EXPLORING HOW TO ADAPTIVELY APPLY TUTORIAL DIALOGUE TACTICS

Pamela Jordan, Patricia Albacete, Sandra Katz

Learning Research and Development Center, University of Pittsburgh

Paper presented at the IEEE 16th International Conference on Advanced Learning Technologies (ICALT2016)

Austin, Texas, USA July 2016

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A130441 to the University of Pittsburgh. The opinions are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Exploring How to Adaptively Apply Tutorial Dialogue Tactics

Sandra Katz, Pamela Jordan and Patricia Albacete
Learning Research & Development Center
University of Pittsburgh
Pittsburgh, USA
katz@pitt.edu

Abstract—**Prior research aimed at identifying linguistic features of tutoring that predict learning found interactions between student characteristics (e.g., incoming knowledge level, gender, and affect) and learning. This paper addresses the question: *What do these interactions suggest for developing adaptive natural-language tutoring systems?* We summarize two studies which investigated interactions between gender and learning in a tutoring system for conceptual physics. We argue that student characteristics such as gender are insufficient to guide the development of adaptive tutorial dialogue systems. A more fruitful approach is to consider the underlying factors and cognitive mechanisms that potentially mediate the relationship between student characteristics and learning from particular tutoring tactics.**

Keywords-adaptive tutoring; natural-language tutoring systems; Rhetorical Structure Theory; physics education

I. INTRODUCTION

Since language is the primary medium of instruction during tutoring, developers of natural-language tutoring systems have emphasized the need for research to identify linguistic mechanisms that support learning during tutoring (e.g., [1-3]). Current approaches (e.g., [2-4]) focus on *interactions* between the student and tutor, instead of on the contributions of either party, in order to identify features of human tutorial dialogue that predict learning and thereby should be simulated in natural-language tutoring systems.

The results of this research indicate that *one size does not fit all*, when it comes to connections between the language of tutoring and learning. In general, linguistic mechanisms do not benefit all students equally. Various factors moderate the effects of tutorial interaction, including students' aptitude, prior knowledge, and affective factors (e.g., [2-4]). For example, [4] focused on cohesive mechanisms between student and tutor dialogue turns and found that simple mechanisms such as word repetition predicted learning for low prior knowledge students but not for more knowledgeable students.

In this paper, we describe two studies that examined interactions between gender and learning from a natural-language tutoring system for physics. The first study investigated the potential benefit of implementing a set of decision rules to guide automated tutorial dialogue, which was derived from corpus analyses of the rhetorical relations that hold between student-tutor dialogue exchanges [3]. The second study focused on one tutoring tactic that these decision rules expressed, in isolation: summarization. Each study found interactions between students' gender and learning. The main question that we explore in this paper is, *how can observed interactions between student characteristics (such as gender) and learning inform the development of adaptive, natural-language tutoring systems?* The findings from these two studies suggest that investigating interactions between gender and learning is inadequate to guide tutoring system design. Researchers need to go further and try to explain observed interactions in terms of potential underlying, mediating mechanisms. We demonstrate why we believe that this level of analysis holds the most promise for guiding adaptive tutoring.

II. A STUDY OF DECISION-RULE GUIDED TUTORING

In order to operationalize the notion of "interactivity" during tutorial dialogue, [3] specified various ways that human tutors use and build upon part or all of students' dialogue turns, while formulating their responses. They observed human physics tutors as they guided students in reasoning about conceptual "reflection questions" that a physics tutoring system posed to students, after students solved quantitative physics problems. Extending prior research that revealed correlations between lexical cohesion mechanisms such as word repetition and learning (e.g., [4]), [3] focused on other linguistic features as possible predictors of learning—in particular, rhetorical relations between student-tutor dialogue turns, using Rhetorical Structure Theory [5] as a framework to describe these relations.

Katz and Albacete [3] then identified features of the discourse context in which rhetorical relations whose frequency predicted learning gains tend to occur, in order to derive a set of domain-neutral "decision rules" that could potentially guide effective coaching in natural-language tutoring systems. For example, the rule, "The tutor should ask 'why' questions when the student does not provide an explanation to support a claim, especially with less knowledgeable students", stems from their finding that tutors' prompts for the antecedent(s) of students' claims predicted learning. We will refer to decision rules such as this as *tutorial dialogue tactics.*

To test the value of explicitly incorporating these tutorial tactics in a natural-language tutoring system, [6] compared student learning from a version of a tutoring system that deliberately implemented the full set of tactics that [3] specified, with a control version that did not explicitly incorporate these tactics, except as necessary to preserve coherence. A web-based natural-language tutoring system for conceptual physics called Rimac served as a research platform for this study. Rimac engages students in conceptual discussions (reflective dialogues) after they solve quantitative physics problems [3, 6].

