Modeling Learning as a Cooperative Multi Agent Interaction

Mathew Davies and Elizabeth Sklar
Department of Computer Science
Columbia University
1214 Amsterdam Avenue, Mailcode 0401
New York, NY 10027 USA
mdavies,sklar@cs.columbia.edu

Abstract. In the Meta-Game of Learning, fashioned after the Iterated Prisonner's Dilemma, Teacher and Student are agents that interact in a synchronized, iterated game. The Teacher can pose to the Student either HARD or EASY questions; the Student can respond with either RIGHT or WRONG answers. When the hard questions are answered correctly, the student is *learning* — the result of a *cooperative* action on the part of the two participating agents. In this paper we expand on the MGL, modelling the Student's behavior and attempting to capture some of the complexity that accounts for the student providing a response to a question. We consider *motivation, emotion* and *ability* as contributing factors and present results of preliminary experiments applying our model and varying each of these factors.

1 Introduction.

Based on the assumption that learning is fundamentally interactive, we are studying the types of interactions that occur in a traditional educational setting. In earlier work [3], we modeled the interactions between teachers and students and the manner in which these agents are rewarded as a *Meta-Game of Learning* or MGL [2], which is fashioned after the Iterated Prisonner's Dilemma (IPD) [1]. In the MGL, we consider the Teacher and Student as two agents in a synchronized version of the IPD. Each player can make one of two moves at each iteration. The Teacher goes first and can present to the Student either a HARD or an EASY question. The Student will respond with either the RIGHT or WRONG answer.

We define *learning* — the overarching goal of the education system — as occurring through a mix of each of the four possible states, with a primary drive towards the HARD question / RIGHT answer state. We describe these behaviors, that advance both agents towards the goal, as *Cooperation*, while behaviors that advance only one agent at the expense of the other is referred to as *Defection*. This reflects that notion that interactions which result in learning are fundamentally cooperative. We use the term *meta-game* to refer to the overarching system of interactions between multiple agents of each type over many iterations.

Our longterm direction with this work is to explore decision-making in pedagogical agents who choose between presenting students with HARD and EASY questions at various stages throughout a series of iterated interactions, with the overarching goal of the student *learning*. In the work presented here, we expand the MGL by focusing on modeling complexities in the student's behavior. We posit that the student's responses to questions are affected by a combination of several fundamental attributes:

- *Ability (A)* — we take ability to indicate how easily a student can learn a concept.
- *Motivation (M)* — in general, a high level of motivation is commensurate with cooperation (i.e., responding with the RIGHT answer when the student is trying) and the converse for defection (i.e., responding with the WRONG answer when the student is not trying).
- *Emotional (E)* — we model a student's emotional state, or relative happiness, contentment, etc., as a variable related to both ability and motivation.

Although in theory these attributes may be considered continuous (possibly vector-valued) functions over time, we attempt to capture their essential behavior with models employing discrete (binary or n-ary) values that change through the course of interaction according to simple rules. The specification of a model must determine how an agent's attribute values map to the agent's actual choice to cooperate or defect.

We have defined a set of simple rules to simulate each of these behaviors (cooperation and defection). The next two sections describe these rules. Section 4 shows some preliminary experimental results of our simulations. We close with a discussion of current and future work.

2 Motivation rules.

A Student's motivation level (i.e., desire to cooperate or defect) may change depending both on the course of interaction with the Teacher, and on changes in the Student's emotional state. However, our foundational rule involves only the two variables of the interaction, i.e., the choice of each agent to cooperate or defect. We state the rule as follows: a reward for a given choice encourages making the same choice next time; a punishment for a given choice encourages making the opposite choice next time.

This rule is expressed in terms of the motivation variable, using the payoff matrix shown in Figure 1. For example, if the Teacher defects and the Student cooperates, the student's motivation M to cooperate decreases, making the student less likely to cooperate the next time. Note that the rationale behind this rule applies to any goal-oriented meta-game, not just our variants of the MGL. In the context of education, a realistic model may require different reactions from students and teachers, but our foundational rule assumes that the direction of changes in motivation are symmetric with respect to the agents' respective choices. (though we do not assume that the degree of change is similar.)

