Capturing Student-Robot Interactions for a Data-Driven Educational Dialogue RL Environment

Tristan Maidment, Mingzhi Yu, Erin Walker, Adriana Kovashka, Diane Litman, Timothy Nokes-Malach
Intelligent Systems Program, Computer Science Department, and Learning Research and Development Center
University of Pittsburgh
tdm51@pitt.edu

ABSTRACT

Intelligent social agents provide means to improve student learning and motivation in an educational setting, and support students in a personalized manner. Current educational theory suggests that learning in an interactive setting is best when students are participating equally. We aim to use reinforcement learning (RL) to teach an intelligent social agent to use gaze, gesture, and dialogue to maintain and improve students' participation. Performing reinforcement learning in the real world is often intractable, generally requiring some pre-training in an artificial domain. We examine a data-driven approach using previously collected data to simulate interactions in an educational group setting.

Keywords

Reinforcement learning, teachable robots, balance of participation, lexical entrainment

1. INTRODUCTION

Consider the scenario of interacting with an intelligent teachable robot, which is designed to facilitate collaborative learning, by encouraging student engagement and providing intelligent feedback to students. Reinforcement Learning (RL) provides means for this teachable robot to adapt to the students, which can result in a more personalized learning experience. We are investigating how a teachable robot can use gesture, gaze, and dialogue, to enhance students' collaborative learning within math.

2. BACKGROUND

Teachable robots have demonstrated potential improvements for student learning and motivation, in an educational setting [8][9]. For our setting, we investigate a teachable agent (named Emma) in the role of an intelligent social agent, used to teach and learn with a small group of students. Previous work has demonstrated that empathetic robots in this peer tutoring scenario have been favorable when compared to a human facilitator, by producing more meaningful dialogue between the students [1]. In a one-on-one setting, teachable robots have also demonstrated improvements in student learning and motivation [13][7]. Here we consider the dyadic setting, where the robot learns math with two students. This collaborative learning setting offers benefits in terms of educational theory but creates additional complexity for RL; the intelligent agent needs to keep track of both interactions between itself and the students, as well as the interactions between the two students.

2.1 Theoretical Frameworks

Our current work follows the theory set forth by the interactive-constructive-active-passive (ICAP) framework of cognitive engagement [2]. ICAP hypothesizes that learning in an interactive setting is best when both students contribute constructively and participate equally. Lexical entrainment (LE) measures the similarity between the language used by speakers in a group over time. LE has demonstrated a correlation to student success in group settings [6]. Our teachable robot will attempt to use a combination of gaze, gesture, and dialogue to maintain and improve the group's balance of participation, and promote lexical entrainment.

2.2 Reinforcement Learning

To learn the best usage of gaze and gesture, we look to RL to automatically learn which behaviors best encourage balanced participation and LE. However, reinforcement learning in the real world is particularly difficult, due to the sample inefficiency of online RL algorithms, and the high level of environmental variance that can be present in the real world. For that reason, many deployments of RL in the real world involve using a policy that is initially trained in an artificial domain, typically a computer simulation, and then is transferred, and trained further in the real world.

Reinforcement learning in education typically involved modeling the learner to provide a more personalized learning experience [5][3], although more recent works are model-free [12]. We describe our work towards developing a data-driven simulation of Emma's student interactions for a more personalized learning experience, without modeling the learner.

3. METHODS

3.1 Data Collection

Virtual sessions were held using an online video-chat service to connect the students with Emma, the teachable robot. For our pilot studies, we collected data from undergraduate students, who interacted with the robot in groups of two.

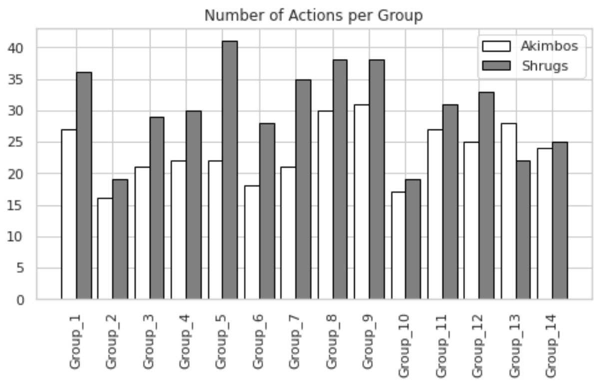

Figure 1: The distribution of actions that were taken by the teachable robot during each session (group).