Albacete et al. [6] found that students learned from both versions of the tutoring system. However, a between-group comparison revealed that students in the Experimental condition (N=125), in which the set of tactics were deployed, outperformed students in the Control condition (N=131), in which the tutoring tactics were not intentionally used ($t(254)=2.078$, $p<0.04$). This effect could not be attributed to between-group differences in prior knowledge or time on task [6]. However, Albacete et al. [6] found interactions due to gender. For open-response test items, which require students to reveal more conceptual knowledge, they found that for all students combined, females (N=117) outperformed males (N=137) ($t(252)=3.025$, $p<0.01$). Comparing female vs. male performance within conditions, they again found that females outperformed males on open-response items in the Experimental condition (54 females; 69 males) ($t(121)=3.190$, $p<0.01$), but not in the Control condition (63 females; 68 males). Thus the version of the tutor that contained the full set of tutorial tactics seems to have benefitted female students more so than male students.

Why might female students benefit more from a decision rule driven version of a tutoring system than male students? Research investigating possible causes of poor retention of young women in physics after high school indicates general differences in how male and female students study physics. As reported in [7], female high school students typically state that they tend to memorize facts instead of striving for understanding. Consequently, it is possible that [6] observed larger gains for female students because these students needed the emphasis on conceptual understanding that the tutorial dialogue tactics prompted. In other words, students' disposition towards learning a body of material might mediate the observed relation between gender and learning.

Certainly, it would be simplistic, and potentially damaging to many students, to develop different versions of a tutoring system for female and male students—one version that frequently and explicitly deploys tutorial dialogue tactics and another version that does not, respectively. However, if further research proves our proposed interpretation to be correct, this would suggest that it would be beneficial for tutoring systems to identify and adapt to students' approach to learning, and/or learning goals, irrespective of gender.

As with the findings for the study reported on in this section, the results of a study on dialogue summarization presented in Section III highlight the need to search for possible hidden factors that might explain observed correlations between tutoring tactics, readily specified student characteristics (e.g., gender and incoming knowledge level), and learning. (At this writing, the study summarized in Section III has been presented only in poster format [8].) It may be just as important, or more so, for a tutoring system to adapt with respect to these underlying factors as to gender.

III. A Study of Dialogue Summarization

End-of-dialogue summaries are intended to simulate human tutors' instructional goals. With reference to unskilled tutors, [9] stated that a post-dialogue summary "serves the function of succinctly codifying a lengthy, multi-turn, collaborative exchange when a question is answered or a problem is solved." (p. 40). For example, consider the following problem and question that a tutoring system asks, after the student has solved the problem:

> **Problem:** Suppose a 0.40 kg football traveling at 27.0 m/s strikes a receiver's chest and comes to rest in about 0.05 s when the receiver catches it.…
> **Reflection question:** Would it be possible for the receiver to stop the ball without having a force exerted on him?

As do several other natural-language tutoring systems (e.g., [10]), Rimac presents a summary at the end of a reflective dialogue, in order to codify its underlying line of reasoning as, for example, the following summary:

> We were trying to determine if it was possible for the receiver in our problem to catch the football (bringing it to rest) without experiencing a net force acting on him...We started out by saying that the football was slowing down. This meant that its velocity was changing.…Finally, we used Newton's third law to determine that if the receiver was exerting a force on the football, then the football must have been exerting a force on the receiver, which thus answered our original question. Good work.

In addition to testing whether receiving a LOR summary after a reflective dialogue within Rimac was better than not receiving a summary, we also tested two summary variants that we observed in human-human tutorial dialogues. In one type of summary, the tutor reminds the student of the main question and its conclusion, but leaves out the detailed, intermediate LOR. We refer to these as Conclusion summaries. An example for the football problem and reflection question is:

> We were trying to determine if it was possible for the receiver in our problem to catch the football (bringing it to rest) without experiencing a net force acting on him. We've determined that the receiver **must** experience a force from the football in order for him to stop it.

Unlike LOR summaries, Conclusion summaries encourage the student to recall the intermediate line of reasoning between question and answer. In the second variation of LOR summaries, the tutor presents the question and its answer, as in Conclusion summaries, but adds additional advice, such as how the discussion could be applied more generally to similar types of problems, which might differ from the given problem in particular ways. We call these Advice summaries; for example:

> We were trying to determine…without experiencing a net force on him. We've determined that the receiver **must** experience a force from the football in order for him to stop it. One of the consequences of Newton's third law is that all forces exist in *pairs*. This means that it would be impossible for the football player to exert *any* kind of force (field or contact force) without experiencing a force *in return*. The only way for the receiver to experience *no* force of any kind would be if he did not exert *any* kind of force on *anything*.

We hypothesized interactions between prior knowledge and learning gains, depending on condition. However, we found that students across conditions learned from interacting with the tutor, and learned similar amounts.