$Student:$	Cooperate (C)	Defect (D)
$Teacher:$ Cooperate (C)	$M+$	$M-$
Defect (D)	$M-$	$M+$

Figure 1: Changes in Student's Motivation.

3 Emotion rules.

Since the primary focus of this meta-game is on the student's learning, we suppose that an agent's emotional state is affected mainly by the student's actions. If we extend the example illustrating the motivation rule, our supposition suggests that when a student cooperates, both student and teacher tend to feel encouraged and experience a positive change in their emotional state. Conversely, when a student defects (for instance, because the student's motivation level has dropped, or the student becomes tired), both student and teacher may feel discouraged and experience a negative change, although not necessarily to the same degree. This rule is expressed schematically in figure 2.

$Student:$	Cooperate (C)	Defect (D)
$Teacher:$ Cooperate (C)	$E+$	$E+$
Defect (D)	$E-$	$E-$

Figure 2: Changes in Student's Emotion.

4 The Lecture model: a simple Multi Agent MGL.

We combine these rules to create a simple Lecture model, where one pedagogical agent (the Teacher) interacts simultaneously with many (n) Student agents. The Teacher presents a series of concepts and asks the students questions about them. We partition the simultaneous interactions between the teacher and the students into n two-agent meta-games, each one involving the Teacher and a Student interacting over a series of concepts. We represent the concepts as objects with attributes, such as difficulty, that take discrete scalar values, similarly to agent attributes. We assume that the Teacher's questions are the same for each Student, but the Students' responses are taken to be independent.

Since the Teacher's effective action (cooperation or defection) depends on the Student's perception, we introduce a rule to determine the effective action taken by the teacher in the student's meta-game, such as:

```
if (student.ability > concept.difficulty)
  teacher.action ← Defect
else
  teacher.action ← Cooperate
```

In other words, if the student's ability is sufficiently high, the concept seems easy, and the teacher appears not to be challenging the student to learn, i.e., the teacher appears to be defecting. Conversely, if the concept is difficult to the student, the teacher appears to be cooperating.

To describe the student's effective action, as perceived by the teacher, we introduce a rule that relates a student's motivation and emotional state to the student's likelihood of answering a hypothetical evaluation question correctly. If the student is both happy and highly motivated (E high, M high), then the student is presumed to have learned the concept well enough to answer correctly no matter the difficulty of the concept, which is effectively cooperation. If the student is both unhappy and unmotivated (E low, M low), then the student is presumed inattentive and will answer incorrectly no matter the difficulty, which is effectively defection.

If E and M are not binary, we can define numerical thresholds that determine what values are "low" and "high". If the E and M states are dissimilar — E high and M low, or the converse — then we suppose the student has some probability of answering correctly in proportion to the relative difficulty (`delta`) of the concept, which we express as:

```
delta ← (concept.difficulty - student.ability)
```

If `delta` > 0, the likelihood is greatest that the student answers incorrectly (the question was hard), whereas if `delta` < 0, the likelihood is greatest that the student answers correctly (the question was easy). We express this rule as follows:

```
if (E low & M low)
  student.action ← Defect
else
  if (E high & M high)
    student.action ← Cooperate
  else
    if chance < delta
      student.action ← Defect
    else
      student.action ← Cooperate
```

5 Experiments.

We have built a simple simulation of these rules and run a series of preliminary experiments using different ranges of ability and difficulty, different degrees of discreteness for M (motivation) and E (emotion), as well as varying the number of students in the model and the number of iterations. At the beginning of each experiment, the students are assigned random values for ability, emotion and motivation. At each iteration, all the students are presented with the same question (i.e., of the same difficulty level). The difficulty level increases uniformly for all students, one level per iteration. Each student's response to the question is a function of that student's ability, emotion and motivation values. The emotion and motivation determine whether the student will *cooperate* or *defect*, and the corresponding rule is applied (as described in the previous sections).

