Integrating Learner Help Requests Using a POMDP in an Adaptive Training System

Jeremiah T. Folsom-Kovarik[1], Gita Sukthankar[2], and Sae Schatz[3]

[1] Soar Technology, Inc., 1942 West County Road 419 #1060, Oviedo, FL 32766

[2] School of Electrical Engineering and Computer Science, University of Central Florida, 4000 Central Florida Blvd, Orlando, FL 32816

[3] MESH Solutions, LLC, 12601 Research Parkway, Orlando, FL 32826

Abstract

This paper describes the development and empirical testing of an intelligent tutoring system (ITS) with two emerging methodologies: (1) a partially observable Markov decision process (POMDP) for representing the learner model and (2) inquiry modeling, which informs the learner model with questions learners ask during instruction. POMDPs have been successfully applied to non-ITS domains but, until recently, have seemed intractable for large-scale intelligent tutoring challenges. New, ITS-specific representations leverage common regularities in intelligent tutoring to make a POMDP practical as a learner model. Inquiry modeling is a novel paradigm for informing learner models by observing rich features of learners' help requests such as categorical content, context, and timing. The experiment described in this paper demonstrates that inquiry modeling and planning with POMDPs can yield significant and substantive learning improvements in a realistic, scenario-based training task.

Introduction

Trainers and teachers become more effective when they can tailor their instruction to the needs of individual learners, rather than trying to find one lesson that can reach a large group of learners. Because adapting instruction makes it more effective, people who learn in a one-on-one setting can greatly outperform those who spend the same amount of time studying the same material in a classroom setting (e.g., Bloom 1984; Koedinger et al. 1997).

Adaptive trainers and other intelligent tutoring systems (ITSs) are computer programs that train or teach individual learners in personalized ways. The goal of adapting instruction with an ITS is to act more like effective, one-on-one tutors than like classroom lecturers.

One of several ways in which current ITSs still differ from human tutors lies in the questions learners ask during instruction. When interacting with a human tutor, learners ask many times more questions than they do in a classroom setting (Graesser and Person 1994). Intuitively, an effective ITS should empower learners to ask questions as well.

Learners who ask questions during instruction can enjoy benefits such as improved recall and deeper understanding (e.g., Harper, Etkina, and Lin 2003). Additionally in the ITS field, *inquiry modeling* describes the idea that learners' questions can provide valuable new input to an ITS's learner model (Folsom-Kovarik et al. 2010a). Questions can reveal an individual's knowledge, affect, and other states or traits that can help an ITS adapt its instruction. The more freedom to ask for help a learner has, the more information an ITS could draw from each help request.

The present paper describes an evaluation of an inquiry modeling ITS, along with a second emerging method: the ITS implements inquiry modeling with a partially observable Markov decision process (POMDP) learner model.

POMDPs are a mathematical representation of sequential decision-making under uncertainty (Kaelbling, Littman, and Cassandra 1998). With a POMDP model, an ITS can interpret uncertain inputs, plan a course of action, and modify its plan whenever new information appears.

Planning might improve instruction in any domain when a greedy policy is not best. For example an ITS could plan ahead and improve long-term learning by deciding not to correct a learner's mistake immediately, or it might spend early instruction time helping a learner feel less frustrated in order to make later instruction more effective.

However, POMDPs' powerful planning requires great complexity. POMDP solvers' search space grows exponentially with the number of modeled states. An ITS might model over 100 orthogonal facts about each learner (e.g., Payne and Squibb 1990), making general problem representations intractable. This paper presents an evaluation of new, specialized representations for the ITS domain that let POMDPs model and reason about these large problems.

This paper also describes the implementation and evaluation of an ITS in a military training scenario. It contributes empirical support for the intuition that inquiry modeling and planning with POMDPs can improve an ITS's instructional effectiveness in a realistic training domain. The experimental findings demonstrate that inquiry modeling and POMDP ITSs are promising emerging applications.