Twenty-eight participants were provided a series of math questions, specifically about ratios, and were instructed to work together to answer the questions and describe the solution to the teachable robot. In total, 14 sessions have been held between Emma and the students. Each session lasted approximately 30 minutes. Students used an in-house web interface to see the math problem that the robot was solving, to advance to the next step of the problem, to "push to speak" with the robot, and to see what the robot most recently said (as a backup to hearing it over the video-chat service).

The video feeds from each session (group) were recorded, along with the audio from each participant. The participants also completed surveys assessing student beliefs and goals related to math learning, technology, and collaboration. They worked on two sets of ratio problems, one before the interaction with the robot, and one after. The study was done under IRB protocol. The audio feeds for each participant were transcribed using an automatic Speech-to-Text system, with word timings maintained. The face area of each participant was extracted from the video feeds for analysis.

3.2 Indicators for Participant Responses

Our approach for creating a data-driven simulation involves capturing the change in various indicator metrics after Emma performs an action, which is used to approximate the balance of participation, lexical entrainment, and emotional response. Recent studies indicate that a learner's emotions, specified valence, and arousal, have a significant effect on some learning outcomes [4]. We call these changes in indicators an "indicator response".

The indicator response is formally defined as the difference between the indicator at some time point during the session and the average indicator value over the next 30 seconds. An "action response" describes the indicator response following an action performed by Emma. Capturing action responses allows for analysis regarding Emma's effects on students.

We consider participants' action responses to Emma using four indicators, Measure of Participation (MoP), Word Co-Occurrence (WCO), Valence (Val), and Arousal (Aro).

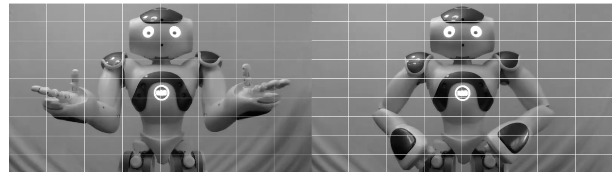

Figure 2: The two actions, "shrug" and "akimbo", that Emma can perform.

These indicators are used to approximate the balance of participation, lexical entrainment, and emotional response. The action responses captured by these indicators will be directly used to construct an environment that is described as a function of the set of indicators. When the simulated robot acts, the indicators transform based on a sample response in the data.

3.2.1 *Measure of Participation (MoP)*

MoP indicates the balance of group participation as proposed by Paletz and Schunn [11]. MoP computes the average level of participation in the group, scaled between 0 (*equal*) and 1 (*dominated*) participation. This metric "provides an unbiased estimate across groups of different sizes and across those that change size over time" [11]. However, in our simplified setting, our groups are consistent in size and do not change in size over time ($n = 2$).

3.2.2 *Word Co-Occurrence (WCO)*

WCO provides a simple estimate of lexical entrainment. WCO is defined as the number of words said that are common between both participants. The size of the shared set of words increases only if both participants are contributing and talking about similar material.

3.2.3 *Valence (Val)*

Facial expressions are used to estimate the students' emotions, using the off-the-shelf package, AffectNet [10]. Valence describes the spectrum of facial emotions from positive to negative. We use this indicator to approximate the immediate emotional response to Emma's actions from both students.

3.2.4 *Arousal (Aro)*

Arousal describes student emotion concerning how aroused (excited) the individual is, on a spectrum from "active" to "passive". Arousal is also captured using facial expressions and AffectNet [10].

3.3 Robot Actions

During the pilot studies, Emma performed two actions, using a hand-coded deterministic policy. Emma is (1) capable of shrugging and (2) standing with her arms akimbo, called "shrug" and "akimbo" respectively.

