Building an Intelligent PAL from the Tutor.com Session Database - Phase 1: Data Mining

Donald M. Morrison, Benjamin Nye, Borhan Samei,
Vivek Varma Datla, Craig Kelly, & Vasile Rus
Institute for Intelligent Systems, University of Memphis

ABSTRACT

In this poster**,** we describe a new research project involving the analysis of nearly 250,000 human-human tutorial dialogue transcripts (in Algebra and Physics) supplied by Tutor.com, a leading provider of online tutorial services for children and young adults. This project involves training a panel of Subject Matter Experts (SMEs) recruited from among Tutor.com's expert tutors to hand-tag a "gold standard" training set of as many as 1,500 transcripts, involving hundreds of different tutors, and potentially totaling more than 100,000 separate utterances. The SMEs will use a theory-based coding scheme to classify utterances into *dialogue acts* and *mode switches*, i.e., dialogue acts that serve to initiate a change in dialogue mode. The resulting training set will be used to train a dialogue act classifier to automatically tag dialogue acts and modes in the remaining transcripts. Machine learning techniques will be used to discover patterns (e.g., sequences, clusters, Markov chains) associated with successful and less successful sessions, where success is measured by internal evidence of learning and also the learner and tutor ratings available in the transcript metadata. Due to the large number of sessions and tutors studied, this research promises to expand our understanding of the prevalence and types of strategies and tactics used by human tutors. Preliminary findings from this data set will be presented during the poster session.[1]

Keywords

Tutorial dialogue; Human tutoring; Data mining; Intelligent tutoring; Computational linguistics; Machine learning; Big data.

1. INTRODUCTION

In recent years, artificial intelligence researchers have begun to apply machine-learning techniques to the analysis of interaction logs generated by online, chat-based (keyboard-to-keyboard) tutorial systems, e.g., [1]. Generally speaking, this approach involves some combination of human tagging of session features (e.g., utterance types), automatic feature detection, and identification of sequential feature clusters. For example, in [1] the researchers tagged the various dialogue acts in a relatively small corpus of tutorial dialogue sessions, then used Hidden Markov Modeling to discover mixtures of dialogue acts associated with identifiable tutorial "modes" [2].

The work described here extends this research, focusing on a large database of nearly 250,000 transcripts of chat-based tutorial dialogues, a subset of a rapidly expanding database of more than 10 million sessions conducted to date by Tutor.com tutors. Our approach features hand-tagging of dialogue acts and mode switches in a training set consisting of more than 1,000 transcripts, the development of an automatic context-sensitive dialogue-act classifier, and a "top-down, bottom-up" cluster analysis aimed at identifying dialogue features associated with positive outcomes, as measured both by the participant quality ratings (available in the transcript metadata). Internal evidence of learning during sessions will also be considered, such as the tutor's feedback on student contributions or student expressions of new understanding, ("Oh, I get it now.")

2. CONCEPTUAL FRAMEWORK

The theory-based coding scheme we are developing views a tutorial dialogue as a special form of human conversation, a joint activity [3] consisting of a sequence of back and forth utterances, each of which represents one or more *dialogue acts* [1].

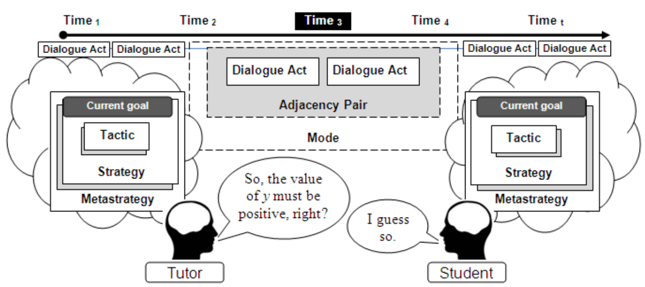

Figure 1. Anatomy of a tutorial dialogue

Dialogue acts are viewed as tactical choices, representing the interlocutors' hidden intentions and strategies, subject to biocultural constraints such as the need to establish common ground [4] and make contributions "relevant" [5]. As such, the significance of a given utterance must be understood in respect to previous utterances (e.g., *adjacency pairs*—[7]), and other higher-level organizational structures such as dialogue modes. Some dialogue modes, such as *openings* and *closings* [7], are common to most human conversation; others, such as *lecturing* and *collaborative problem-solving*, are characteristic of particular kinds of conversation, including the tutorial dialogues we focus on in this study. Successful tutorial dialogues, we hypothesize, are those in which the participants, both tutors and learners, manage to cooperatively align and accomplish their individual goals, drawing on sets of *tactics* (specific dialogue moves), *strategies* (algorithms or "policies" for selecting from among available tactics based on unfolding circumstances), and *metastrategies* (algorithms for selecting from among available strategies). This conceptual framework is explored more fully in related work [6].

[1] Corresponding Author: Donald M. Morrison (chipmorrison@gmail.com)

3. TUTORIAL TRANSCRIPT CORPUS

Our corpus consists of a set of 245,192 tutorial session transcripts shared with us by our partner, Tutor.com, a leading provider of online (chat-based) tutorial services. While this is only a subset of the 10 million (and counting) sessions available, we believe it is orders of magnitude greater than most prior analyses of human tutoring. The sessions represent attempts to help students solve problems and understand related concepts involving selected subtopics in Algebra (65% of the transcripts), and Physics (35%). The transcripts consist of more than 25 million time-stamped lines (corresponding roughly to utterances), representing more than 80,000 hours of dialogue, and containing more than 1,200,00 unique tokens (words and mathematical expressions). Each transcript is linked to a set of metadata, including both tutor and student ratings of session quality.