The IMP Adaptive Trainer

The Inquiry Modeling POMDP (IMP) adaptive trainer was created to evaluate inquiry modeling and planning with POMDPs in human studies. IMP's learner model leverages certain properties common to intelligent tutoring problems to make its problem representation tractable. While the contents of its model target a specific instructional domain, described below, the IMP architecture has the potential to train or teach in a range of realistic instructional scenarios.

One important property of ITS work, in general, is that instructional material often has a natural presentation order. For example, it is difficult to teach multiplication before addition, or exponentiation before multiplication. Although the order is not always absolute and a tutor can often skip material or backtrack as needed, it is possible to assign dependencies between knowledge units the ITS teaches. An intervention is more likely to be effective if the learner understands its prerequisites, and less likely if not.

The fact that content has an order lets IMP adapt its behavior despite the large number of knowledge gaps that an ITS must model. To do this, IMP uses a *state queue* representation (Folsom-Kovarik et al. 2010b). Rather than tracking each knowledge feature individually, IMP arranges the features in a queue and models only a learner's current position in the queue. Observations from the real world let IMP move up and down or skip over items in the queue, rather than teaching every item in a fixed order.

A second property of the intelligent tutoring domain is that learner performance can give evidence about multiple knowledge and cognitive states. For example, different answers to the problem "$2 \times (5 + 3)$" could give information about a learner's ability to apply addition, multiplication, and order of operations. This is especially true in an open-ended domain such as a training simulation, where actions are not easily broken into subtasks and every learner action requires the learner to apply several skills at once.

In order to make these highly informative observations tractable in a POMDP learner model, IMP uses an *observation chain* representation (Folsom-Kovarik et al. 2010b). This representation breaks each high-information observation into orthogonal features and transmits them to the POMDP over multiple action-observation cycles.

Finally, a property of many ITS domains is that one learner mistake can have several possible root causes. In order to target its interventions an ITS must decide which knowledge is actually missing or which misconception is present. Inquiry modeling is integral to IMP's accomplishing this common ITS task.

Functionally, IMP runs on a standard laptop concurrently with an existing simulator and overlays the simulator screen with two additional windows: IMP can choose to display text or graphic messages over the screen center, and it always displays a menu of questions the learner can ask at the bottom of the screen. This question user interface (QUI) presents up to 28 questions learners might want to ask during training. In addition to supporting inquiry modeling, the menu contents scaffold effective help seeking (Nelson-Le Gall 1981) in terms of help request content and construction, and IMP can also choose to scaffold help request timing by highlighting a particular QUI item when its model believes the learner should ask that question.

The timing and categorical content of each QUI help request are integrated into IMP's model as a POMDP observation. Other observation types include trainee performance in the simulator and the passage of time with no performance. Since many performance events contain information about multiple model states, these events have over 100,000 different possible values. However, IMP's observation chain representation needs only 48 unique observations to transmit the same information to the POMDP.

The states IMP models represent missing or incorrect learner knowledge and transient cognitive states that affect learning. IMP's knowledge model includes 17 domain-specific gaps that IMP needs to tutor, represented with a state queue. For cognitive states, IMP models four affective states that can make instruction more or less effective (Craig et al. 2004). Boredom, confusion, frustration, and flow are orthogonal to the state queue in a fully enumerated representation, with transitions between states based on prior research (Baker, Rodrigo, and Xolocotzin 2007). The state queue representation lets IMP model all of a learner's knowledge and cognitive features with only 144 POMDP states, well below the current practical maximum.

Finally, IMP has the ability to carry out 30 instructional actions: 17 actions presenting hints, 12 actions highlighting QUI items, and a no-op. Each hint lets IMP target a different knowledge gap. Hint actions and most question answers remind trainees about information they learned before practice, rather than teaching new material.

