Modeling One-on-one Tutoring Sessions

Karmen Dykstra [1], Jacob Whitehill [1], Linda Salamanca [2], Monique Lee [2], Amber Carini [2],
Judy Reilly, PhD [2] and Marni Bartlett, PhD [1]
1. Institute for Neural Computation at the University of California, San Diego
2. Department of Psychology at San Diego State University

***Abstract*—The overarching goal of this project is to develop computational models of the nonverbal behavior and interactive strategies observed during face-to-face teaching. This project will help advance the science of learning and teaching by improving our understanding of the dynamics of nonverbal behavior in teaching at a computational level across multiple scales, including low-level facial movements, cognitive and affective processes, and higher level strategic behaviors. In this work we connect higher level teacher and student behaviors to lower level eye gaze dynamics to inform the development of models of nonverbal behavior. Specifically, we use student and teacher behaviors to predict teacher-to-student gaze onset. We additionally model student and teacher behaviors as a probabilistic finite state machine to examine cross-session teaching policy. Findings suggest a relationship between tutoring events and teacher gaze to student, with cross-validation yielding a mean 2AFC performance of .69. Future analysis will connect automated detection of facial affect to the behavioral events explored in this paper.**

I. Introduction

There is growing recognition that social interaction between students and teachers plays a crucial role in the effectiveness of learning. While major research efforts have focused on cognitive aspects of adaptive one-to-one tutoring such as adjusting to student performance [1], nonverbal behaviors present in one-on-one tutoring, such as the use of appropriate facial expressions and gestures by teachers, has been associated with greater student learning [2], student state motivation [3], and student attendance and participation [4]. Motivated by this empirical evidence there has been a growing thrust to develop tutoring systems that respond to student emotional and cognitive states [5]. This project aims to model nonverbal behavior and teaching policies observed in human tutoring sessions to aid development of student-responsive automated tutoring systems.

Preliminary video analysis showed that teacher gaze to student is relatively infrequent during tutoring. Given that an effective teacher, whether human or automated, responds to the cognitive and affective state of the student, and supposing that the teacher gazes at the student to monitor his/her state, knowing when teacher gaze to student occurs could inform when an automated tutoring system should sample student affective state. In the analyses presented in this paper we 1) explore whether we can predict teacher gaze to student using higher-level behavior states and 2) model teacher and student behaviors as a probabilistic finite state machine to gain insight into the flow of teacher-student interaction.

Fig. 1: Student and teacher camera view of a tutoring session.

II. Dataset

We collected video data from 20 face-to-face tutoring sessions on the subject of logarithms. The tutors were two accredited middle school math teachers (1M, 1F) and the participants were 20 typically developing 7th and 8th graders (10M, 10F). Participants were issued a 10 minute pre-test, followed by a 40 minute tutoring session, and a 10 minute post-test. Video was collected simultaneously from 4 camera angles; a view of teacher and student (Fig. 1), a view of teacher only, a view of student only, and an overhead view.

The videos were first transcribed and then each speech unit was labeled according to its function in the tutoring session, like teacher "prompting" or student "correct attempt". An example of a speech-label pair would be student speech unit "Easy" labeled as student "express comprehension". Student and teacher eye gaze were labeled for the duration of the session, where eye gaze labels were 1) gaze to paper 2) gaze to tutor/tutee or 3) gaze elsewhere. Thus far, we have annotated 10 tutoring sessions, all of which were included in the analyses. The functional labels of speech were used as the behavioral labels for analysis.

III. Predicting Teacher Gaze to Student

We fit a Multivariate Logistic Regression (MLR) model to a feature vector describing the presence or absence of 10 high-level behavioral events within a preceding time window of t seconds. The events were: *teacher* explanation, prompting, indirect negation, confirmation, check for comprehension, and present problem, and *student* correct attempt, incorrect attempt, express comprehension and express lack of comprehension. We formed an nxm predictive binary feature matrix X and an nx2 binary label matrix Y (Fig. 2), where n is the number of teacher-to-student (+) or non-teacher-to-student (-) gaze examples and m is the number of behavioral predictors. Negative gaze examples were randomly selected time points