Table 1: Summary transcript statistics

Subject		Mean	Std. Dev	Total
Physics	Minutes	24.6	21.6	2,123,429
	Tutor lines	68.0	52.4	5,875,944
	Student lines	49.7	39.8	4,530,487
Algebra	Minutes	18.3	18.1	2,897,482
	Tutor lines	55.3	40.0	8,773,353
	Student lines	41.9	30.7	6,355,446

4. RESEARCH PLAN

This research involves five distinct development tracks, as summarized in Figure 1.

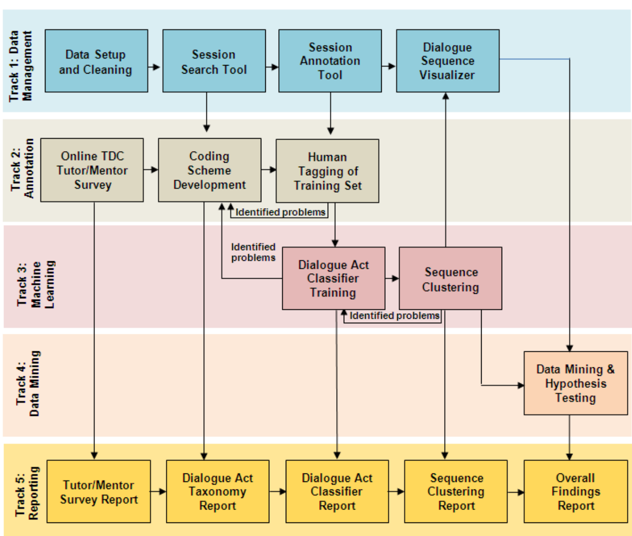

Figure 1: Research Plan Tracks

At this writing we are in the process of data cleaning and descriptive analysis, as well as development of the toolset we will use for session searches, annotation and visualization. One critical task is the development of a web-based annotation environment that will be used to train the human taggers, to hand-tag selected transcripts, and to review and revise transcripts tagged using automated tools.

We have also conducted an online survey of 250 Tutor.com tutors and tutor mentors, consisting of a set of open-ended questions aimed at eliciting the respondents' expert opinions regarding choices of particular tactics and strategies in different circumstances. We are using this data for two purposes: (1) to select a "blue ribbon" panel of tutors and mentors to serve as the Subject Matter Experts (SMEs); and (2) to ensure that our theory-based coding scheme is consistent with how the SMEs themselves think about the dynamics of the tutorial process.

A panel of 15 to 20 SMEs are being recruited to help modify the coding scheme, test the annotation environment, and hand-tag as many as 1,500 session transcripts for both dialogue acts and mode switches, i.e., dialogue acts that have the effect of turning on a particular mode ("Welcome to Tutor.com") or switching from one mode to another ("So, have you tried to do this problem yourself?").

Based on this training set, a dialogue act classifier will be tuned, which we will then use to auto-tag the remaining transcripts in the database. Finally, we will use sequencing and clustering algorithms to discover hidden patterns (interpretable as tactics and strategies) associated with successful and less successful sessions. Sequence-mining is one method we will use to detect patterns within sessions. Since Hidden Markov Modeling has a history of success for this type of analysis, we expect this to be one key technique [1][2]. Clustering will be applied to identify traits that characterize certain types of successful (or less successful) sessions.

The results of this data mining are intended to inform the design of future tutoring and adaptive learning systems. The project is the first phase in a planned multiyear research and development effort funded by the U.S. Department of Defense Advanced Distributed Learning (ADL), aimed at developing hybrid human and artificially-intelligent tutoring systems compatible with ADL's Personal Assistant for Learning (PAL) architecture.

5. ACKNOWLEDGMENTS

The research described here is supported by Tutor.com, an IAC company, under a contract with the U.S. Department of Defense Advanced Distributed Learning Initiative (W911QY-14-C-0019).

REFERENCES

[1] Boyer, K. E., Ha, E., Wallis, M. D., Phillips, R., Vouk, M. A., & Lester, J. C. (2009, July). Discovering tutorial dialogue strategies with Hidden Markov Models. In *AIED* (pp. 141-148).

[2] Cade, W. L., Copeland, J. L., Person, N. K., & D'Mello, S. K. (2008, January). Dialogue modes in expert tutoring. In *Intelligent tutoring systems* (pp. 470-479). Springer Berlin Heidelberg.

[3] Clark, H. H. (1996). *Using language*. Cambridge University Press.

[4] Garrod, S., & Pickering, M. J. (2007). Alignment in dialogue. *The Oxford handbook of psycholinguistics*, 443-451.

[5] Grice, P. 1975. Logic and conversation. *Syntax and semantics*, *3*, 41-58.

[6] Morrison, D.C. & Rus, V. (2014, to appear). Is it a strategy or just a tactic?: A Martian perspective on the nature of human pedagogical dialogue. In Sottilare, R., Hu, X., Graesser, A. and Goldberg, B. (Eds.) *Design Recommendations for Adaptive Intelligent Tutoring Systems: Adaptive Instructional Strategies, Volume II.* Army Research Laboratory.

[7] Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*(4), 289-327