After defining a POMDP to model the observations, states, and actions in IMP's instructional domain, a policy was found to choose an approximately optimal action given any instructional situation. Because IMP's ITS-specific representations can work with standard POMDP algorithms, an existing algorithm called SARSOP was chosen to search for a policy (Kurniawati, Hsu, and Lee 2008). Policy searches were run for approximately 48 hours on a single core of a 3-GHz processor with 8 GB of RAM before reaching an arbitrary time limit. The resulting policies, one for each experimental condition, were fairly substantial. Each contained between 2,000 and 2,500 α-vectors partitioning belief space into potentially different courses of action depending on the instructional situation, but the policies still executed in real time on the target laptops.

IMP's utility and practicality is not specific to one domain. Inquiry modeling and the structures outlined here could also model other ITS domains such as algebra or computer programming, as long as they display common characteristics such as some order of material presentation and some ability to observe learner performance.

Related Work

Although there is previous work on question-answering in tutoring systems, the utility of incorporating user questions into the learner model has not been well explored. Many ITSs let learners request help with an interface like a hint button (e.g., Koedinger and Aleven 2007). Others update a learner model based on the timing or existence of hint button clicks (e.g., Ainsworth and Grimshaw 2004). However, a single button gives limited information, so these ITSs do not draw evidence from help request content.

Three ITSs give learners more freedom to ask different questions. STEVE (Rickel and Johnson 1999) is an adaptive trainer for machine maintenance that can answer trainee questions about what to do next, how, and why. Autotutor (Graesser et al. 1999) teaches physics with natural-language dialogue. When learners ask Autotutor questions, it searches a text for keywords and presents a list of sections that might be related. Quantum Simulations, Inc. is a commercial enterprise building ITSs that teach subjects like math, chemistry, and accounting (Johnson, Phillips, and Chase 2009). Learners can choose from a question menu while they practice, but the ITS does not attempt to trace learner knowledge. Though these systems give learners more freedom to ask for help, none completely integrates information from help requests into a learner model.

Three existing systems employ POMDPs for tutoring and training applications. Two use POMDP representations that are not specific to ITSs, and they model few facts about learners in order to remain tractable. A coin tutor for elementary students models three binary knowledge features with eight states (Theocharous et al. 2009). A second POMDP that controls lesson selection in a military training setting contains an unknown number of features, but only 11 states (Levchuk, Shebilske, and Freeman in press).

A third system, the Reachable Anytime Planner for Imprecisely sensed Domains (RAPID), uses a domain-specific search algorithm to leverage some of the same features of the ITS problem as IMP (Brunskill and Russell 2010). In simulations, RAPID was able to encompass problems with about as many learner features. However, RAPID assumes stricter restrictions on learner state transitions and observation information. These restrictions make it difficult to apply RAPID in some realistic domains such as the practice scenarios of the present paper, where prerequisite relations between knowledge features are not deterministic and each observation relates to many features.

Finally, there are several approaches to making POMDPs tractable that are not tied to the ITS domain. Some domain-agnostic approaches include state aggregation (Boutilier and Poole 1996), macro-actions (Hauskrecht et al. 1998), hierarchical POMDPs (Theocharous and Mahadevan 2002), and observation factorization (Hoey and Poupart 2005). In the future, such approaches can potentially be combined with ITS-specific representations to achieve even greater efficiency. However, there still remains a need for domain-specific enhancements to take advantage of regularities in real-world ITS problems.

The Call for Fire Training Domain

IMP was evaluated in a realistic military training domain. The task to be trained, known as *call for fire* (CFF), is performed by Forward Observers (FOs). FOs work on the front lines of battle to observe enemy positions and direct attacks from allied units. When FOs *call for fire*, they transmit an enemy's location and description so that distant artillery and other units can target that enemy with precise fire while avoiding harm to friendly units or civilians.

Performance in the CFF domain requires *target selection* and *target engagement*. First, an FO must identify enemy units and prioritize them based on the threat they pose. Second, each CFF the trainee transmits must contain accurate information about the FO's location, one target's relative position, the type of ammunition to use, and the firing pattern to use. The training domain is interesting because errors in a CFF transmission can have different underlying causes, such as incorrectly identifying a unit's type or misremembering the prescribed method to attack that unit.

