Learning Tutorial Rules Using Classification Based On Associations

Xin Lu[1], Barbara Di Eugenio and Stellan Ohlsson
Department of Computer Science, University of Illinois at Chicago, USA

Abstract. Rules have been showed to be appropriate representations to model tutoring and can be easily applied to intelligent tutoring systems. We applied a machine learning technique, Classification based on Associations, to automatically learn tutorial rules from annotated tutoring dialogues of a human expert tutor. The rules we learn concern the tutor's attitude, the domain concepts to focus on, and the tutor moves. These rules have very good accuracy. They will be incorporated in the feedback generator of an Intelligent Tutoring System.

Keywords. Tutorial rules, Natural language feedback, Expert tutoring

Introduction

To bridge the gap between current Intelligent Tutoring Systems and human tutors, previous studies[1][2] proved that natural language (NL) interfaces could be one of the keys. But it is still not clear what type of NL feedback, and when and how to deliver it in ITSs to engender significantly more learning than simple practice. We are specifically interested in building a computational model of expert tutoring, and to describe how expert tutors will give natural language feedback to their students. In this paper, we present a rule based model of how tutors generate their feedback. This model is motivated by a previous study of ours, in which we found that the expert tutor is indeed significantly more effective than the non-expert tutor[3]. In that study, students were tutored on extrapolating complex letter pattern, such as inferring EFMGHM from ABMCDM. We have already developed a baseline ITS for this task. Our goal is to build a natural language feedback generator for this ITS which will use the tutorial rules we have learned.

Based on the ACT-R theory[4], production rules can be used to realize any cognitive skill. Therefore we can use production rules as a formalism to computationally model expert tutoring. The rules can be designed manually or learned from the human tutoring transcripts. The dialogue management of AUTOTUTOR[5] embeds a set of 15 fuzzy production rules to select the next dialogue move for the tutoring system. The CIRCSIM group has applied machine learning to discover how human tutors make decisions based on the student model[6]. They used Quinlan's C4.5 decision tree learning algorithm[7] to find tutoring rules. They obtained about

[1] Correspondence to: Xin Lu, Computer Science (M/C 152), University of Illinois at Chicago, 851 S Morgan St., Chicago IL, 60607, USA. Email: xlu4@uic.edu

80%~88% accuracy only within the training data set which is also an extremely small sample. Since they only reported accuracy on training but not on testing, it's very hard to understand how good their approach is.

1. Method

Classification based on associations (CBA) [8] which integrates classification and association rule mining can generate class association rules and can do classification more accurately than C4.5. Classification association rules (CARs) are association rules with the target on the right hand side of the rules. A CAR is an implication of the form: $X \rightarrow y$, *where* $X \subseteq I$, *and* $y \in Y$. *X* is a set of features. *I* is the set of all features. *y* is the target class. *Y* is the set of all classes. CBA also provides strength measurements for the CARs:

- **Support:** The rule holds with support *sup* if *sup*% of cases contain *X* or *y*.
- **Confidence:** The rule holds with confidence *conf* if *conf*% of cases that contain *X* also contain *y*.

So when CBA does classification, more than one rule can fit a certain case and the final class will be derived from the rule with highest confidence. If the confidence of the rules that apply is the same, the rule with highest support will be picked. Again if the support is also equal, CBA will classify the case according to the rule which is generated earlier than the others. Of course, there will be some cases that no CARs can classify the case. CBA saves a default class to deal with this kind of situation.

2. Experiments and Results

We collected tutoring dialogues in tutoring the letter pattern extrapolation task with three tutors, one expert and two non-experts. All the sessions are video recorded and 12 dialogue excerpts were transcribed from 6 different subjects with an expert tutor solving two problems in the curriculum. On the transcript we annotated tutor move, tutor's attitude, student move, correctness of student move, student action, student input, student's confidence, hesitation time, the letter relationship currently focused on, and the relationship scope within the problem.

Features used in the rules are the annotations of the tutoring dialogues and the student's knowledge state on each type of letter relationship, which is computed from other annotations within each dialogue excerpts. CBA will automatically generate rules with a subset of these features. Tutorial dialogues are time series data, which means that the prediction of what the expert tutor should do now should be based on information of the last few utterances. In our experiments, we used the features from only the last utterance.

Using 12 dialogues with 6-way cross validation, we did 4 experiments to learn tutorial rules for choosing the tutor's attitude, the letter relationship which the tutor will talk about, the relationship scope within the problem which the tutor will focus on, and the tutor move. Tutor's feedback to students can be divided into 3 categories according to the tutor's attitude towards students: positive, neutral and negative. The letter relationship is the basic concept in the letter pattern task. The relationship scope

Table 1. Experiment Results

Prediction Category	Accuracy in Training	Accuracy in Testing
Tutor's attitude	93.628%	89.490%
Letter relationship	95.987%	90.035%
Relationship scope	95.097%	88.421%
Tutor move	78.261%	56.789%

concerns the coverage of each type of letter relationship. During tutoring, tutors need to choose the concepts to teach students and discuss with them, and also need to decide how to break down the problem and choose an appropriate coverage. A tutor move is akin a response strategy.

Table 1 reports accuracy on training and testing in learning these four sets of tutorial rules. The accuracy for tutor moves is not as high (only about 78% in the training data and 57% in the testing data). One of the possible reasons is that among 8 categories of tutor move, three (summarizing, prompting, instructing) are very difficult to distinguish, even for human annotators. For example, we found that the two annotators disagreed a lot as regards "summarizing" and "instructing". However, the results are sufficient as a basis for our experiments, since our ultimate evaluation measure is whether the natural language feedback generated based on these rules can improve learning.

Acknowledgements

This work was supported by grant N00014-00-1-0640 from the Office of Naval Research, and partly, by awards IIS0133123 and ALT 0536968 from the National Science Foundation.

References

[1]B. Fox. The Human Tutorial Dialogue Project. *Lawrence Erlbaum Associates, Hillsdale*, NJ, 1993.

[2]B. Di Eugenio, D. Fossati, D. Yu, S. Haller and M. Glass. Aggregation Improves Learning: Experiments In Natural Language Generation For Intelligent Tutoring Systems. *The 42nd Meeting of the Association for Computational Linguistics*, Ann Arbor, MI, 2005.

[3]B. Di Eugenio, T. C. Kershaw, X. Lu, A. Corrigan-Halpern, and S. Ohlsson: Toward a computational model of expert tutoring: a first report. In *Proceedings of 19th International conference of Florida Artificial Intelligence Research Society*, 2006, 503-508.

[4]J. R. Anderson, C. F. Boyle, A. T. Corbett, and M. W. Lewis: Cognitive modeling and intelligent tutoring. *Artificial Intelligence*, 42(1), 1990.

[5]A. Graesser, N. K. Person, and D. Harter: Teaching tactics and dialog in autotutor. *International Journal of Artificial Intelligence in Education*, 2005.

[6]Y. Zhou: Building A New Student Model to Support Adaptive Tutoring in a Natural Language Dialogue System. Doctoral dissertation, Illinois Institute of Technology, 2000.

[7]R. J. Quinlan: C4.5: Programs for Machine Learning, 1993.

[8]B. Liu, W. Hsu, and Y. Ma: Integrating classification and association rule mining. In *Knowledge Discovery and Data Mining*, 1998, 80–86.