See Figure 2 for a demonstration of each pose. For the shrug action, the robot raises both arms, with open palms towards the sky. The robot puts her arms on her hips for the akimbo action, with her elbows extended outwards. These actions are performed to accompany her speech. For instance, she

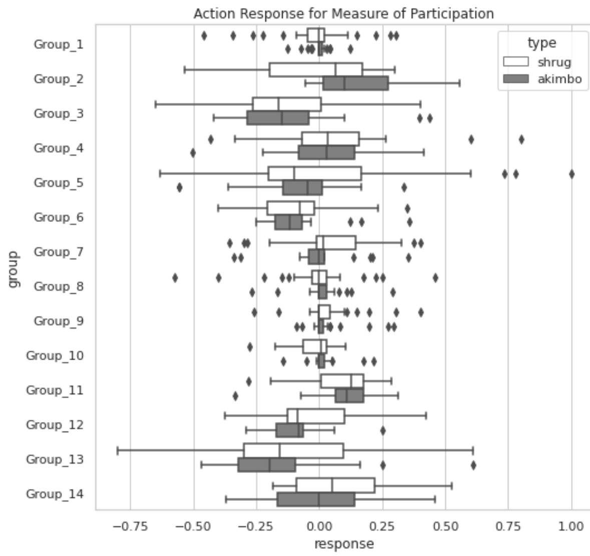

Figure 3: A comparison between the action responses for each group using the Measure of Participation indicator.

performs the "shrug" action when asking a question. The distribution of the action usage is described in Figure 1.

4. RESULTS

4.1 Action Indicator Responses

4.1.1 Measure of Participation

Figure 3 shows the observed action indicator responses for each group and the difference in response per action type. On average, the akimbo action results in a more equal MoP, with an average of -0.016, compared to an average response of -0.007 for the shrug (lower is better).

4.1.2 Word Co-Occurrence

WCO varies depending on the group, but less so than Measure of Participation. We see a higher average increase in WCO after the shrug action compared to akimbo. Shrugs provide a mean Word Co-occurrence action response of 0.115, while akimbo provides an average response of 0.090, higher being better. Figure 4 depicts the action responses for each group.

4.1.3 Valence

Figure 5 describes the change in group valence after actions. The shrug action increased valence and while the akimbo saw a slight decrease, with values 0.017 and -0.0007 respectively. There is a notable difference in groups' action indicator response variance, with groups 2, 7, and 12 demonstrating particularly high variance.

4.1.4 Arousal

Figure 6 provides the action responses per group for the arousal indicator. Shrugs see a slight decrease of arousal, -0.026, whereas the akimbo action provides a slight positive change, 0.005. The trend in groups' action indicator response variances continues; the same set of groups exhibit high variance.

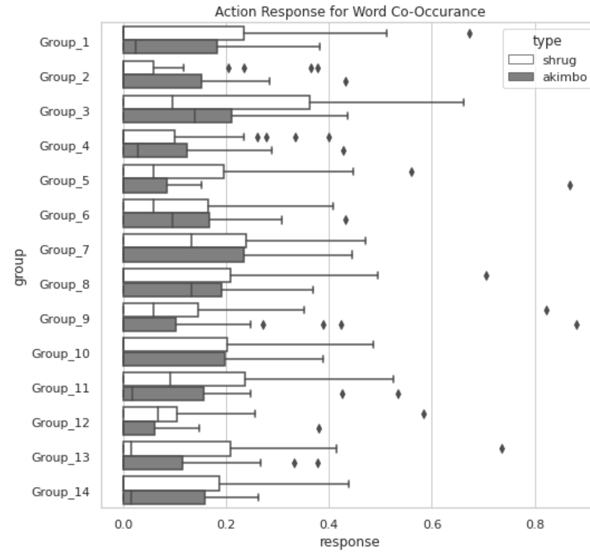

Figure 4: A comparison between the action responses for each group using the Word Co-Occurence indicator.