Besides the domain tasks to be trained, the practice environment used to evaluate IMP also presents interesting technical challenges to a traditional ITS approach. In the United States Marine Corps, FOs currently practice CFF tasks with a laptop-based simulator called the Combined Arms Network (CAN) (Bailey and Armstrong 2002). The CAN does not output fine-grained information that could help diagnose trainees' needs, such as their mouse clicks or progress on partial tasks that contribute to a CFF. Instead, the only times IMP can observe trainee performance are after a successful CFF and the resulting fire. CFFs that are unsuccessful because of malformation are rejected by the CAN with no report to IMP. As a result, IMP's learner model must estimate trainees' needs based on limited performance observations. In such an environment, questions learners may ask during training give valuable extra input.

Finally, because there are only a few moments when IMP receives new information about training progress, there are similarly few opportunities for IMP to attempt new training interventions. IMP needs to use every opportunity wisely and select its interventions effectively.

For these reasons, training personnel to call for fire is a realistic task with implications for current military needs. The CFF domain also presents interesting instructional and technical challenges to making training truly adaptive.

Method

An experiment was conducted to evaluate IMP's ability to train adaptively in the CFF domain. The experiment lasted up to two hours. Participants learned CFF procedures and practiced in the CAN simulator twice. Each practice scenario was preceded by a three-minute video introducing the material the trainees needed to practice. Participants were given pen and paper and instructed to take notes. After practicing, the participants performed in two graded scenarios and answered written questions to test their skill proficiency and declarative knowledge.

The IMP evaluation used a between-subjects design in order to avoid carryover effects. Participants were randomly assigned to one of three experimental conditions. The three groups received the same introductory material and simulator scenarios but differed in the kind of support participants received during practice. Participants in *group O*, the control condition, trained with expository videos only and received no support during practice. Condition O represented the *status quo* for U.S. Marine Corps trainees' use of the CAN for practice. Participants in *group Q* received support from IMP during the 26 minutes of the two simulator practice sessions. Finally, participants in *group P* received practice support from an ablated version of IMP. The ablated IMP maintained a POMDP learner model and provided adaptive practice support, but did not display a QUI and had no inquiry modeling functionality.

Participants ($N = 106$) were recruited from Craig's List and were paid \$10 per hour. All participants were United States citizens at least 18 years old ($\mu = 24.2$, $\sigma = 6.6$); 64% were male and 87% had at least some college education. No participants had military training or experience.

A statistical power analysis projected that 33 participants were required for each experimental condition. Based on experience of a previous study using similar training tasks in the CAN, a prerequisite criterion was established for inclusion in data analysis. Participants were not included if they failed to take any measurable action during both practice sessions. Participants who displayed any measurable performance in either practice session were retained in the analysis, even if they did not perform in the test sessions. Data analysis finally included 106 participants: 35 in the control condition (O), 37 in the +POMDP condition (P), and 34 in the +POMDP +QUI condition (Q).

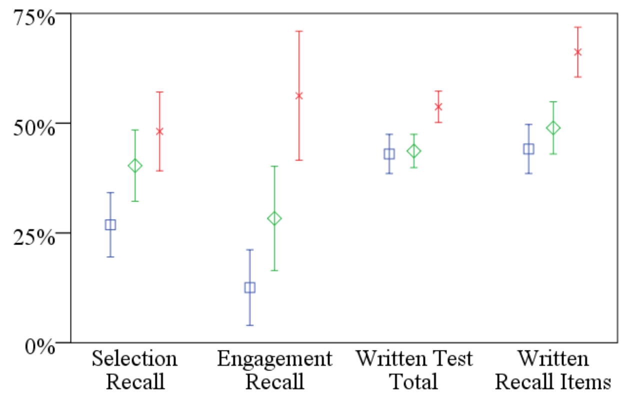

Figure 1. Four learning outcomes showing significant improvement. Each column of three results compares control (□), POMDP (◊), and full IMP (×) conditions. Bars show 95% CI.