As in the study summarized in Section II [6], we observed interactions between gender and learning gains. Prior to investigating these interactions, we compared incoming knowledge (as measured by pretest) of females vs. males as a whole and in each condition and found no statistically significant differences. Likewise, considering males and females separately, we found no difference in their prior knowledge between conditions. However, female students (N=49) gained significantly more than male students (N=47), ($t(94)=2.096$, $p=.039$). Within conditions, this difference held only for Conclusion summaries ($t(23)=2.081$, $p=.049$), with a trend for these summaries to be better for females (N=13) than males (N=12), but only for test items rated as "difficult" ($t(23)=2.000$, $p=.057$) after we classified items as "easy", "medium", or "difficult". A further analysis of the performance of females in problems of a higher difficulty level revealed that the only condition in which there was a statistically significant difference between pretest and posttest scores for these types of problems was the Conclusion summary condition.

The Conclusion summary may trigger females to self-explain, which has consistently been shown to support learning (e.g., [11]). Alternatively, and in keeping with our discussion about the first study, Conclusion summaries may support a disposition towards rote learning strategies, such as memorization of facts, by crystallizing a question and its answer. In the latter case, the Conclusion summary might be counterproductive for fact-oriented learners, assuming that it is better to be an understanding-oriented learner.

IV. Conclusion and Future Work

Several linguistic mechanisms and tutoring strategies have been found to predict student learning, but not equally for all students. These findings raise the challenge considered in this paper: *how can we use these findings to guide adaptive, automated tutoring?* The two studies summarized in this paper indicate that examining interactions between student characteristics such as gender and learning from tutorial dialogue is too coarse-grained to meet this challenge. However, considering the underlying factors that potentially mediate these interactions seems more promising.

For example, in the first study, the observed gender differences in learning from a tutoring system which incorporated a set of empirically derived tutoring tactics possibly stem from gender-based differences in learning goals and strategies, thereby suggesting that it may be useful to gather information about students' learning dispositions, irrespective of gender, and to tailor instruction accordingly. The second study also indicated an interaction between gender and learning, in this case from a particular type of dialogue summary (Conclusion summaries), which encapsulate a reflection question and its answer. This finding could also be due to gender differences in learning goals/strategies, and suggests a potential benefit of having students who tend to be fact-oriented co-construct dialogue summaries with the tutor—again irrespective of gender. We are currently conducting a study that tests this hypothesis.

In future work, we will continue to explore possible interactions between student characteristics and particular tutoring tactics, and to test hypotheses about *why* (i.e., through what underlying mechanisms?) gender and other student characteristics predict learning from these tactics.

Acknowledgment

We thank Dennis Lusetich, Svetlana Romanova, and Scott Silliman. This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A130441 to the University of Pittsburgh. The opinions expressed are those of the authors and do not necessarily represent the views of the Institute or the U.S. Department of Education.

References

1. Boyer, K.E., et al. *Characterizing the effectiveness of tutorial dialogue with hidden markov models*. in *Intelligent Tutoring Systems*. 2010. Springer.
2. Litman, D. and K. Forbes-Riley, *Correlations between dialogue acts and learning in spoken tutoring dialogues.* Natural Language Engineering, 2006. **12**(02): p. 161-176.
3. Katz, S. and P.L. Albacete, *A tutoring system that simulates the highly interactive nature of human tutoring.* Journal of Educational Psychology, 2013. **105**(4): p. 1126.
4. Ward, A. and D. Litman. *Semantic cohesion and learning*. in *Intelligent Tutoring Systems*. 2008. Springer.
5. Mann, W.C. and S.A. Thompson, *Rhetorical structure theory: Toward a functional theory of text organization.* Text-Interdisciplinary Journal for the Study of Discourse, 1988. **8**(3): p. 243-281.
6. Albacete, P., P. Jordan, and S. Katz. *Is a dialogue-based tutoring system that emulates helpful co-constructed relations during human tutoring effective?* in *AIED: 17th International Conference on Artificial Intelligence in Education*. 2015.
7. Hazari, Z., P.M. Sadler, and R.H. Tai, *Gender differences in the high school and affective experiences of introductory college physics students.* The Physics Teacher, 2008. **46**(7): p. 423-427.
8. Katz, S., P. Albacete, and P. Jordan, *Do summaries support learning from post-problem reflective dialogues?*, in *Proceedings of the 13th International Conference on Intelligent Tutoring Systems*, A. Micarelli and J. Stamper, Editors. 2016, Springer: Zagreb, Croatia.
9. Graesser, A.C., et al., *AutoTutor: A simulation of a human tutor.* Cognitive Systems Research, 1999. **1**(1): p. 35-51.
10. Evens, M. and J. Michael, *One-on-one tutoring by humans and computers*. 2006: Psychology Press.
11. Chi, M.T., et al., *Self-explanations: How students study and use examples in learning to solve problems.* Cognitive science, 1989. **13**(2): p. 145-182.