Towards Integrating Human and Automated Tutoring Systems

Steve Ritter, Stephen E. Fancsali, Susan Berman
Carnegie Learning, Inc.
437 Grant St.
Pittsburgh, PA 15219
{sritter, sfancsali, sberman}@carnegielearning.com

Michael Yudelson
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
yudelson@cs.cmu.edu

ABSTRACT

We envision next generation learners having access to both automated and human sources of instruction in a variety of learning contexts. In such contexts, it will be most effective if students can be assisted to appropriately navigate between these sources of instruction. For example, human tutors, when helping a struggling student, might benefit from having access to the learning profile an automated tutor possesses on the student, including what the student already knows, detected misconceptions, inferred affective state and details about the student's work with the automated system before requesting human help. Similarly, an automated tutoring system would benefit from knowledge of interactions during human tutoring session. To facilitate student transitions between these types of systems, we need to understand the factors that best aid students in transitioning between such systems. This poster reports preliminary analyses, suggesting that students who are struggling with the course are more likely to take advantage of the optional human tutoring support and that such use is associated with increased course completion rates, regardless of the student's level of preparation.

Keywords

Human tutoring, intelligent tutoring system, blended approach.

1. INTRODUCTION

Intelligent tutoring systems (ITSs) frequently seek to mimic the best practices of one-on-one human tutors to drive improved student learning outcomes in a manner that is both scalable and cost effective. While extensive research considers a learning context in which a student uses an ITS while having a human instructor available (e.g., in K-12 computer labs), little work considers situations in which students use an automated tutoring system like an ITS alone (e.g., in their homes) while having human tutors available optionally for tutoring sessions via online chat. Data collected under such circumstances has the potential to generate important insight into how instructional "hand-offs" should proceed between such instructional modalities as well as general best practices for human and automated tutoring.

This project builds on more than a decade and a half of research on Carnegie Learning's Cognitive Tutor (CT) ITS [1]. The project leverages a unique dataset comprised of detailed learning records for thousands of students taking an online developmental math course. Students had required CT assignments as well as access to an online chat-based human tutoring service. This dataset allows us to explore the reasons that may lead students to choose to seek help from human tutors while using an intelligent tutoring system. The project also heavily draws on extensive work on tutorial dialogue data [2-3], allowing us to understand the human tutoring interactions that lead to the greatest learning gains within this context. At a technical level, the work further extends prior work exploring tutorial dialogue interactions and their automated classification by incorporating new and previously unavailable machine tutor data.

To the best of our knowledge, the proposed approach we are starting to work towards is the first attempt to address the creation and evaluation of an integrated approach to capitalize on the joint compensatory nature and data exchange between computerized tutors like ITSs and human tutors. We expect tools and results to generalize beyond the specific automated and human tutoring systems examined. For example, we expect knowledge gained from this work to inform us about how to better educate teachers about how to assist students in classrooms using the educational software in physical classrooms and how to build better reporting systems for human tutors helping students in a wide variety of educational applications.

As our first step in understanding how students navigate between CT and human tutoring (HT), we were particularly interested in understanding whether the subset of students who chose to use HT differed substantially in their use of CT and in their outcomes from students who did not use HT. In order to understand whether student preparation for the course affects use of HT, we use student performance in the first week of the course as a proxy for their initial ability in the course.

2. DATA

We collected data from two developmental college mathematics courses (one is a prerequisite for the other) deployed online at a degree-granting institution. Each course took place over five weeks, and the assignment for each week consisted of one large CT module. Each of these modules was broken into sections of content that grouped roughly similar problems. The instructional model within CT employs a mastery learning approach, in which, new problems are given until the CT's estimates of the underlying skills surpasses mastery thresholds. New sections of each math course begin every week; our dataset consists of all CT and HT interactions taking place from June 1 to December 31, 2014. The subject population consists of 16,905 CT users, approximately 3,300 of whom opted to request HT help during the selected period. These students produced over 19,000 human-tutored sessions, with an average length of 22 minutes. Students were predominantly adult learners of college age and older.

3. RESULTS

Table 1 shows primary descriptive statistics for these populations. Statistics for both courses were merged for simplicity since they are quite similar. The data indicate that students who opt to use HT struggled with the courses more than students who did not take advantage of HT. Students using HT have a higher assistance score (number of hints plus number of errors) in CT, as opposed to those who did not use HT. Perhaps as a result of asking for more hints and making more errors, students using HT worked more slowly, completing fewer sections per hour. The measure of sections per hour has been previously found to be predictive of overall course achievement [4].