Figures 3 and 4 compare the results of running the model under a number of different conditions for both binary degrees of M and E (2 values) as well as n-ary (4 values in this example). Each figure contains two columns of graphs, which are arranged pairwise in rows, one row per experiment. The lefthand graph in each row shows end positions for all of the students after the simulation is complete. The horizontal axis respresents the difficulty level, so for example, a student ending up in the 8th horizontal position only progressed to the 8th level (e.g., out of 20 in figure 3a). Note that if multiple students end up at the same position, then only a single mark is indicated on the plot. The corresponding righthand graph illustrates the distribution of students in each position at the end of the simulation. The top subplot in this column has percent on its vertical axis, indicating the percentage of total students who ended up in each position. The vertical axes for the remaining plots are the range of possible values for each variable.

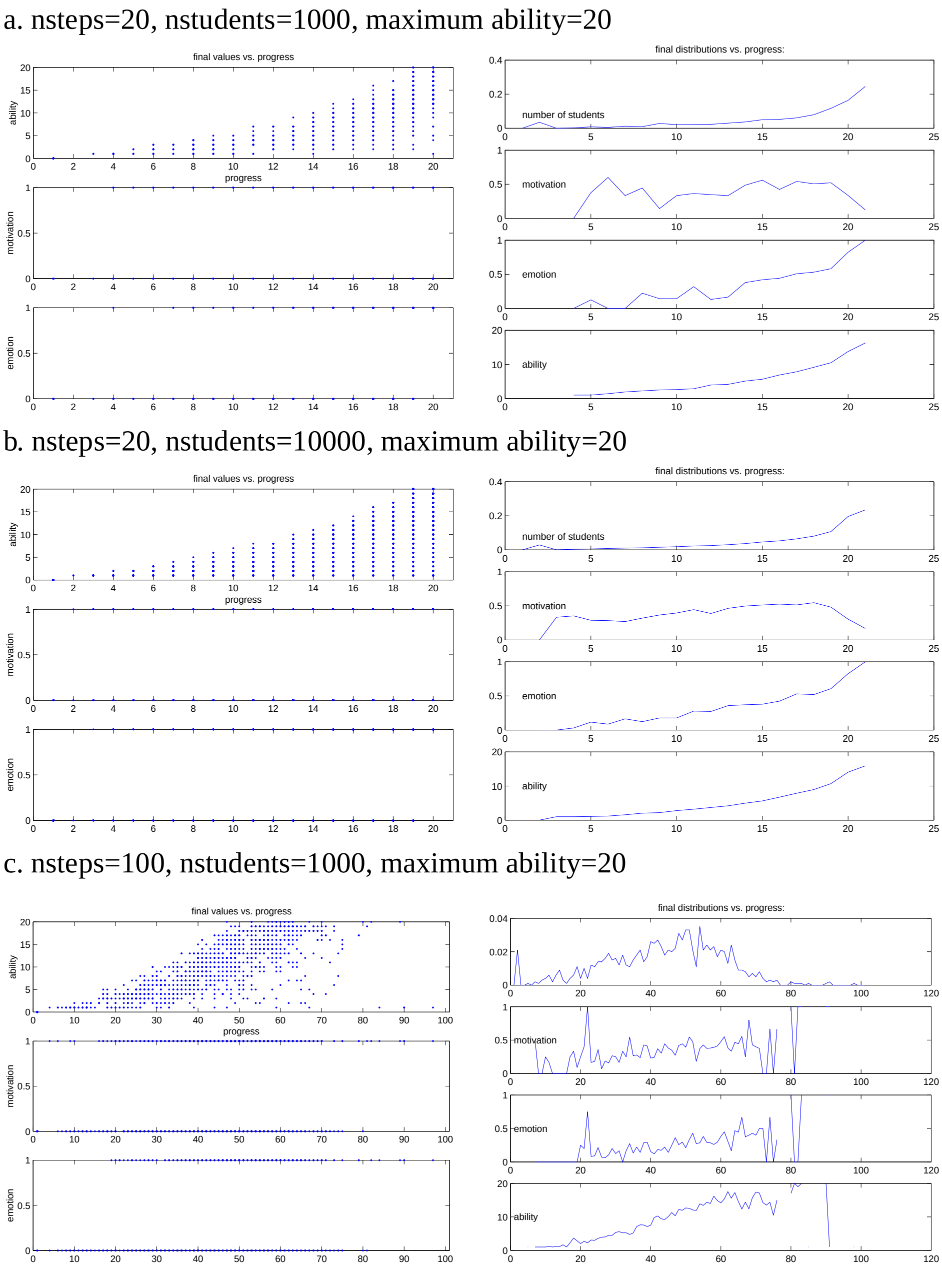
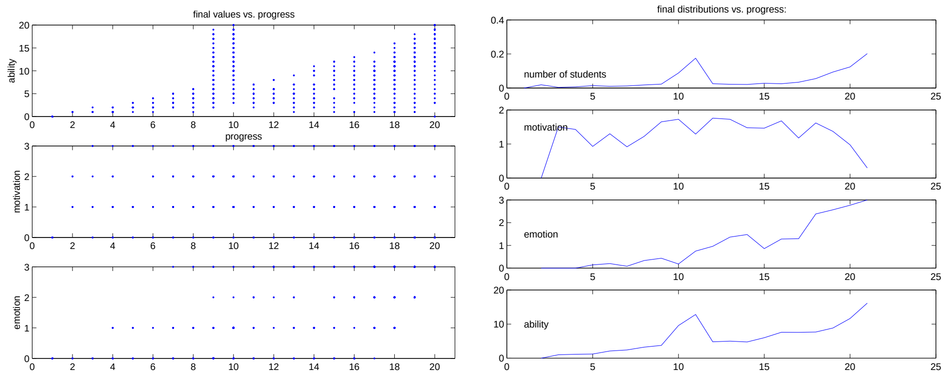