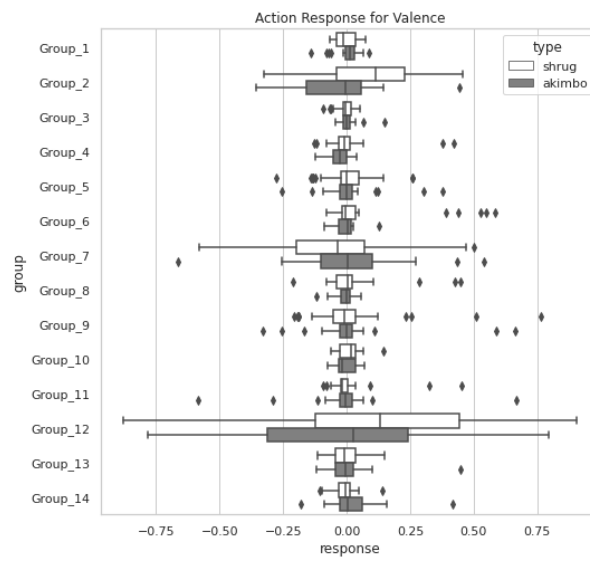

Figure 5: A comparison between the action responses for each group using the Valence indicator.

Metric	Shrug	Akimbo	No Action
MoP	-0.007	-0.016	0.036
WCO	0.115	0.090	0.155
Val	0.017	-0.0007	-0.016
Aro	-0.026	0.005	0.005

Table 1: The average indicator response for each action, and the average indicator response after periods of no action.

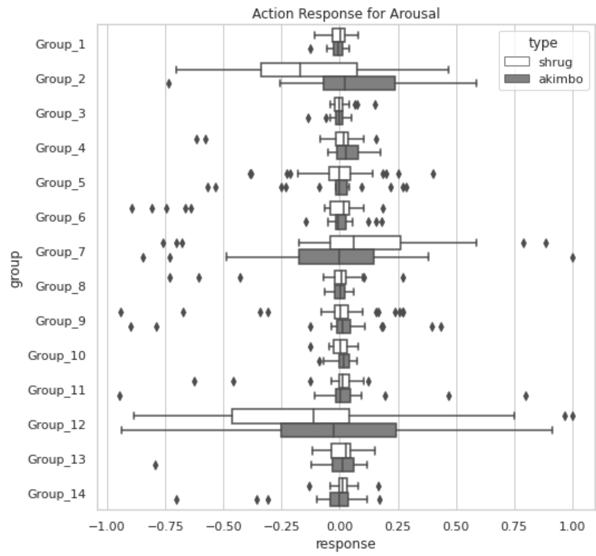

Figure 6: A comparison between the action responses for each group using the Arousal indicator.

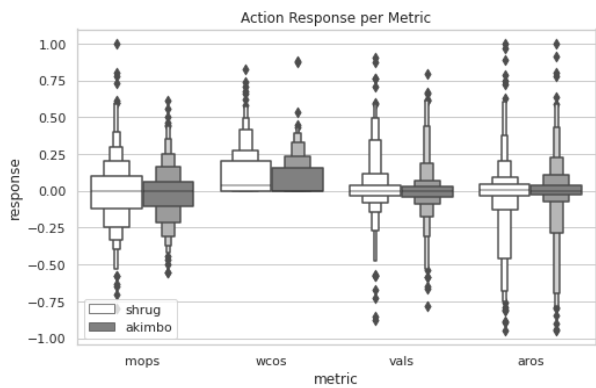

Figure 7: A comparison of the average action response distributions for each action, over all indicators.

4.2 Shrug or Akimbo

In Figure 8, we plot each action's distribution of indicator responses. There appears to be little difference in the distribution of each action's responses, for the population as a whole.

4.3 Action or No Action

To see the utility of the current set of gestures (versus not gesturing), we compare indicator responses during times of robot inaction to the captured action indicator responses. We note more variance after the periods of no action, compared to the periods after an action. We see that the periods following actions result in a slightly more equal measure of participation, compared to periods of no action. However, we see periods following no action have a higher increase in Word Co-Occurrence, than periods after an action. Actions seem to be associated with an increase in valence. Actions do not appear to be associated with arousal.

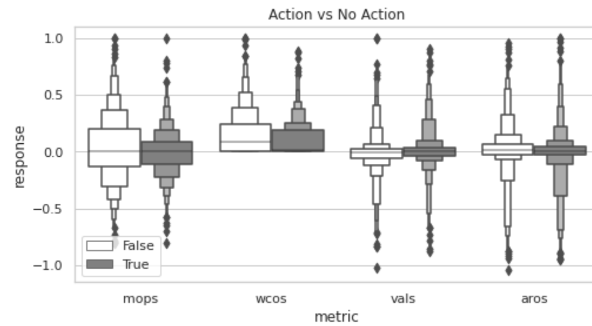

Figure 8: Average change in indicator after an action compared to no action taken.

