Partially Observable Sequential Decision Making for Problem Selection in an Intelligent Tutoring System

EMMA BRUNSKILL AND STUART RUSSELL, University of Califoria, Berkeley

A key part of effective teaching is adaptively selecting pedagogical activities to maximize long term student learning. In this poster we report on ongoing work to both develop a tutoring strategy that leverages insights from the partially observable Markov decision process (POMDP) framework to improve problem selection relative to state-of-the-art intelligent tutoring systems, and evaluate the computed strategy in the classroom. We also highlight some of the challenges in data mining related to automatically constructing pedagogical strategies.

1. INTRODUCTION

We are interested in creating algorithms to automatically construct adaptive pedagogical strategies for intelligent tutoring systems. Our approach assumes as input a student model, which can be learned from educational data. In this paper we assume that such a model is provided, and focus on the task of computing adaptive policies given the student model; however, in the future we plan to learn the student model from data and at the end of this document we briefly sketch some of the data mining challenges involved.

Different pedagogical activities vary in their instructional effectiveness, and in the information they provide about the student's underlying understanding. Our approach explicitly reason about pedagogical activity features when computing a conditional sequence of activities to provide to a student. We pose tutoring policy creation as an instance of Partially observable Markov decision process (POMDP) planning [Sondik 1971]. The student's knowledge state is represented as a set of hidden variables. The value of these variables can change in response to a pedagogical activity, and these hidden variable values may be probed, or partially observed, such as by posing a test question to the student). A probability distribution over possible knowledge states of the student is maintained and updated using Bayesian filtering: examples of prior work using this approach to student modeling include Corbett and Anderson [1995] and Conati et al. [2002].

There has been some very recent interest in using POMDPs to incorporate Bayesian models of student learning into algorithms for constructing sequential pedagogical policies: see Brunskill and Russell [2010], Rafferty et al. [2011], Tehocharous et al. [2009], and Folsom-Kovarik et al. [2010]. While this initial work is encouraging, none of these prior approaches have performed experiments with standard school curriculum, nor evaluated their approaches relative to existing, high-performing problem selection strategies used in intelligent tutoring systems. Here we report on an ongoing effort to examine whether POMDP planning can further improve the effectiveness of an existing intelligent tutoring system for high school math.

2. APPROACH

We follow Corbett and Anderson [1995] and model the student knowledge state as whether the student has mastered or not mastered a set of N skills. Though the algorithm we developed can be applied more broadly, as our goal was to compare our approach to existing state of the art methods, and to conduct classroom studies, we focus on developing an approach for the Algebra I tutor, produced by the software company Carnegie Learning. In this tutor students do a number of different interactive exercises, and our goal was to adaptively select the best exercise to give to the student, given our Bayesian estimate of the student's underlying knowledge state, in order to help the student learn the most in a limited amount of time. The possible set of exercises is large, and, if an individual exercise consisted of K skills, there were 2^K possible observations that could be received after the student completed the exercise, corresponding to whether or

not the student got each skill correct. Standard approaches to solving POMDPs struggle to scale to domains with large numbers of (pedagogical) actions to select among, and a large set of possible observations. We first explored in simulation using a prior approach by Brunskill and Russell [2010], but the high number of observations possible after a single exercise resulted in challenges for that particular method.

Due to space limitations we only briefly sketch out the algorithm we developed. We computed a depth one POMDP forward search tree from a set of initial knowledge states, and used a heuristic estimate of the value at the leaves of the tree: roughly, this heuristic estimate corresponds to how well we expect the student to perform if we used a heuristic method to select pedagogical activities from this point onwards for a fixed number of activities. These heuristic values were calculated assuming that the student would complete a fixed number of pedagogical activities before receiving a post test.[1] The forward search tree values were cached and combined to construct a look up table that was used to select pedagogical activities during tutoring.

3. ONGOING WORK

We have done some preliminary evaluation of our approach in a classroom study with ninth grade students. However, there were several limitations of this initial study. We are currently focusing on using simulation studies to better evaluate and modify our proposed algorithm.

So far the only pedagogical activities considered as part our algorithm were problem exercises. The POMDP framework should be particularly helpful in trading off among activities with variable components of instruction and evaluation. We hope to soon consider a more diverse set of pedagogical activities types.

Finally, the approach we have described depends critically on an input set of parameters which describe the student model. There has recently been some promising work by [Chi et al. 2010] which learned Markov transition dynamics from a set of data collected using random tutoring strategies, and used that to infer a successful tutoring strategy. However, most existing data on student-tutoring system interactions will involve a non-random strategy. This can lead to sparsity in the collected data as many pedagogical activities will never have been tried from particular student states. Constructing good models in these scenarios, particularly when the student state is modeled as hidden, remains an interesting unfinished challenge.

REFERENCES

BRUNSKILL, E. AND RUSSELL, S. 2010. RAPID: A reachable anytime planner for imprecisely-sensed domains. In *Proc. UAI*.

CHI, M., VANLEHN, K., AND LITMAN, D. 2010. Do micro-level tutorial decisions matter: Applying reinforcement learning to induce pedagogical tutorial tactics. In *Proc. ITS*.

CONATI, C., GERTNER, A., AND VANLEHN, K. 2002. Using Bayesian networks to manage uncertainty in student modeling. *Journal of User Modeling and User-Adapted Interaction 12*, 371–417.

CORBETT, A. AND ANDERSON, J. 1995. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction 4*, 253–278.

FOLSOM-KOVARIK, J., SUKTHANKAR, G., SCHATZ, S., AND NICHOLSON, D. 2010. Scalable POMDPs for diagnosis and planning in intelligent tutoring systems. In *AAAI Fall Symposium on Proactive Assistant Agents*.

RAFFERTY, A., BRUNSKILL, E., SHAFTO, P., AND GRIFFITHS, T. 2011. Faster teaching by POMDP planning. In *Proc. AIED*.

SONDIK, E. 1971. The Optimal Control of Partially Observable Markov Processes. *DTIC Research Report AD0730503*.

TEHOCHAROUS, G., BECKWITH, R., BUTKO, N., AND PHILIPOSE, M. 2009. Tractable POMDP planning algorithms for optimal teaching in "SPAIS". In *IJCAI PAIR Workshop*.

---

[1]Ultimately we are interested in scenarios in which there is a fixed amount of time to provide teaching. Here we approximated a fixed amount of time by modeling a fixed number of problems. In reality different students take varying amounts of time, as do different types of problems.