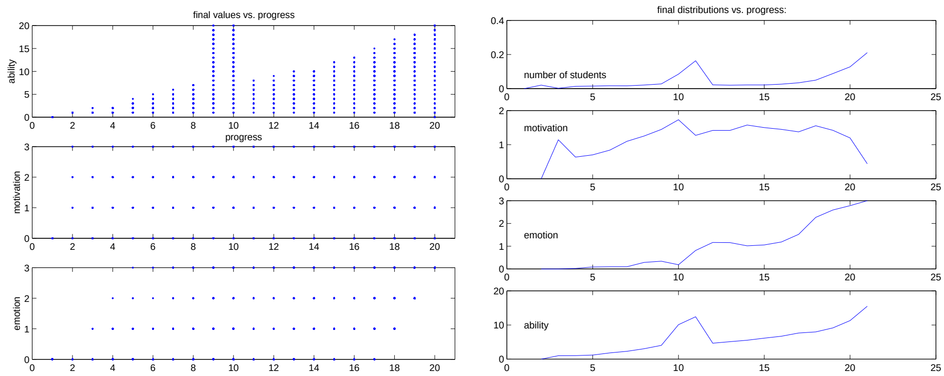
Figure 3: Binary degree.

In varying not just the degree of discreteness, but also the range of abilities, starting level of difficulty, progression of difficulties (linearly increasing, flat, random), number of iterations and number of students, we found the behavior of the model to be quite stable, with the relative progress of students tending to increase with ability and the smoothness of the distribution graphs improving with the number of students simulated. However, within this broad stability we found a number of significant features that are logical, but not immediately obvious, results of the model.

