CycleTalk: Supporting Reflection in Design Scenarios with Negotiation Dialogue

Carolyn Penstein Rosé, Vincent Aleven
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh PA, 15216 USA
cprose,aleven@cs.cmu.edu

Cristen Torrey
Language Technologies Institute
Carnegie Mellon University
5000 Forbes Ave., Pittsburgh PA, 15216 USA
ctorrey@andrew.cmu.edu

ABSTRACT

In this paper we discuss the motivation for a novel style of tutorial dialogue system that emphasizes reflection in a design context by engaging students in negotiation dialogues. Our current research focuses on the hypothesis that negotiation-style dialogue will lead to better learning than previous tutorial dialogue systems because (1) it motivates students to explain more in order to justify their thinking, and (2) it supports the students' meta-cognitive ability to ask themselves the right questions about the choices they make.

Author Keywords

Tutorial Dialogue, self-explanation, design

INTRODUCTION

Among the most important skills one can develop during one's education regardless of the chosen profession or trade is the ability to think critically and construct sound arguments. For example, these skills are foundational to effective conflict resolution, which is a basic facet of all business relations. Furthermore, widespread accessibility of information in recent years, such as through the internet, is empowering individuals to take the initiative to educate themselves about their legal rights and medical options. However, in order for people to benefit from this initiative, they must be prepared to evaluate, filter, and synthesize potentially conflicting information from a wide assortment of sources in order to be able to argue their own interpretation of this morass of information. The centrality of reflection, critical thinking, and argumentation is perhaps most clearly seen in the scientific arena. Understanding how science connects with the real world on a conceptual level involves building mental models, in other words arguments, that use scientific principles to explain why objects interact the way they do. On another level, critical thinking and argumentation are at the heart of the scientific method. Finally, argumentation comes into play when science is applied in an engineering scenario when design trade-offs are evaluated in the light of scientific understanding. Because of these reasons, we are developing the CycleTalk tutorial dialogue system that supports the development of critical thinking and argumentation skills by engaging students in negotiation dialogues in natural language to immerse students in scientific inquiry at these three key levels: (1) *understanding science* at a conceptual level, (2) *doing science* by forming and then testing predictions using a simulator, and (3) *using science* for evaluating design trade-offs. A key feature of our approach is to engage students in negotiation dialogues for the purpose of stimulating reflection and drawing out their reasoning along these three lines and encouraging them to clarify their own thinking.

MOTIVATION FROM PREVIOUS WORK

The important role of language in learning has been affirmed many times from many different angles and with respect to many different types of subject matter. Recent research on student self-explanations supports the view that when students explain their thinking out loud it enhances their learning [6,7,8,17]. A human tutoring study in the Basic Electricity and Electronics domain [19] revealed a trend for Socratic style tutoring dialogues to be more effective for learning than didactic style ones. A possible explanation for this result is that students learn more effectively when they are given the opportunity to reflect and discover knowledge for themselves in an active way [4,14,16]. Stevens, Collins, and Goldin (1979) report that the best teachers tend to use a Socratic tutoring style. A follow-up analysis of the BEE corpus [9] demonstrates a significant correlation between ratio of student words to tutor words and learning, underscoring the importance of encouraging students to talk more during tutorial dialogue. In support of this, an analysis of the WHY2 human tutoring corpus has demonstrated a significant correlation between average student turn length and learning [21]. In further support of the importance of students expressing themselves through language as part of their learning, Chi et al. (2001) demonstrate that students in a pure self-explanation condition performed no worse than students in a human tutoring condition.

These results have spawned an optimistic view about the potential for building highly effective tutorial dialogue systems, capable of combining the advantages of both individualized instruction and interaction in natural language. Significant progress has been made with respect to this research agenda. Many tutorial dialogue systems have been built and have been evaluated with students, often in realistic educational settings [1,18,10,13,3,11]. These formative evaluation studies demonstrate that state-of-the-art computational linguistics technology is sufficient for building tutorial dialogue systems that are robust enough to be put in the hands of students and to provide useful learning experiences to students. A number of these studies show that tutorial dialogue systems have advantages over instructional treatments that do not involve dialogue. At times, however the comparison results were inconclusive, demonstrating that the field is still young and that there is much room for growth.

