connections or oversights in the original model. For example, an implementation of early social development models of joint attention revealed connections between two distinct modules that were thought to be unrelated [20].

4 Legacy

The Cog Project began with the hypothesis that physical embodiment was a critical part of building complex intelligence. While proving this to be necessarily true for all instances of complex intelligence is beyond the reach of any project of this nature, Cog served as a proof-of-concept that physical embodiment had value when considering the construction of complex intelligent systems. The project highlighted research areas (such as social interaction and developmental learning) that had been largely ignored and at the same time established new conventions that are still popular today. In many ways, this project can be seen as a link between the insect-like "behavior based" robots of the 1980s and early 1990s and the humanoid robots focused on cognitive skills that proliferated in the late 1990s and 2000s.

Many of the central ideas introduced during the decade of work on Cog have inspired research disciplines outside of humanoid robotics. The focus on building behaviors following a developmental trajectory became the founding focus of the Autonomous Mental Development (AMD) and Developmental Robotics communities, and is well represented today by an IEEE journal and conference series (the IEEE Transactions on Cognitive and Developmental Systems and the joint IEEE International Conference Developmental Learning and Epigenetic Robotics). The work on building socially interactive systems was part of the impetus for the highly successful ACM/IEEE International Conference on Human-Robot Interaction and the International Conference on Social Robotics, along with their respective journals. The use of physical embodiment to aid or replace computation was highly influential in work on bipedal locomotion, including some work on passive dynamic walking systems. Finally, the use of robots as models for cognitive behavior has produced special sessions dedicated to robotic models at conferences in more human-centered disciplines (including the Annual Meeting of the Cognitive Science Society and the International Conference on Infant Studies).

Perhaps the most lasting impact of this project was the collection of researchers that were brought together to collaborate on this platform. This collection included both the graduate students that Rodney Brooks brought to the project (many of which have had influential positions in academia and industry) and a broad range of researchers and visitors from computer science, philosophy, cognitive science, and developmental psychology. Starting with six graduate students in the first 5 years of the project (Cynthia Breazeal, Robert Irie, Matthew Marjanovic, Yoky Matsuoka, Brian Scassellati, and Matthew Williamson) the project expanded in later years to include contributions from many other students (including Bryan Adams, Artur Arsenio, Lijin Aryananda, Jessica Banks, Aaron Edsinger, Paul Fitzpatrick, Charles Kemp, Eduardo Torres-Jara, Paulina Varchavskaia, and Juan Velasquez). The project was also shaped through the interactions with a large number of collaborators and visiting researchers included Dave Cliff, Dan Dennett, Hideki Kozima, Giorgio Metta, Lorenzo Natale, Una-May O'Reilly, Rolf Pfeiffer, Polly Pook, Takinori Shibata, Sajit Rao, Giulio Sandini, and Manuela Veloso.

References

1. R. Brooks, C. Breazeal, M. Marjanovic, B. Scassellati, M. Williamson, The cog project: Building a humanoid robot, in *Computation for Metaphors, Analogy and Agents, Vol. 1562 of Springer Lecture Notes in Artificial Intelligence*, ed by C. Nehaniv (Springer, New York, 1998), pp. 52–87
2. R. Brooks, L. Stein, *Building Brains for Bodies*, AI Lab Memo, vol 1439 (MIT, Cambridge, 1993)
3. G. Pratt & M. Williamson, Series Elastic Actuators, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-95),*Pittsburgh, vol 1 , 1995, pp. 399–406
4. B. Scassellati, *A Binocular, Foveated Active Vision System*. Technical Report 1628 (MIT, Cambridge, 1998)
5. B. Adams, C. Breazeal, R. Brooks, B. Scassellati, Humanoid robots: A new kind of tool. IEEE Intell. Syst. Appl. **15**(4), 25–31 (2000)
6. R. Brooks, C. Breazeal, R. Irie, C. C. Kemp, M. Marjanovic, B. Scassellati, M. Williamson, Alternate essences of intelligence. in *15th National Conference on Artificial Intelligence (AAAI-98)*, Madison, Aug 1998
7. M. Marjanovic, B. Scassellati, M. Williamson, Self-taught visually-guided pointing for a humanoid robot, in *4th International Conference on Simulation of Adaptive Behavior (SAB-96)*, Cape Cod, June 1996
8. B. Scassellati, Eye finding via face detection for a foveated, active vision system, in *15th National Conference on Artificial Intelligence (AAAI-98)*, Madison, Aug 1998
9. B. Scassellati, Theory of mind for a humanoid robot, in *1st IEEE/RSJ International Conference on Humanoid Robotics (Humanoids 2000)*, Cambridge, Sept 2000
10. B. Scassellati, Imitation and mechanisms of joint attention: A developmental structure for building social skills on a Humanoid Robot, in *Computation for Metaphors, Analogy and Agents*, *Vol. 1562 of Springer Lecture Notes in Artificial Intelligence*, ed. by C. Nehaniv (Springer-Verlag, Berlin, 1998), pp. 176–195
11. C. Breazeal, B. Scassellati, How to build robots that make friends and influence people, in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-99)*. Kyonjiu, Aug 1999

12. B. Scassellati, Investigating models of social development using a humanoid robot, in *Biorobotics*, ed. by B. Webb, T. Consi (MIT Press, Cambridge, 2000)
13. M.M. Williamson, Rhythmic Robot Arm Control Using Oscillators, in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS-98), Victoria, 1998
14. A. Arsenio, Tuning of neural oscillators or the design of rhythmic motions, in *International Conference on Robotics and Automation (ICRA 2000)*, (IEEE, San Francisco, 2000)
15. C. Breazeal, A. Edsinger, P. Fitzpatrick, B. Scassellati, P. Varchavskaia, Social constraints on animate vision. IEEE Intell. Syst. **15**(4), 32–37 (2000)
16. G. Metta and P. Fitzpatrick, Better Vision through Manipulation, in *Second International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems (EPIROB)*. Edinburgh, Aug 2002
17. B. Adams, C. Breazeal, R. Brooks, B. Scassellati, Humanoid robots: A new kind of tool. IEEE Intell. Syst. **15**(4), 25–31 (2000)
18. C. Breazeal, B. Scassellati, Challenges in building robots that imitate people, in *Imitation in Animals and Artifacts*, ed. by K. Dautenhahn, C. Nehaniv (MIT Press, Cambridge, 2000)
19. R. Brooks, From earwigs to humans. Robot. Auton. Syst. **20**(2–4), 291–304 (1997)
20. B. Scassellati Investigating models of social development using a humanoid robot, in *Biorobotics* ed. by B. Webb, T. Consi (MIT Press, Cambridge, 2001)