Improving Student Problem Solving in Narrative-Centered Learning Environments: A Modular Reinforcement Learning Framework

Jonathan P. Rowe and James C. Lester

Center for Educational Informatics, North Carolina State University, Raleigh, NC 27695
{jprowe, lester}@ncsu.edu

Abstract. Narrative-centered learning environments comprise a class of game-based learning environments that embed problem solving in interactive stories. A key challenge posed by narrative-centered learning is dynamically tailoring story events to enhance student learning. In this paper, we investigate the impact of a data-driven tutorial planner on students' learning processes in a narrative-centered learning environment, CRYSTAL ISLAND. We induce the tutorial planner by employing *modular reinforcement learning*, a multi-goal extension of classical reinforcement learning. To train the planner, we collected a corpus from 453 middle school students who used CRYSTAL ISLAND in their classrooms. Afterward, we investigated the induced planner's impact in a follow-up experiment with another 75 students. The study revealed that the induced planner improved students' problem-solving processes—including hypothesis testing and information gathering behaviors—compared to a control condition, suggesting that modular reinforcement learning is an effective approach for tutorial planning in narrative-centered learning environments.

Keywords: Narrative-Centered Learning Environments, Tutorial Planning, Modular Reinforcement Learning, Game-Based Learning.

1 Introduction

Over the past decade, the education research community has shown growing interest in digital games, largely inspired by a key question: how can we motivate and engage students in learning? A promising class of games is narrative-centered learning environments, which integrate the motivational qualities of stories, along with the adaptive pedagogy of intelligent tutoring systems, to foster student engagement in learning and problem solving. When students use narrative-centered learning environments, they become active participants in ongoing narratives whose outcomes are shaped by students' learning behaviors. As a result of recent advances in game engines and authoring tools, there are now a range of narrative-centered learning environments under investigation across different domains, including language learning [1], anti-bullying education [2], biosafety training [3], and science inquiry [4].

A key benefit of narrative-centered learning environments is their capacity to discreetly support students' learning processes by integrating pedagogical and narrative elements. This form of scaffolding depends upon the presentation of events

adfa, p. 1, 2011.

that fulfill dual roles: advancing problem-centric storylines, and providing tutorial support such as feedback or hints. Yet, despite a substantial research base on the cognitive principles of student learning [5], there is limited research on how to effectively design narrative-centered learning environments. If designed or deployed ineffectively, narrative-centered learning environments risk the introduction of seductive details, which can be harmful for learning [6]. Moreover, a one-size-fits-all approach to the design of narrative-centered learning environments has important limitations, due to the role of students' individual differences in learning.

To address these challenges, we conceptualize adaptive scaffolding in narrative-centered learning environments as an instance of tutorial planning. We seek to devise computational models for generating, sequencing, and personalizing story events in a narrative-centered learning environment, with the explicit aim of enhancing student learning and engagement. To solve this problem, we employ a data-driven framework for inducing *narrative-centered tutorial planners* that leverages modular reinforcement learning. This formulation is made possible by the observation that tutorial planning in narrative-centered learning environments can be decomposed in terms of multiple independent sub-problems, each focused on a particular class of scaffolding events. Our framework is inspired by work on reinforcement learning methods for tutorial dialogue management [7], adapting and extending these techniques to meet the requirements of narrative-centered learning.

To evaluate our framework, we present results from an experiment investigating the impact of an induced tutorial planner integrated with the CRYSTAL ISLAND narrative-centered learning environment. Empirical findings indicate that the induced planner improves students' problem-solving behaviors, including hypothesis-testing and information-gathering processes, compared to a control condition. The results suggest that our modular reinforcement-learning framework is a promising method for devising data-driven tutorial planners that scaffold learning effectively in narrative-centered learning environments.

2 Tutorial Planning with Modular Reinforcement Learning

We formalize tutorial planning as a modular reinforcement learning problem. Modular reinforcement learning is a multi-goal extension of classical single-agent reinforcement learning [8,9]. In reinforcement learning, an agent learns a policy for selecting actions in an uncertain environment, guided by delayed rewards, in order to accomplish a goal [10]. The agent utilizes an environment-based reward signal in order to learn a policy, denoted π, which maps observed states to actions and maximizes total accumulated reward. Agents in reinforcement learning problems are typically modeled with Markov decision processes (MDPs).