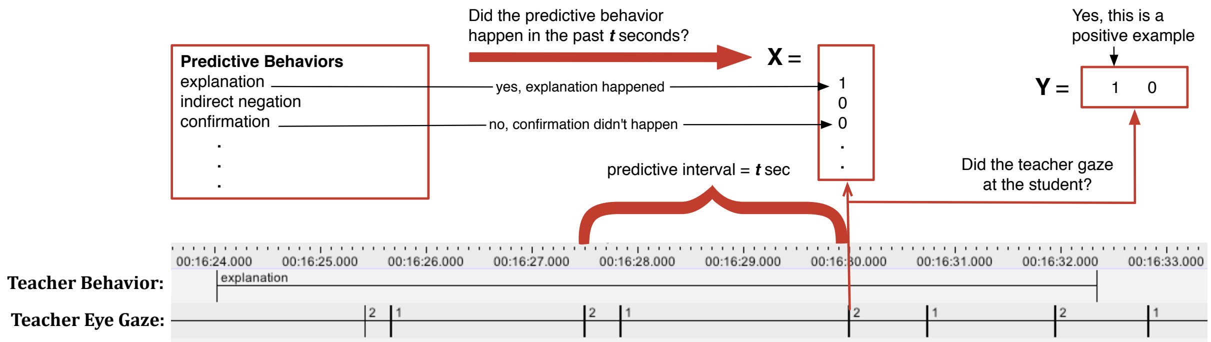

Fig. 2: Binary feature matrix X indicates the occurrence or non-occurrence of behavioral predictors in a 2 second time window before teacher-to-student gaze onsets represented in binary label matrix Y.

within the tutoring session that did not overlap with positive gaze events plus a two second buffer before and after. A behavior's representation within X was formed by looking for its occurrence in a time interval of t seconds before the onset of each gaze event represented in Y. The time interval t was set to 2 seconds by selecting the time interval that gave optimal prediction accuracy for the greatest number of predictors.

The behavioral predictors with the greatest positive weighting were teacher explanation (weight = .0114), check for comprehension (weight = .004), and prompting (weight = .0022), and the most negatively-weighted predictor was teacher present problem (weight = -.0052). These weights indicate that a teacher is more likely to gaze to student following explanation, prompting or checking for comprehension and less likely to gaze to student following presentation of a problem. To evaluate our model we used the 2 Alternate-Forced-Choice (2AFC) metric which evaluates the model's ability to distinguish between negative and positive gaze examples. Within subject cross-validation yielded a mean 2AFC performance of .69, where naive performance would equal .50.

IV. State Machine of Teacher/Student Behavior

We also investigated the behavioral dynamics during the tutoring sessions using probabilistic finite state machines. In particular, we estimated the probability that action a directly precedes action b, as well as the probability that action a directly follows action b, for each pair of actions a and b using Maximum Likelihood Estimation. Results were averaged across all 10 sessions and are displayed in Figure 3.

We noted a .7 outgoing transition probability from student correct attempt to teacher confirmation, indicating that a student's correct answer almost always receives teacher confirmation. Also, student expression of lack of comprehension is more likely to transition to teacher prompting (probability of .31) than to teacher explanation (probability of .29), suggesting frequent use of questions for guidance (Fig. 3).

V. Conclusion

Results from modeling behavior transition probabilities and teacher eye gaze provide guidance as to where sampling student affect could inform teaching policy. Continued work

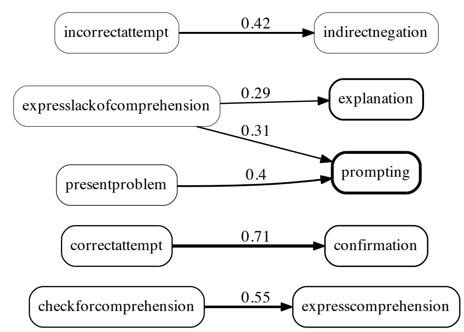

Fig. 3: The highest-probability outgoing behavior transitions averaged across sessions.

will investigate relationships between student affect output by CERT [6] and high-level behaviors to further understand when and how nonverbal behavior impacts one-on-one tutoring.

Acknowledgment

Support for this work was provided by NSF grants IIS-0905622 and SBE-0542013 and NSF REU. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] J. R. Anderson, C. F. Boyle, and B. J. Reiser, "Intelligent tutoring systems," *Science*, vol. 228, pp. 456–462, 1985.

[2] T. P. Mottet, J. Parker-Riley, J. Cunningham, C. Beebe, and S. A. Raffeld, "Testing the neutralizing effect of instructor immediacy on student course workload expectancy violation and tolerance for instructor unavailability," *Communication Education*, vol. 55, pp. 147–166, 2006.

[3] D. M. Christophel, "The relationships among teacher immediacy behaviors, student motivation, and learning," *Communication Education*, vol. 39, pp. 323–340, 1990.

[4] K. A. Rocca, "College student attendance: Impact of instructor immediacy and verbal aggression," *Communication Education*, vol. 53, pp. 185–195, 2004.

[5] B. Kort, R. Reilly, and R. W. Picard, "An affective model of interplay between emotions and learning: Reengineering educational pedagogy - building a learning companion," in *Proc. 2nd IEEE Int'l Conf. Advanced Learning Technologies (ICALT 2001)*, 2001, p. p. 43.

[6] M. S. Bartlett, Littlewort-Ford, and J. R. Movellan, "Computer expression recognition toolbox," in *Proc: 8th Int'l IEEE Conference on Automatic Face and Gesture Recognition, Amsterdam*, 2008.