Exploring the Effectiveness of Knowledge Construction Dialogues

C. P. Rosé, D. Bhembe, S. Siler, R. Srivastava, K. VanLehn
Learning Research and Development Center, University of Pittsburgh, Pittsburgh PA, 15260

Abstract.
The goal of the Atlas project has been to provide opportunities for students to construct their own knowledge and to learn actively by conversing with a natural language-based ITS. We report the results of an evaluation comparing student learning of basic qualitative physics concepts when they engage in natural language dialog, specifically in Knowledge Construction Dialogs (KCDs), with student learning when they simply read about the physics concepts in minilessons. A 2-tailed paired t-test computed over all matched pairs in our student population demonstrates a trend in favor of KCDs with students who had no previous college level physics courses.

1 Introduction

Natural language dialogue offers a wide range of attractive features for intelligent tutoring systems above what is already provided by typical Model Tracing Tutors (MTTs). For example, MTTs in a wide range of domains have previously been criticized for failing to encourage deep learning [6]. If students do not reflect upon the hints they are given, but instead simply continue guessing until they perform an action that receives positive feedback, they tend to learn the right actions for the wrong reasons [1]. Dialogue is one way to encourage students to reflect upon what they have been taught. Tutorial dialogue technology allows us to build tutoring systems that tutor in a style that is closer to that of human tutors.

The goal behind the design of Knowledge Construction Dialogues (KCDs) is to increase the opportunities for students to construct their own knowledge by conversing (in typed form) with a natural language-based ITS. The domain independent Atlas system [2, 5] provides a general purpose planning engine and robust input understanding component that can be used to augment any tutoring system with dialogue capabilities. A set of accompanying authoring tools [4] makes it possible for domain experts to author the lines of reasoning underlying the KCDs. These authoring tools have been used successfully by domain experts with no technical or linguistic background whatsoever. The KCDs used in the study reported here were developed by a team including one physics professor, one computer science PhD, and a psychology graduate student with a physics background. KCDs are interactive directed lines of reasoning that are each designed to lead students to learn as independently as possible one or a small number of concepts. As students are engaged in KCDs using a web interface, they are able to see the dialogue history so far as well as the current tutor turn and a text box for entering their response. When a question is presented to a student, the student types a response in the text box in natural language. If the student enters a wrong or empty response, the system engages the student in a remediation sub-dialogue designed to lead the student to the right answer to the corresponding question. The system selects a subdialogue based on the content of the student's response using authored correspondences between answer classes and remediation sub-dialogue goals, so that incorrect responses that provide evidence of an underlying misconception can be handled differently than responses that simply show ignorance of correct concepts. Once the remediation is complete, the KCD returns to the next question in the directed line of reasoning.

2 Fall 2002 Study

In the current study we contrast KCDs with non-interactive minilessons that present all of the same information provided by the main line of reasoning from the corresponding KCD. To test whether KCDs would be more effective than a minilesson control condition with naive learners, we ran a study in the Fall of 2002. Subjects for this study were University of Pittsburgh undergraduates who had never taken a college level physics course. 35 students have participated in the study so far, which were a combination of paid volunteers and psychology students who earned course credit for their participation. We are continuing to collect data from additional students.

Students first completed a pretest that consisted of 22 multiple-choice questions designed to assess students' knowledge of the target concepts taught by the KCDs or minilessons. Since the subjects participating in this study had never taken college level physics before, we then provided them with a 6 page document summarizing the conceptual physics topics that were going to be covered in the KCDs and minilessons, which were extracted from [3]. Students spent on average 30 minutes reading the overview.