Modular reinforcement learning tasks are formally defined in terms of N concurrent MDPs, $M = \{M_i\}_1^N$, where each $M_i = (S_i, A_i, P_i, R_i)$, corresponding to a sub-problem in the composite reinforcement learning task. Each agent M_i has its own state sub-space S_i, action set A_i, probabilistic state transition model P_i, and reward model R_i. The solution to a modular reinforcement learning problem is a set of N policies, $\pi^* = \{\pi_i^*\}_1^N$, where π_i is the optimal policy for the constituent MDP M_i. Any circumstance where two policies π_i and π_j with $i \neq j$ recommend different actions in the same state requires the application of an arbitration procedure.

Tutorial planning in narrative-centered learning environments is naturally represented as a modular reinforcement learning problem: *state* consists of the learner's state and history as well as the learning environment's; *actions* represent the pedagogical decisions the planner can perform; a *probabilistic state transition model* encodes how learners, and the learning environment, respond to the planner's tutorial decisions; and a *reward model* encapsulates measures of students' learning outcomes, which the tutorial planner seeks to optimize. The solution to a modular reinforcement-learning problem is a set of *policies*, or mappings between states and tutorial actions, that govern how the tutorial planner scaffolds students' learning. If two policies conflict, externally defined arbitration procedures specify which policy prevails.

By decomposing tutorial planning into multiple sub-problems, we can reduce the complexity of reinforcement learning by reframing the task in terms of several smaller, concurrent Markov decision processes. To perform this decomposition, we employ the concept of an *adaptable event sequence* (AES), an abstraction for a series of one or more scaffolding-related events that, once triggered, can unfold in several different ways within the learning environment [11]. To illustrate the concept of an AES, consider an example of an event sequence that occurs when a player asks a non-player character (NPC) about her backstory. The NPC could respond in one of several ways: 1) providing a detailed explanation and a hint about how her backstory information is useful, 2) providing an explanation but no hint, 3) responding suspiciously and revealing only a few details, or 4) not responding at all. Each of these four responses is an alternate manifestation of the *NPC Backstory* event sequence. Each option is coherent within the storyline, can be interchanged with any other, and provides a distinct level of problem-solving support. We refer to the event sequence as *adaptable*, or in other words, it is an adaptable event sequence (AES).

AESs can encode a broad range of scaffolding types. For example, an AES could specify the location of an important object, or determine what level of hint to provide to a student, or select whether to prompt a student to self-explain their problem-solving strategy or not. Further, multiple AESs can be interleaved. AESs encode distinct threads of story events, each potentially involving multiple decision points spanning an entire story. For this reason, AESs are sequential and operate concurrently. Each AES is modeled separately as a MDP, and tutorial decisions about scaffolding are determined through modular reinforcement learning.

Leveraging the concept of an AES, narrative-centered tutorial planning can be cast as a collection of sequential decision-making problems about scaffolding student learning within a narrative-centered learning environment. Modular reinforcement learning is applied as follows. Each AES is modeled as a distinct Markov decision process, M_i. For each AES, every occurrence of the event sequence corresponds to a decision point for M_i. The set of possible scaffolding options for the AES is modeled by an action set, A_i. A particular state representation, S_i, is tailored to the AES using manual or automatic feature selection techniques. Rewards, R_i, can be calculated from formative or summative assessments of student learning, such as a post-test. A state transition model P_i encodes the probability of transitioning between two specific states during successive decision points for the AES. To estimate the values of these parameters, we can collect training data from students by deploying a tutorial planner that selects actions randomly, in effect sampling the space of tutorial policies and rewards [7]. Leveraging this mapping between AESs and MDPs, and a training corpus of random tutorial decision data, we can employ model-based reinforcement learning techniques to induce policies for tutorial planning. Specifically, we utilize dynamic programming methods (e.g., value iteration) to compute solution policies for each MDP using estimates of the state transition model and reward model inferred from the training corpus [7,10]. In cases where two policies conflict, we utilize *greatest mass arbitration*, a domain-independent arbitration procedure that selects the action with the largest Q-value calculated during policy induction [8,9]. In combination, this formulation provides a method for formulating narrative-centered tutorial planning as an instance of modular reinforcement learning.

3 Corpus Collection

To investigate our modular reinforcement learning framework for tutorial planning, we used CRYSTAL ISLAND, a narrative-centered learning environment for middle school microbiology (Figure 1). The version of CRYSTAL ISLAND used in this study was built on Valve Software's Source™ engine. The environment features a science mystery in which students investigate the identity and source of an infectious disease that is plaguing a research team on a remote island. Students adopt the role of a medical detective who must save the research team from the outbreak. Over the past decade, CRYSTAL ISLAND has been the subject of extensive empirical investigation, and has been found to provide substantial learning and motivational benefits [12].