An evaluation of the AutoTutor system, a tutorial dialogue system in the domain of computer literacy, showed an advantage over re-reading of the textbook of about 0.5 standard deviations [15]. The textbook re-reading condition itself was no better than a no-treatment control condition. Similarly, a recent evaluation of WHY-AutoTutor, a system based on the same architecture as the original AutoTutor but applied to the domain of qualitative physics, demonstrates a significant advantage of this system over a textbook reading control [12]. However, in a different experiment the learning results obtained with WHY-AutoTutor were no better than those in a control condition in which students read targeted "mini-lessons," short texts that covered the same content as that presented in the dialogue [11]. The mini-lesson condition is different from reading textbook text in that mini-lessons tend to be focused specifically on the knowledge and potential misconceptions involved in a specific exercise. It appears to be a high standard against which to compare. Even human tutors are not always more effective a mini-lesson control, although human tutors are significantly more effective than a mini-lesson control condition with students who have no prior background with the subject material [21].

An evaluation of Andes-Atlas, a tutorial dialogue system for the domain of physics, which leads student through directed lines of reasoning, implemented by means of Knowledge Construction Dialogues (KCDs), demonstrated a significant advantage of Andes-Atlas without a significant increase in time-on-task, compared to an otherwise equivalent version of Andes which provided hints rather than dialogues [18]. The results of this study are however somewhat difficult to interpret due to a very high dropout rate (57%).[VA1] While Atlas' KCDs were shown to be more effective than hints in this evaluation of Andes-Atlas, in a different experiment they were not more effective than mini-lessons of the same kind as were used in the evaluation of WHY-AutoTutor, mentioned above [20,23].

A third tutorial dialogue system, the Geometry Explanation Tutor, which is still under development, was evaluated in two classroom studies. As students solve geometry problems, the system helps them through a restricted form of dialogue to state general explanations for their problem-solving steps. In the two evaluation studies, this system was compared against a version that was the same in all respects, except that students explained their steps by means of a simple menu instead of in a dialogue. In the first study, the students who explained in a dialogue had higher learning gains than students who explained by means of a menu [1]. However, the detailed pattern of results was difficult to interpret, in terms of the underlying knowledge that the students may have acquired, rendering the results somewhat inconclusive. In the second classroom study, carried out in a different school with better-prepared students, there was little difference between the two conditions [2].). The inconclusive result is likely to be due the fact that the students already had significant geometry knowledge.

Thus, tremendous progress has been made in the tutorial dialogue community in the past few years. Tutorial dialogue systems have been shown to lead to improved learning, compared to such as controls as textbook reading. At the same time, we know of no studies that have demonstrated conclusively that tutorial dialogue systems provide more effective or efficient instruction than some of the alternatives to which they have been compared, including an otherwise equivalent targeted "mini-lesson" based approach [11,20,23] and a "2nd-generation" intelligent tutoring system with simple support for self-explanation [1]. However, the situation sketched here does present a challenge. How does one develop a tutorial dialogue system that is more effective than the ones developed so far, especially where many of the systems built so far have a solid basis in empirical studies of human tutors and/or results in the cognitive science literature?

CYCLETALK

Our current research focuses on the hypothesis that negotiation-style dialogue will lead to better learning because (1) it motivates students to explain more in order to justify their thinking, and (2) it supports the students' meta-cognitive ability to ask themselves the right questions about the choices they make. Furthermore, we hypothesize that a more effective tutorial dialogue system would move beyond engaging students in understanding science into actually doing science and using science. In order to test that hypothesis, we are developing a novel style of tutorial dialogue system that pushes beyond the limitations of current tutorial dialogue technology by engaging students in negotiation dialogues in a design context. Specifically, we propose to develop CycleTalk, a tutorial dialogue system that builds on an existing "articulate simulator" in the field of thermodynamics. Building upon this foundation, the CycleTalk tutorial dialogue system will engage students in dialogues in which they negotiate the pros and cons of alternative designs for thermodynamic cycles, such as those that form the foundation for steam power plants or refrigerators.

Thus, CycleTalk will support students in understanding science by engaging them in discussions about how principles of thermodynamics play out in simulations of thermodynamic cycles. It will support them in actually doing science, by encouraging students to construct and defend predictions about how changes to example cycles would affect the cycle's efficiency or effectiveness and then testing those predictions using the simulator. Furthermore, it will engage students in using science by asking them to apply their understanding of thermodynamics to make and defend design decisions.

