Tutorial Dialog in an Equation Solving Intelligent Tutoring System

Leena M. Razzaq, Neil T. Heffernan
100 Institute Road, Computer Science Department
Worcester Polytechnic Institute
Worcester, MA, USA

Abstract. A new intelligent tutoring system is presented for the domain of solving equations. This system is novel, because it is an intelligent equation-solving tutor that combines a cognitive model of the domain with a model of dialog-based tutoring. The tutorial model is based on the observation of an experienced human tutor and captures tutorial strategies specific to the domain of equation-solving. In this context, a tutorial dialog is the equivalent of breaking down problems into simpler steps and asking new questions before proceeding to the next step. The resulting system, named E-tutor, was compared, via a randomized controlled experiment, to a traditional model-tracing tutor that does not engage students in dialog. Preliminary results using a very small sample size showed that E-tutor capabilities performed better than the control. This set of preliminary results, though not statistically significant, shows promising opportunities to improve learning performance by adding tutorial dialog capabilities to ITSs. The system is available at www.wpi.edu/~leenar/E-tutor.

1 Introduction

This research is focused on building a better tutor for the task of solving equations by replacing traditional model-tracing feedback in an ITS with a dialog-based feedback mechanism. This system, named "E-tutor", for Equation Tutor, is novel because it is based on the observation of an experienced human tutor and captures tutorial strategies specific to the domain of equation-solving. In this context, a tutorial dialog is the equivalent of breaking down problems into simpler steps and then asking new questions before proceeding to the next step. This research does not deal with natural language processing (NLP), but rather with dialog planning.

Studies indicate that experienced human tutors provide the most effective form of instruction known [2]. They raise the mean performance about two standard deviations compared to students taught in classrooms. Intelligent tutoring systems can offer excellent instruction, but not as good as human tutors. The best ones raise performance about one standard deviation above classroom instruction [7]. Although Ohlsson [9] observed that teaching strategies and tactics should be one of the guiding principles in the development of ITSs, incorporating such principles in ITSs has remained largely unexplored [8].

2 Our Approach

E-tutor is able to carry on a coherent dialog that consists of breaking down problems into smaller steps and asking new questions about those steps, rather then simply giving hints. Several tutorial dialogs were chosen from the transcripts of human tutoring sessions collected to be incorporated in the ITS. The dialogs were designed to take the place of the hints that are available in the control condition. E-tutor does not have a hint button. When students make errors they are presented with a tutorial dialog if one is available. The student must respond to the dialog to exit it and return to solving the problem in the problem window. Students stay in the loop until they respond correctly or the tutor has run out of dialog. This forces the student to participate actively in the dialog. It is this loop that we hypothesize will do better at teaching equation-solving than hint sequences do. When the tutor has run out of dialog, the last tutorial response presents the student with the correct action and input similar to the last hint in a hint sequence. A close mapping between the human tutor dialog and the ITS' dialog was attempted.

Evaluation. E-tutor was evaluated with a traditional model-tracing tutor as a control. We will refer to this tutor as "The Control." The Control did not engage a student in dialog, but did offer hint and buggy messages to the student. Table 1 shows how the experiment was designed.

Because of the small sample size, statistical significance was not obtainable in most of the analyses done in the following sections. It should be noted that with such small sample sizes, detecting statically significant effects is less likely. A large note of caution is also called for, since using such small sample sizes does make our conclusions more sensitive to a single child, thus possibly skewing our results.