Figure 1. CRYSTAL ISLAND narrative-centered learning environment.

To investigate narrative-centered tutorial planning in CRYSTAL ISLAND, we developed a modified version of the system that includes 13 AESs. We selected 13 AESs in order to incorporate a broad range of scaffolding capabilities. Space limitations preclude a detailed description of every AES, but they included decisions about whether to provide hints and prompts during the mystery (e.g., prompt the student to record her findings in a diagnosis worksheet, prompt the student to self-explain her problem-solving strategy); whether to administer embedded assessments of content knowledge; which disease and transmission source caused the outbreak; how much detail should NPCs provide about their symptoms; what level of feedback should be provided on students' proposed diagnoses; and manipulations to the number of hypotheses that students can test in the virtual laboratory. For additional details, a more comprehensive discussion of the AESs is available in [13].

To illustrate how AESs unfolded within CRYSTAL ISLAND, consider the following scenario. When a student begins the narrative, the *Mystery's Solution* AES occurs behind the scenes, selecting one of 6 possible solutions to the mystery. The tutorial planner selects *salmonellosis* as the mystery disease and *contaminated milk* as the disease's transmission source. This AES decision is invisible to the student, but the selection dictates which symptoms and medical history are reported by the sick characters. As the student explores the camp, she initiates a conversation with a sick scientist named Teresa. The student asks Teresa about her symptoms, triggering a decision point for the *Details of Teresa's Symptoms* AES. This AES controls how much information Teresa provides in her response. The tutorial planner has Teresa provide minimal information, leading Teresa to groan and explain that she has a fever. If the student chooses to ask Teresa about her symptoms again later, the planner may choose a different response to help the student narrow on a diagnosis. Next, a decision point for the *Record Findings Reminder* AES is triggered, because the student has just received useful information for diagnosing the illness. The tutorial planner chooses whether to hint to the student that she should take a note about the symptom information. The narrative continues in this manner, driven by the student's actions, and periodically triggering scaffolding events that shape how the experience unfolds.

After modifying CRYSTAL ISLAND to incorporate AESs, we conducted a pair of classroom studies to collect training data for inducing a tutorial planner. The first study involved 300 students from a North Carolina middle school, and the second study involved 153 students from another middle school. All students used the same version of CRYSTAL Island, followed the same study procedure, and used the game individually. One week prior to using CRYSTAL ISLAND, students completed a pre-test, which collected data on students' demographics, game-playing experience, and microbiology content knowledge. The microbiology content test consisted of 19 multiple-choice questions, and was created iteratively by the research team and a group of eighth-grade science teachers. During the studies, students interacted with CRYSTAL ISLAND until they solved the mystery, or 55 minutes elapsed, whichever occurred first. Immediately afterward, students completed a post-test, which included the same content knowledge assessment as the pre-test, as well as several self-report measures of engagement. Both the pre- and post-tests lasted no more than 30 minutes.

While using CRYSTAL ISLAND, students unknowingly encountered AESs several times. At each AES decision point, the environment selected a scaffolding-related event according to a uniform random policy. By logging these tutorial planning decisions, as well as students' responses, the environment broadly sampled the space of policies for controlling adaptable event sequences. The data from both studies were combined into a single corpus consisting of two parts: students' interaction logs, and students' pre- and post-test results. After removing incomplete or inconsistent records, there were 402 participants remaining. The resulting data consisted of 315,407 events. In addition to student actions, there were 10,057 instances of AESs in the corpus, which corresponded to approximately 25 tutorial planning decisions per student.

4 Implemented Planner

Using the corpus, we induced a policy for each MDP to control CRYSTAL ISLAND's scaffolding features, with the exception of one AES for which we had insufficient training data (off-task behavior discouragement). All of the MDPs shared the same state representation, which consisted of 8 binary features drawn from three categories: narrative features, individual difference features, and problem-solving features. We limited the state representation to 8 binary features to mitigate potential data sparsity issues. The first four features were narrative-focused. Each feature was associated with a salient plot point from CRYSTAL ISLAND's narrative and indicated whether the plot point had been completed thus far. The next two features were based on students' individual differences. The first feature was computed from a median split on students' microbiology pre-test scores, and the second feature was computed from a median split on students' self-report data about how often they played video games. The final two state features were computed from students' observed problem-solving behaviors. Specifically, we computed running median splits on the frequency of students' lab-testing and book-reading behaviors within CRYSTAL ISLAND.