These results are consistent with the idea that students who are struggling with the course are more likely to take advantage of HT. It seems unlikely that use of HT would have strong effects on course-level measures like amount of assistance or completion of sections per hour, since, on average, students who used HT used it fewer than 6 times in a course covering between 25 and 50 topics.

In contrast to these indicators that students using HT struggle with the course is the data showing that such students are more likely to complete sections in the course. That is, despite the fact that students turning to HT struggle with the course, they complete more sections of the course, indicating that HT may have a broad effect on student persistence.

To further investigate this effect, we use performance in the first module in the course as a proxy for students' initial preparation for the course. To better align Course 1 and Course 2, module 1 performance was converted to a z-score relative to the mean for that course and binned. Bin size was set to 0.5 standard deviations. Figure 1 shows means of course completion probability for each bin for users and non-users of HT with the number of students printed next to each point. At all levels of course preparation, students using HT, although, as we have seen, struggling, are more likely to complete the CT course material.

Table 1. CT and HT statistics: means (standard errors).

	Students using HT	**Students not using HT**
Parameter	**Course 1**	**Course 1**
CT sections attempted	50.25 (0.25)	50.25 (0.25)
CT problems attempted	493.22 (3.39)	359.38 (2.95)
CT assistance score	3003.16 (47.40)	2621.52 (47.70)
CT assistance score per section	62.78 (1.05)	71.36 (1.24)
CT time per student (hours)	35.41 (0.47)	35.85 (0.50)
CT sections mastered per hour	1.57 (0.03)	0.99 (0.02)
HT time per student (minutes)	110.05 (5.14)	N/A
HT utterances per student	352.82 (17.13)	N/A

4. Conclusion

These preliminary analyses provide a basis for understanding the factors that lead students to use HT and for understanding the broad influence of HT on students. These data are suggestive that students who are struggling with mathematics are more likely to use HT. Interestingly, the data are also suggestive that use of HT may have a broad affective influence on students. Despite the relatively small amount of contact with human tutors during the course, it appears that students who take advantage of such contact appear to be more willing to stick with the course and complete more work, despite their struggles with the mathematics.

5. ACKNOWLEDGMENTS

This work is supported by the contract with Advanced Distributed Learning agency of the Department of Defence (award W911QY-15-C-0070).

6. REFERENCES

[1] Ritter, S., Anderson, J.R., Koedinger, K.R., & Corbett, A. (2007). The Cognitive Tutor: Applied research in mathematics education. Psychonomic Bulletin & Review, 14(2), pp. 249-255.

[2] Morrison, D. M., Nye, B., & Hu, X. (2014). Where in the data stream are we?: Analyzing the flow of text in dialogue-based systems for learning. In R. A. Sottilare, X. Hu, H. Holden, & K. Brawner (Eds.) Design Recommendations for Intelligent Tutoring Systems: Volume 2: Adaptive Instructional Strategies and Tactics (pp. 217–223). U.S. Army Research Laboratory.

[3] Rus, V., D'Mello, S., Hu, X., & Graesser, A.C. (2013). Recent Advances in Conversational Intelligent Tutoring Systems, AI Magazine, 34(3):42-54.

[4] Ritter, S., Joshi, A., Fancsali, S.E., and Nixon, T. (2013). Predicting Standardized Test Scores from Cognitive Tutor Interactions. In Proc. of the 6th International Conf. on Educational Data Mining (Memphis, TN, July 6-9, 2013). 169-176.

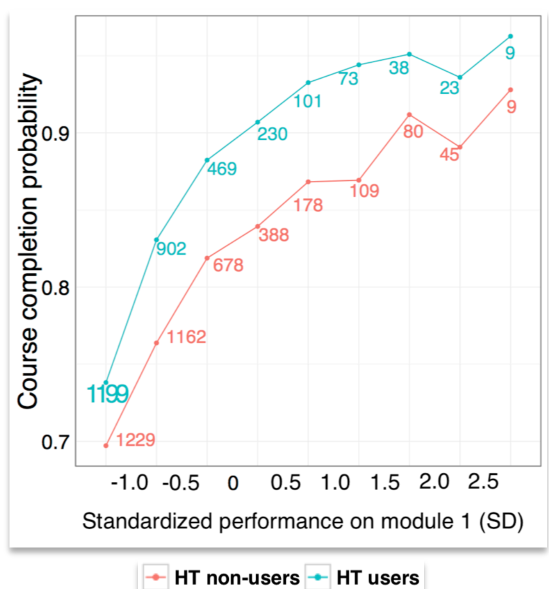

Figure 1. Standardized performance on module 1 vs. overall course completion probability.