Results and Discussion

Two test scenarios measured how well participants demonstrated target skills in a performance environment. The first test evaluated proficiency in the same scenario as the participants had practiced, while the second test evaluated proficiency in a new scenario that required participants to adapt the behaviors they had been practicing. The novel scenario challenged participants to try to measure deeper understanding of the skills they had learned and practiced.

First, if IMP's actions interfered with trainees' ability to practice, the experimental conditions would have differed in number of calls for fire transmitted or enemy targets destroyed. However, no significant differences appeared in these measures. Therefore, IMP's interventions did not significantly interrupt practice or otherwise decrease trainees' automaticity and speed in calling for fire.

Next, for both test scenarios, participants' target selection and target engagement were scored. Shapiro-Wilk tests for normality showed these scores were not normally distributed, so they were analyzed with non-parametric Kruskal-Wallis one-way ANOVAs. Both target selection and target engagement differed significantly in the recall proficiency test (Figure 1), although scores did not differ on the test requiring participants to create new behaviors.

On the recall proficiency test, significant differences were found between conditions in target engagement scores ($H = 22.93$, d.f. $= 2$, $p < 0.001$). Post-hoc tests for least significant difference of mean ranks showed significant differences between all conditions. Participants in the P condition outperformed control participants by 0.64 σ, and in turn participants in the Q condition outperformed P participants by 0.82 σ. Participants in the Q condition also outperformed control participants by 1.77 σ. This measure represented a particularly large difference between conditions: IMP's trainees scored 4.5 times higher than control participants and twice as high as those who used ablated IMP with no inquiry modeling functionality.

Differences were also found between conditions in target selection scores ($H = 12.91$, d f. = 2, $p < 0.01$). Post-hoc tests showed that participants in the P and Q conditions performed significantly better than the control condition. Trainees who used ablated IMP outscored the control participants on average by 0.64 σ, and with full IMP, the improvement was 0.94 σ.

After the test scenarios, a written test of declarative knowledge measured how well participants were able to recall and apply knowledge outside of a simulator setting. Outcome differences were checked with parametric linear regression. Participants in different conditions accomplished significantly different outcomes on the written test ($F = 9.494$, d.f. = (2, 103), $p < 0.001$). Post hoc tests using Tukey's honestly significant difference method showed that IMP's trainees significantly outscored control participants by 0.83 σ and trainees using ablated IMP by 0.85 σ.

Like the proficiency tests, the written measure showed a trend that practice with IMP produced large learning improvements on the material it could choose to tutor, although not on new material that IMP's hints did not directly cover. Support during practice strongly affected direct recall test items ($F = 17.586$, d f. = (2, 103), $p < 0.001$), but not extension items that required knowledge application in new situations ($F = 1.439$, d f. = (2, 103), *n.s.*).

Next, in order to give insight into how IMP produced its positive effects on participants' learning, IMP's interventions for participants in the P and Q conditions were examined for patterns. If IMP repeated the same sequences for many participants, it would indicate IMP did not adapt to individuals' needs. Instead, IMP's intervention choices branched quickly. After its initial action, IMP successfully moved on to more remedial or more advanced interventions based on its observations of individual trainees. As Figure 2 shows, the intervention sequences IMP chose were unique for almost 90% of participants after only five turns. IMP's adaptive behaviors were responsive to the performance and the training needs of each participant.

Finally, 60% of actions IMP recommended during practice did not directly address the estimated most-probable gap. Considering recommendations IMP made immediately after observing one or more trainee errors, 51% did not directly address any of the errors just observed. These two patterns support the assertion that IMP chose interventions to effect a longer-term teaching plan, rather than applying a more obvious reactive or greedy intervention strategy.

In summary, the analysis and outcomes suggest that IMP's POMDP learner model successfully interpreted sparse simulator performance observations and then, rather than reacting to the latest observation or helping everyone the same way, used planning to intervene in ways that helped each individual practice effectively and learn more.

