2007 Student Research Symposium

Language Technologies Institute

Using Transactivity in Conversation for Summarization of Educational Dialog

Mahesh Joshi

Abstract:
This is interdisciplinary work at the intersection of Text Mining and Computer Supported Collaborative Learning (CSCL). We are working on Text Mining techniques for automatic analysis of conversational data in collaborative learning groups. Although computer-based Intelligent Tutors present several pedagogical and technological benefits for individuals, it has been found that collaboration between students, and especially certain types of collaborative learning discussions, have a positive effect on learning over and above what is provided by tutoring technology. Our focus is to apply and extend language technologies to make such collaborative learning interactions more productive.

The long-term goal of this project is to build an automatic conversation summarization system that will help the facilitators or instructors in a collaborative learning setting to quickly identify the groups that need help. For example, given a chat conversation between a pair of students working together to solve a Thermodynamics design problem, the task is to identify portions of their conversation that are good predictors of whether the conversation was beneficial from the perspective of learning Thermodynamics. These portions can then be a part of the summary presented to the instructors, helping them identify groups that need assistance.

It has been found that effective learning in collaborative groups is linked to the process by which learners work on the learning task together and how they build on the contributions of their learning partners [1, 2], otherwise known as transactivity. In our work, we leverage this theoretically motivated idea of transactivity in discourse, and integrate it with a design methodology for conversation summarization. Rather than present to facilitators a condensed version of the interaction meant to convey purely the information content of a conversation, the excerpts selected and included in the summary are meant to give a sense of the nature of the conversation, since the nature of that interaction, such as how transactive it was, is an important piece of evidence for predicting the effectiveness of the interaction for supporting learning. We have applied an existing coding scheme for transactivity, designed by Berkowitz & Gibbs [1], to conversational data from the Thermodynamics domain and have demonstrated that transactive contributions that relate to reasoning of both the students in a pair (dyadic transacts) are better predictors of learning gain than a purely bottom-up approach that performs regression using corpus based features. We have also empirically validated the utility of transactive components in a summary by having human judges rank students

based on their collaborative conversation, with and without the transactive contributions highlighted. To be able to automatically include transactive contributions in a conversation summary, we have experimented with machine learning based classification of conversation contributions as either transactive or non-transactive. To our knowledge, this is the first attempt to automatically identify transactive contributions in a conversation.

For all of the machine learning based evaluations, we carry out leave-one-dialog-out cross-validation to obtain the most valid evaluation metrics. The bottom-up approach uses corpus based features such as unigrams, bigrams, Part-of-Speech bigrams, punctuation and contribution length, as well as meta-level features computed using the contributions of both the students in the pair (such as whether a student was the initiator in a conversation segment, whether he/she was the major contributor and the cosine similarity of the contributions of the two students). Among the several different regression algorithms that we tried, the best correlation between rankings derived from predicted and true learning gain was 0.17 using Support Vector Machine regression with a linear kernel. For the transactivity-based analysis, we computed the percentage of different types of transacts in each student's contributions, using the manually annotated data. The regression model that uses the percentage of dyadic transacts in a student's contributions as the independent variable achieves a correlation of 0.297 ($p<0.05$), a significant improvement over the bottom-up regression approach, although the value is still low. On the task of empirical evaluation by human judges, the average correlation of two judges with the true ranking when using highlighted transactive contributions was 0.299 as compared to a correlation of 0.018 when not using any highlighting. Finally, for the transactive vs. non-transactive classification task, using off-the-shelf classifiers, the best F1-score obtained is 0.506. Plotting the learning curve for this task indicates that performance might be boosted further with more data.

Joshi, M. & Rosé, C. P. (to appear). Using Transactivity in Conversation for Summarization of Educational Dialog. In Proceedings of the Workshop on Speech and Language Technology in Education (SLaTE 2007).

[1] Berkowitz, M., & Gibbs, J. Measuring the developmental features of moral discussion. Merrill-Palmer Quarterly, 29, 399-410, 1983.

[2] Teasley, S. D. Talking about reasoning: How important is the peer in peer collaboration? In L. B. Resnick, R. Säljö, C. Pontecorvo & B. Burge (Eds.), Discourse, tools and reasoning: Essays on situated cognition, 361-384, 1997.