a. nsteps=20, nstudents=1000, maximum ability=20



b. nsteps=20, nstudents=10000, maximum ability=20



c. nsteps=100, nstudents=1000, maximum ability=20

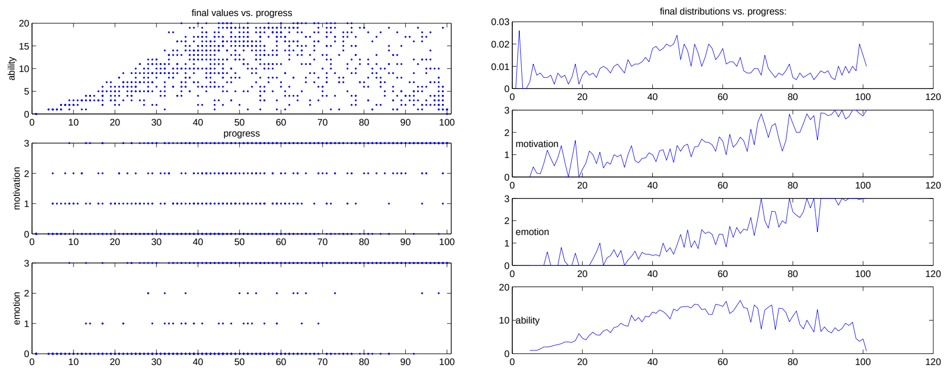


Figure 4: N-ary degree.

In the beginning stages of iteration, when the ability of any student may be greater than the difficulty of the question, there is a significant difference between binary and n-ary degrees of discreteness for emotion and motivation. Careful inspection of the model shows that, for the binary case, if a student defects on an easy question, the student is likely to cooperate on the next question[1]. However, in the n-ary case, if a student with low emotion level and borderline motivation cooperates on an easy question, the student is doomed to defect at least once before cooperating. This is the case because while emotion does not cross the threshhold to “high”, motivation does cross the threshold to “low”. With both low emotion and low motivation, the student will defect no matter what. This pattern may then repeat. In short, students high emotion levels will progress on virtually every iteration, while those with low emotion will progress on at most every other iteration. This is seen in the simulation as a band of students of every ability moving (over time) half as quickly as the students with uniformly high emotion level (see in particular figures 4a and 4b). As degree of discreteness increases, the width of this band increases. In short, the size of possible changes in emotion or motivation per iteration relative to the degree of discreteness has a significant effect on the behavior of the model.

Although a student’s progress is obviously proportional to the number of times the student cooperated, the student’s final motivation level is not a strong correlate with progress until questions become very hard. Of the statistics we tracked, the best correlate of progress turns to be not ability, but the student’s emotion level. Although ability is obviously a strong determinant of progress, we found consistently that students with low ability may, under the right circumstances, actually perform consistently better than students with high ability. We surmise that this is because students with low ability find more of the questions challenging, and are thus rewarded with an increase in motivation when they answer correctly. Such students may be seen as the rightmost outliers in the distribution graphs in both figures.

6 Discussion.

At this point our initial Lecture Model has been specified. Our next step is to simulate many student-teacher interactions over a series of concepts and track how quickly students progress from concept to concept. We can then record information such as the number of interactions it takes to learn a sequence of concepts, the number of interactions per concept across students and per student across concepts, and the number of concepts learned after a certain number of interactions. We are currently experimenting with personalizing the pedagogical agents so that for each student there is an ideal teacher who selects the right question at the right time.

Note that currently, we are neglecting any interactions among students. The problem of agents’ perceptions of each other’s behavior can be illustrated with the following example. Consider a classroom model with one teacher and a number of students, all of whom may interact. Ideally, each agent can perceive whether other agents cooperated or defected on a given interaction, but in practice — especially in a group setting — different individuals may perceive the same behavior differently. For example, the teacher cannot generally pose questions that challenge all students to the same degree; to some students, the teacher is posing easy questions, while to others the questions are very hard. Thus, a teacher’s effective action (as seen by the student) may not be the same as the teacher’s actual or intended action (as seen by the teacher). Similarly, the teacher cannot always tell whether a given student

[1] unless the size of delta is very small

is trying to learn or not; the teacher only knows whether the student answers a question correctly, which in turn depends on the difficulty of the question for that student. Students' responses, even in the ideal case, will thus never be uniform unless the students are identical.

References

[1] R. Axelrod. *The Evolution of Cooperation*. Basic Books, 1984.

[2] J. B. Pollack and A. D. Blair. Co-evolution in the successful learning of backgammon strategy. *Machine Learning*, 32:225–240, 1998.

[3] E. Sklar, A. D.Blair, and J. B. Pollack. Co-evolutionary learning: Machines and humans schooling together. In *Workshop on Current Trends and Applications of Artificial Intelligence in Education: 4th World Congress on Expert Systems*, 1998.