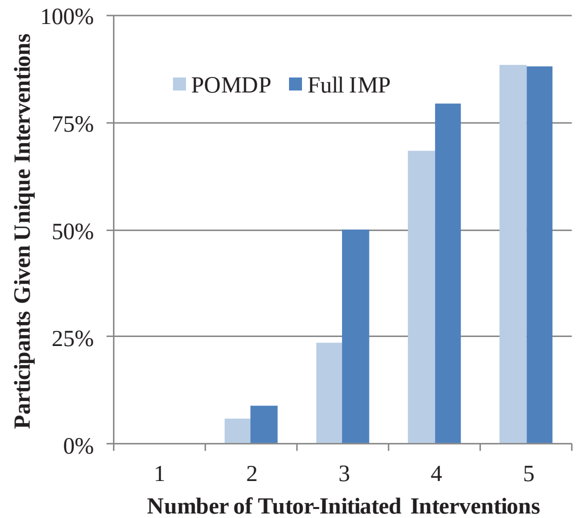

Figure 2. Both the ablated tutor with a POMDP learner model and full IMP, using a POMDP and inquiry modeling, successfully adapted interventions to individual trainees' needs.

Conclusions and Future Work

This experiment demonstrated that inquiry modeling and planning with POMDPs can yield significant and substantive learning improvements in a realistic training task.

The new, ITS-specific representations in IMP produced these learning improvements despite discarding some information to remain tractable. Going forward, the improved scaling of the state queue and observation chain representations will let POMDP ITSs model more information or more detailed information about each learner.

Inquiry modeling let IMP produce its best learning outcomes. The question of how best to integrate information from learner help requests remains a topic for ongoing research. For example, questions learners asked usually let IMP present advanced material earlier. However, there were also cases when IMP incorrectly assumed learners knew material after they had asked about it, leading it to skip basic corrections the learners needed. In particular, one pattern of asking many questions seemed to indicate floundering and not evidence of improved knowledge.

One path for future work might include updating IMP's POMDP evidence interpretation based on the actual learners' experiences gathered in this experiment. The ability to improve a model with general machine-learning algorithms is an advantage of parametric models such as POMDPs.

The ITS-specific representations IMP uses also suggest another interesting possibility. The representations' simplifying independence assumptions have the potential to make the POMDP's decisions and design more transparent than might otherwise be true. There seems to be an opportunity to open up the model to instructors and other endusers for their input and understanding. One could imagine individual instructors without programming skills might be able to correct IMP's wrong belief, or tweak IMP's instruction to focus on a specific learner population, or ask IMP for model information about particular learners so as to help them better with other training. Such changes would continue to enhance the inquiry modeling POMDP ITS and make it more usable in practical teaching domains.

Acknowledgement

This work is supported in part by National Science Foundation award IIS-0845159 and by Office of Naval Research grant N0001408C0186. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NSF, ONR, or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

Ainsworth, S. and Grimshaw, S. 2004. Evaluating the REDEEM authoring tool: Can teachers create effective learning environments? International Journal of Artificial Intelligence in Education 14(3): 279–312.

Bailey, M.P. and Armstrong, R. 2002. The Deployable Virtual Training Environment. In Proceedings of the Interservice/Industry Training, Simulation, and Education Conference: National Training and Simulation Association.

Baker, R.S.J.d., Rodrigo, M.M.T., and Xolocotzin, U.E. 2007. The Dynamics of Affective Transitions in Simulation Problem-Solving Environments. In Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction, 666–677. Berlin / Heidelberg: Springer.

Bloom, B.S. 1984. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher 13(6): 4–16.

Boutilier, C. and Poole, D. 1996. Computing optimal policies for partially observable decision processes using compact representations. In Proceedings of the 13th National Conference on Artificial Intelligence, 1168-1175. Menlo Park, CA: AAAI Press / MIT Press.

Brunskill, E. and Russell, S. 2010. RAPID: A Reachable Anytime Planner for Imprecisely sensed Domains. In Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence, 83–92. Arlington, VA: AUAI Press.

Craig, S.D., Graesser, A.C., Sullins, J., and Gholson, B. 2004. Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. Journal of Educational Media 29: 241–250.