We then assigned students to one of two conditions. In the first condition, namely the KCD condition, students participated in 10 KCDs covering vector components, speed versus velocity, computing average velocity, computing average acceleration, centripetal acceleration, computing weight force, freefall acceleration, dynamic friction force, and Newton's Third Law. In the second condition, students instead read minilessons covering the main lines of reasoning from their corresponding KCDs. In both conditions, after each KCD/minilesson, students were asked to enter a summary of a few sentence giving an overview of the KCD or minilesson. In the KCD condition, students were prompted to elaborate their summaries once after their initial summary. In our pilot testing we noticed that students in the KCD condition typed in extremely terse summaries, in contrast to the minilesson condition students. Even with an additional prompt, the summaries entered by students (adding together the initial one and the additional one) were half as long as those entered by the students in the minilesson condition.

We assigned students to conditions based on pretest score in such a way as to balance our conditions as much as possible. Currently, 16 students have participated in the Minilesson condition and 19 in the KCD condition. We controlled for topic coverage but not for time on task. In this study we found that it took students approximately twice as long to engage in a KCD about a topic than to read a minilesson about that topic. Students in the minilesson condition took on average about a half hour to read all 10 minilessons, while students in the KCD condition took on average about 1 hour to go through all of the corresponding KCDs. After students completed the KCDs/minilessons, they took a post-test, which was identical to the pre-test. We found no significant correlation between time on task and learning either within condition or over the entire population.

We found a trend in favor of the KCD condition. Here we present scores for pre and post test as fractions of 1, indicating the percentage of total points earned by students. Students in the KCD condition had an average pre-test score of .458 with standard deviation .131. They received on average a post-test score of .593 with standard deviation .157. Students in the minilesson condition received on average a pre-test score of .423 with standard deviation .141. They earned on average a post-test score of .571 with standard deviation .144. No subjects reached the maximum possible score on either the pretest or post test; thus, there was no ceiling effect.

We computed a 2-tailed paired t-test with the 14 pairs of subjects from the two conditions that had identical pretest scores. On this subset of the population, average pre-test score, which was identical for both populations, was .470 with standard deviation .120. Average post test score for students in the minilesson condition was .570 with standard deviation .150. Average post-test score for students in the KCD condition was .600 with a standard deviation of .160. The result was significant, ($t(13)=1.13$; $p=.28$).

3 Conclusion and Current Directions

The current study demonstrates a trend in favor of KCDs over a reading control. Thus, the current study provides further evidence of the effectiveness of KCDs for instruction beyond that presented in [5]. We are continuing to collect data in order to further investigate the effectiveness of KCDs in comparison with minilessons.

4 Acknowledgments

This research was supported by the Office of Naval Research, Cognitive Science Division under grant number N00014-0-1-0600 and by NSF grant number 9720359 to CIRCLE, Center for Interdisciplinary Research on Constructive Learning Environments at the University of Pittsburgh and Carnegie Mellon University.

References

[1] V. Aleven, K. R. Koedinger, H. C. Sinclair, and J. Snyder. Combating shallow learning in a tutor for geometry problem solving. In *Proceedings of the Intelligent Tutoring Systems Conference*, 1998.

[2] R. Freedman, C. P. Rosé, M. A. Ringenberg, and K. VanLehn. Its tools for natural language dialogue: A domain-independent parser and planner. In *Proceedings of the Intelligent Tutoring Systems Conference*, 2000.

[3] P. G. Hewitt. *Conceptual Physics*. Adison Wesley, 1987.

[4] P. Jordan, C. P. Rosé, and K. VanLehn. Tools for authoring tutorial dialogue knowledge. In *Proceedings of AI in Education*, 2002.

[5] C. P. Rosé, P. Jordan, M. Ringenberg, S. Siler, K. VanLehn, and A. Weinstein. Interactive conceptual tutoring in atlas-andes. In *Proceedings of Artificial Intelligence in Education*, pages 256–266, 2001.

[6] K. VanLehn, P. Jordan, C. P. Rosé, and The Natural Language Tutoring Group. The architecture of why2-atlas: a coach for qualitative physics essay writing. Proceedings of the Intelligent Tutoring Systems Conference, 2002.