Design skills are essential and yet are difficult for students to acquire. Beyond understanding thermodynamics concepts and how and why individual factors can affect the efficiency of a cycle, design requires students to weigh and balance alternative choices in order to accomplish a particular purpose. Furthermore, design requires not only a theoretical understanding of the underlying science concepts but also a practical knowledge of how these concepts are manifest in the real world under non-ideal circumstances. Because of the intense demands that design places on students, we hypothesize that design problems will provide the ultimate environment in which students will be stimulated to construct knowledge actively for themselves.

For example, for each of the thermodynamic cycles that students explore with CycleTalk, the system will engage them in a discussion of the pros and cons of one or more alternative designs, such as changing the temperature or pressure within (certain) components of the cycle, adding components (e.g., a condensor, feedwater heater, re-heater), or changing the working fluid. Typically, the effect of such changes on thermodynamic efficiency will be considered, but the effect on other desirable characteristics (such as low cost of building the cycle or maintaining it) may be considered as well. As discussed, CycleTalk will employ a negotiation style of dialogue to encourage exploration of the space of alternative design choices.

Tutor: How do you think you could improve the efficiency of a simple Rankine cycle?

Student: Wouldn't it be possible to introduce a second turbine in series with the first one since the steam leaving the original turbine is still hotter than what enters the Boiler?

Tutor: A second turbine would improve the efficiency, but what about the possibility of some steam condensing on the blades of the second turbine as it continues to cool? Can you see a problem with this?

Student: No, I am not sure what you are getting at.

Tutor: What happens when steel continues to come into contact with water?

Student: it could rust?

Tutor: very good. Condensed water on the blades of the turbine would damage the blades. Can you think of a way to avoid this?

Student: A reheating unit could heat up the steam coming out of the first turbine so that it is not likely to condense onto the blades of the second turbine.

Tutor: Good idea! That solves the condensation problem, but what is the downside?

Student: Reheating requires an additional input of energy.

Tutor: True, so we should test the extent to which we would come out ahead of the game by doing this. But before we do that, can you think of any other disadvantages? Higher efficiency saves money, but does this come for free?

Student: No, the reheating unit costs money - between that and the additional turbine - not to mention additional maintenance. I wonder if it would be worth it in the long run...

This example illustrates that in CycleTalk, the point of the negotiation is to teach students to ask themselves the right questions, considering general issues such as efficiency, maintainability, durability, cost of parts, time, etc.

ACKNOWLEDGMENTS

This project is supported by ONR Cognitive and Neural Sciences Division , Grant number N000140410107. The authors would like to thank Ken Forbus for providing the CyclePad simulator to build on as a foundation for this work.

REFERENCES

1. Aleven V., Koedinger, K. R., & Popescu, O. (2003). A Tutorial Dialogue System to Support Self-Explanation: Evaluation and Open Questions. *Proceedings of the 11th International Conference on Artificial Intelligence in Education, AI-ED 2003.*
2. Aleven, V., Popescu, O., Ogan, A., & Koedinger, K. R (2003). A Formative Classroom Evaluation of a Tutorial Dialogue System that Supports Self-Explanation. *Supplemental Proceedings of the 11th International Conference on Artificial Intelligence in Education, AI-ED 2003. Volume VI.*
3. Ashley, K. D., Desai, R., & Levine, J. M. (2002). Teaching Case-Based Argumentation Concepts Using

Dialectic Arguments vs. Didactic Explanations. In S. A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), *Proceedings of Sixth International Conference on Intelligent Tutoring Systems, ITS 2002,* 585-595. Berlin: Springer Verlag.

4. Brown, A. L., & Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. *Cognitive Psychology, 20,* 493-523.

5. Chi., M. T. H., Siler, S., Jeong, H., Yamauchi, T., Hausmann, R. (2001). Learning from human tutoring. Cognitive Science, 25(4), 471-533.

6. Chi, M. T. H. (2000). Self-Explaining Expository Texts: The Dual Processes of Generating Inferences and Repairing Mental Models. In R. Glaser (Ed.), *Advances in Instructional Psychology,* (pp. 161-237). Mahwah, NJ: Erlbaum.

7. Chi, M. T. H., de Leeuw, N., Chiu, M. H., LaVancher, C., (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18:3, 439-477.

8. Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., and Glaser, R., (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13:2, 145-182.

9. Core, M. G., Moore, J. D., Zinn, C., (2003) The Role of Initiative in Tutorial Dialogue*, Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics*, Budapest, Hungary.

10. Graesser, A. C., Bowers, C. A., Hacker, D.J., & Person, N. K. (1998). An anatomy of naturalistic tutoring. In K. Hogan & M. Pressley (Eds.), *Scaffolding of instruction.* Brookline Books.

11. Graesser, A., VanLehn, K., the TRG, & the NLT (2002). Why2 Report: Evaluation of Why/Atlas, Why/AutoTutor, and Accomplished Human Tutors on Learning Gains for Qualitative Physics Problems and Explanations, LRDC Tech Report, University of Pittsburgh.

12. Graesser, A. C., Jackson, G. T., Mathews, E. C., Mitchell, H. H., Olney, A., Ventura, M., Chipman, P., Franceschetti, D., Hu, X., Louwerse, M. M., Person, N. K., and the Tutoring Research Group, (2003). Why/AutoTutor: A Test of Learning Gains from a Physics Tutor with Natural Language Dialog. *Proceedings of the Cognitive Science Society.*

13. Heffernan, N. T., & Koedinger, K. R. (2002) An intelligent tutoring system incorporating a model of an experienced human tutor. In S. A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.), *Proceedings of the Sixth International Conference on Intelligent Tutoring Systems, ITS 2002* (pp. 596-607). Berlin: Springer Verlag.

14. Lovett, M. C. (1992). Learning by Problem Solving versus by Examples: The Benefits of Generating and Receiving Information. In *Proceedings of the Fourteenth Annual Meeting of the Cognitive Science Society* (pp. 956-961). Hillsdale, NJ: Erlbaum.

15. Person, N., Bautista, L., Graesser, A., Mathews, E., & The Tutoring Research Group (2001). In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future, Proceedings of AI-ED 2001* (pp. 286-293). Amsterdam, IOS Press.

16. Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. Educational Psychologist, 27, 91-109.

17. Renkl, A. (2002). Learning from worked-out examples: Instructional explanations supplement self-explanations. *Learning & Instruction,* 12, 529-556.

18. Rosé, C. P., Jordan, P., Ringenberg, M., Siler, S., VanLehn, K., & Weinstein, A. (2001a). Interactive Conceptual Tutoring in Atlas-Andes, In J. D. Moore, C. L. Redfield, & W. L. Johnson (Eds.), *Artificial Intelligence in Education: AI-ED in the Wired and Wireless Future, Proceedings of AI-ED 2001* (pp. 256-266). Amsterdam, IOS Press.

19. Rosé, C. P., Moore, J. D., VanLehn, K., Allbritton, D., (2001b). A Comparative Evaluation of Socratic versus Didactic Tutoring, *Proceedings of the Cognitive Sciences Society.*

20. Rosé, C. P., Bhembe, D., Siler, S., Srivastava, R., & VanLehn, K. (2003a). Exploring the Effectiveness of Knowledge Construction Dialogues, *Proceedings of the 11th International Conference on Artificial Intelligence in Education, AI-ED 2003.*

21. Rosé, C. P., VanLehn, K., & the Natural Language Tutoring Group (2003b). Is Human Tutoring Always More Effective Than Reading: Implications for Tutorial Dialogue Systems, *Proceedings of the AIED 2003 Workshop on Tutorial Dialogue Systems: With a View Towards the Classroom.*

22. Rosé, C. P., Bhembe, D., Siler, S., Srivastava, R., VanLehn, K. (2003d). The Role of Why Questions in Effective Human Tutoring, *Proceedings of the 11th International Conference on Artificial Intelligence in Education, AI-ED 2003.*

23. Siler, S., Rosé, C. P., Frost, T., VanLehn, K. & Koehler, P. (2002). Evaluating Knowledge Construction Dialogues (KCDs) versus minilessons within Andes2 and alone, ITS Workshop on Empirical Methods for Tutorial Dialogue Systems, San Sebastian, Spain.

24. Stevens, A. L., Collins, A., Goldin, S. E., (1979) Misconceptions in Student's Understanding, International *Journal of Man-Machine Studies*, 11, 145-156.