Folsom-Kovarik, J.T., Schatz, S., Sukthankar, G., and Nicholson, D. 2010a. What Information Does this Question Convey? Leveraging Help-Seeking Behavior for Improved Modeling in a Simulation-Based Intelligent Tutor. In Proceedings of the 2010 Spring Simulation Multiconference, Military Modeling and Simulation Symposium. Vista, CA: SCS / ACM.

Folsom-Kovarik, J.T., Sukthankar, G., Schatz, S., and Nicholson, D. 2010b. Scalable POMDPs for Diagnosis and Planning in Intelligent Tutoring Systems, FS-10-07, Association for the Advancement of Artificial Intelligence, Arlington, VA.

Graesser, A.C. and Person, N.K. 1994. Question asking during tutoring. American Educational Research Journal 31(1): 104–137.

Graesser, A.C., Wiemer-Hastings, K., Wiemer-Hastings, P., and Kreuz, R. 1999. AutoTutor: A simulation of a human tutor. Journal of Cognitive Systems Research 1: 35–51.

Harper, K.A., Etkina, E., and Lin, Y. 2003. Encouraging and analyzing student questions in a large physics course: Meaningful patterns for instructors. Journal of Research in Science Teaching 40(8): 776–791.

Hauskrecht, M., Meuleau, N., Boutilier, C., Kaelbling, L.P., and Dean, T. 1998. Hierarchical Solution of Markov Decision Processes Using Macro-Actions. In Proceedings of the 14th International Conference on Uncertainty In Artificial Intelligence. San Francisco, CA: Morgan Kaufmann Publishers.

Hoey, J. and Poupart, P. 2005. Solving POMDPs with Continuous or Large Discrete Observation Spaces. In Proceedings of the 19th International Joint Conference on Artificial Intelligence: Professional Book Center.

Johnson, B.G., Phillips, F., and Chase, L.G. 2009. An intelligent tutoring system for the accounting cycle: Enhancing textbook homework with artificial intelligence. Journal of Accounting Education: 30–39.

Kaelbling, L.P., Littman, M.L., and Cassandra, A.R. 1998. Planning and acting in partially observable stochastic domains. Artificial Intelligence 101(1–2): 99–134.

Koedinger, K.R. and Aleven, V. 2007. Exploring the assistance dilemma in experiments with Cognitive Tutors. Educational Psychology Review 19(3): 239–264.

Koedinger, K.R., Anderson, J.R., Hadley, W.H., and Mark, M.A. 1997. Intelligent tutoring goes to school in the big city. International Journal of Artificial Intelligence in Education 8(1): 30–47.

Kurniawati, H., Hsu, D., and Lee, W.S. 2008. SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces. In Proceedings of the 4th Robotics: Science and Systems Conference, 65–72. Cambridge / London: MIT Press.

Levchuk, G., Shebilske, W., and Freeman, J. in press. A model-driven instructional strategy: The benchmarked experiential system for training (BEST). In P. J. Durlach & A. M. Lesgold (eds.), Adaptive Technologies for Training and Education. Cambridge, UK: Cambridge University Press.

Nelson-Le Gall, S. 1981. Help-seeking: An understudied problem-solving skill in children. Developmental Review 1(3): 224–246.

Payne, S.J. and Squibb, H.R. 1990. Algebra mal-rules and cognitive accounts of error. Cognitive Science 14(3): 445–481.

Rickel, J. and Johnson, W.L. 1999. Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. Applied Artificial Intelligence 13(4): 343–382.

Theocharous, G., Beckwith, R., Butko, N., and Philipose, M. 2009. Tractable POMDP Planning Algorithms for Optimal Teaching in "SPAIS" In Proceedings of the 21st International Joint Conferences on Artificial Intelligence workshop on Plan, Activity, and Intent Recognition. Menlo Park, CA: AAAI Press.

Theocharous, G. and Mahadevan, S. 2002. Approximate planning with hierarchical partially observable Markov decision process models for robot navigation. In Proceedings of the IEEE International Conference on Robotics and Automation, 1347–1352.