5. DISCUSSION

As shown in Table 1, the "akimbo" action is on average, followed by a period of more equal MoP for some dyads, thus potentially being correlated to improved student balance of participation. The shrug action also is associated with a period of more equal MoP. To reiterate, Emma's actions accompany her dialogue. Therefore, these indicator responses may be due to students reacting to some combination Emma's of dialogue and gesture. We also note that the shrug action, and associated language, are associated with a higher student valence, and lower arousal.

Figures 3, 5, and 6 highlight a hidden source of variance; student groups respond in significantly different ways to Emma's actions. The group's response distributions have different variances, means, and magnitudes. This source of variance needs to be captured in the virtual environment used to train the RL policy that controls Emma's actions. At the beginning of a session in the virtual environment, a group should be randomly and blindly selected to be used to model the indicator responses.

Notably, some trends are similar between groups. For example, in Figure 6, We see that groups 2 and 12 have similar action response distributions for both the akimbo and shrug action. Groups 2 and 12 also respond similarly in terms of valence, in Figure 5. In Figure 3, groups 3 and 13 appear to respond similarly to actions with respect to MoP. Additional sessions may provide additional insight towards whether or not there are trends that exist between groups.

6. CONCLUSION

This study describes some of the potential effects an intelligent agent's actions have on indicators that depict balance of participation, lexical entrainment, and student emotion, in a setting where two students teach math concepts to an intelligent social robot.

Acknowledgement

This work was partially supported by National Science Foundation Grant No. IIS-2024645 and partially supported by an internal grant from the University of Pittsburgh's Learning Research Development Center.

7. REFERENCES

[1] S. Chandra, P. alves oliveira, S. Lemaignan, P. Sequeira, A. Paiva, and P. Dillenbourg. Can a Child Feel Responsible for Another in the Presence of a Robot in a Collaborative Learning Activity? Oct. 2015.

[2] M. T. H. Chi and R. Wylie. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4):219–243, Oct. 2014. Publisher: Routledge _eprint: https://doi.org/10.1080/00461520.2014.965823.

[3] K. Chrysafiadi and M. Virvou. Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, 40(11):4715–4729, Sept. 2013.

[4] S. Chung, J. Cheon, and S. M. Crooks. How learners' valence and arousal influence learning outcomes in an instructional animation? page 5.

[5] M. C. Desmarais and R. S. J. d. Baker. A review of recent advances in learner and skill modeling in intelligent learning environments. *User Modeling and User-Adapted Interaction*, 22(1):9–38, Apr. 2012.

[6] H. Friedberg, D. Litman, and S. Paletz. Lexical entrainment and success in student engineering groups. pages 404–409, Dec. 2012.

[7] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro. Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human Computer Interaction (Special issues on human-robot interaction)*, 19:61–84, June 2004.

[8] D. Miller, I. Nourbakhsh, and R. Siegwart. Robots for Education. In *Springer Handbook of Robotics*. Jan. 2008. Journal Abbreviation: Springer Handbook of Robotics.

[9] R. Mitnik, M. Nussbaum, and A. Soto. An autonomous educational mobile robot mediator. *Autonomous Robots*, 25:367, Nov. 2008.

[10] A. Mollahosseini, B. Hasani, and M. H. Mahoor. AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. *arXiv:1708.03985 [cs]*, Oct. 2017. arXiv: 1708.03985.

[11] S. B. F. Paletz and C. D. Schunn. Assessing group-level participation in fluid teams: Testing a new metric. *Behavior Research Methods*, 43(2):522–536, June 2011.

[12] H. W. Park, I. Grover, S. Spaulding, L. Gomez, and C. Breazeal. A Model-Free Affective Reinforcement Learning Approach to Personalization of an Autonomous Social Robot Companion for Early Literacy Education. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:687–694, July 2019.

[13] H. W. Park, R. Rosenberg-Kima, M. Rosenberg, G. Gordon, and C. Breazeal. Growing Growth Mindset with a Social Robot Peer. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '17, pages 137–145, New York, NY, USA, Mar. 2017. Association for Computing Machinery.