Table 1. Experimental Design

Time	Control Condition (7 students)	Experimental Condition (8 students)
20 minutes	Paper and pencil pre-test	
5 minutes	Demonstration of both systems	
Control: average 47.8 minutes Experimental: average 55.6 minutes	Used Control Tutor	Used E-tutor
20 minutes	Paper and pencil post-test	

Learning Gains by Condition. To check for learning by condition, a repeated measure ANOVA was performed using experimental or control condition as a factor. The repeated measure of pre-test and post-test was a factor, with prediction of test score as the dependent variable. Due to the small sample size, we found that the experimental group did better on the pre-test by an average of about 1.5 points; the difference was bordering on statistical significance ($F(1,9) = 3.9$, $p = 0.07$). There was marginally statistically significant greater learning in the experimental condition than in the control condition ($F(1,9) = 2.3$, $p = 0.16$). The experimental condition had average pre-test score of 5.67 and post-test score of 6.67, showing a gain score of 1 problem. The control had average pre-test score of 4 problems correct and average post-test score of 4.2 problems correct. The effect size was a reasonable 0.4 standard deviations between the experimental and control conditions, that is, an effect size for E-tutor over the Control.

3 Conclusion

The experiment showed evidence that suggested incorporating dialog in an equation-solving tutor is helpful to students. Although the sample size was very small, there were some results in the analyses that suggest that, when controlling for number of problems, E-tutor performed better than the Control with an effect size of 0.4 standard deviations for overall learning by condition.

There were some limitations in this research that may have affected the results of the experiment. E-tutor presented tutorial dialogs to students when they made certain errors. However, the Control depended on student initiative for the appearance of hints. That is, the students had to press the Hint button if they wanted a hint. Although students in the control group were told that they could request hints whenever they wanted, the results may have been confounded by this dependence on student initiative in the control group. We may also be skeptical about the results because the sample size was very small. Additionally, the experimental group performed better on the pre-test than the control group, so they were already better at solving equations than the control group.

In the future, an experiment could be run with a larger and more balanced sample of students which would eliminate the differences between the groups on the pre-test. The confound with student initiative could be removed for a better evaluation of the two conditions. Another improvement would be to employ more tutorial strategies. Another experiment that controls for time rather than for the number of problems would examine whether E-tutor was worth the extra time.

References

1. Anderson, J. R. & Pelletier, R. (1991). A development system for model-tracing tutors. In *Proceedings of the International Conference of the Learning Sciences*, 1-8. Evanston, IL.

2. Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-one Tutoring. *Educational Researcher*, 13, 4-16.

3. Graesser, A.C., Person, N., Harter, D., & TRG (2001). Teaching tactics and dialog in AutoTutor. *International Journal of Artificial Intelligence in Education.*

4. Heffernan, N. T., (2002-Accepted) Web-Based Evaluation Showing both Motivational and Cognitive Benefits of the Ms. Lindquist Tutor. *SIGdial endorsed Workshop on "Empirical Methods for Tutorial Dialogue Systems" which was part of the International Conference on Intelligent Tutoring System* 2002.

5. Heffernan, N. T (2001) *Intelligent Tutoring Systems have Forgotten the Tutor: Adding a Cognitive Model of Human Tutors.* Dissertation. Computer Science Department, School of Computer Science, Carnegie Mellon University. Technical Report CMU-CS-01-127 <http://reports-archive.adm.cs.cmu.edu/anon/2001/abstracts/01-127.html>

6. Koedinger, K. R., Anderson, J. R., Hadley, W. H. & Mark, M. A. (1995). Intelligent tutoring goes to school in the big city. *In Proceedings of the 7th World Conference on Artificial Intelligence in Education*, pp. 421-428. Charlottesville, VA: Association for the Advancement of Computing in Education.

7. Koedinger, K., Corbett, A., Ritter, S., Shapiro, L. (2000). Carnegie Learning's Cognitive Tutor™: Summary Research Results. http://www.carnegielearning.com/research/research_reports/CMU_research_results.pdf

8. McArthur, D., Stasz, C., & Zmuidzinas, M. (1990) Tutoring techniques in algebra. *Cognition and Instruction*. 7 (pp. 197-244.)

9. Ohlsson, S. (1986) Some principles for intelligent tutoring. *Instructional Science, 17*, 281-307.

10. Razzaq, Leena M. (2003) *Tutorial Dialog in an Equation Solving Intelligent Tutoring System.* Master Thesis. Computer Science Department, Worcester Polytechnic Institute. <http://www.wpi.edu/Pubs/ETD/Available/etd-0107104-155853>