The action sets for the 12 MDPs corresponded to the scaffolding options for the associated AESs. The action sets' cardinalities ranged from binary to 6-way decisions. If the entire planning task were modeled as a single MDP, it would require encoding approximately 1,644,000 parameters to populate the entire state transition model (256 states × 25 distinct actions × 257 states, including the terminal state), although not all state transitions were possible.[1]

Each MDP shared the same reward function, which was based on students' normalized learning gains (NLG). NLG is the normalized difference between participants' pre- and post-study knowledge test scores. To determine reward values in the corpus, NLG was first calculated for each participant, and then a median split was performed. Students who had a NLG that was greater than or equal to the median were awarded +100 points at the conclusions of their episodes. Participants with a NLG that was less than the median were awarded -100 points.

To induce the tutorial policies, we used *value iteration* [10]. The 12 MDPs, one for each AES in CRYSTAL ISLAND, were implemented with a reinforcement-learning library written in Python by the first author. Policies were induced using a discount rate of 0.9. The discount rate parameter governs how rewards are attributed to planner actions during reinforcement learning. Our previous work has found that discount rate has a limited effect on the policies induced for CRYSTAL ISLAND [11].

5 Evaluation Experiment

After inducing tutorial planning policies for each adaptable event sequence, we evaluated the tutorial planner's impact on students' learning experiences in the run-time CRYSTAL ISLAND learning environment. This required incorporating the induced tutorial planning policies into CRYSTAL ISLAND by replacing the exploratory tutorial policies from the corpus collection studies with the newly induced policies.

To evaluate CRYSTAL ISLAND's induced tutorial planner, we conducted a follow-up controlled experiment with middle school students comparing the induced policies to

[1] Several AESs included an action choice of *do nothing*. We count all of these *do nothing* choices as a single action, yielding a total of 25 distinct actions across the 12 AESs.

a control condition. Participants were drawn from a different school than the corpus collection studies. A total of 75 eighth-grade students participated. Among these students, 14 were removed due to incomplete or inconsistent data.

The study had two conditions: an Induced Planner condition and a Control Planner condition. Students in both conditions played CRYSTAL ISLAND, but the conditions differed in terms of the tutorial planning policies employed by the narrative-centered learning environment. The Induced Planner followed policies obtained by inducing solution policies for each Markov decision process associated with an AES in CRYSTAL ISLAND, with conflicts resolved via greatest mass arbitration [9]. The Control Planner employed a uniform random policy, where tutorial decisions were selected randomly whenever the planner encountered a decision point. This was the same policy used by the exploratory planner during the corpus collection studies.

Students were randomly assigned to the two conditions when they entered the experiment room. Among students with complete data, 33 were randomly assigned to the Induced Planner condition, and 28 were assigned to the Control Planner condition. Students played until they solved the mystery or the interaction time expired, whichever occurred first. The study procedure, pre-test, and post-test were otherwise identical to the corpus collection studies.

6 Results

Analyses of students' learning gains found students achieved significant improvements in microbiology content knowledge in both experimental conditions. In the Induced Planner condition, students significantly improved their content test scores by 1.6 questions on average from pre-test (M = 7.8, SD = 2.2) to post-test (M = 9.4, SD = 3.6), $t(32) = 2.67$, $p < .02$. In the Control Planner condition, students also achieved significant improvements in content test score from pre-test (M = 7.2, SD = 2.5) to post-test (M = 9.5, SD = 3.4), $t(27) = 4.09$, $p < .001$, a gain of 2.3 questions on average. A comparison between the two conditions' average post-test scores did not find evidence of a significant condition effect on microbiology content learning. Similarly, no condition effects were observed on students' normalized learning gains or self-reported engagement. In hindsight, the lack of a condition effect on learning is unsurprising. A majority of the AESs provided scaffolding for students' inquiry behaviors, rather than microbiology content exposure, which was the focus of the pre- and post-tests. Students in both conditions had the same access to the game's microbiology content. Additionally, we had anticipated a potential *test effect* from the *Knowledge Quiz* AES, which controlled decisions about whether to administer embedded assessments in CRYSTAL ISLAND, and would be hypothesized to yield increased learning gains [5]. However, the Induced Planner tended to not deliver the assessments, surprisingly, making it unlikely to find such an effect.

Next, we investigated students' problem-solving processes in CRYSTAL ISLAND. In particular, we sought evidence of deliberate problem solving, in contrast to strategies that involve extensive guessing or non-purposeful behavior. To perform this investigation, we calculated several metrics that had previously yielded insights about problem solving in CRYSTAL ISLAND, including measures of hypothesis testing efficiency [14] and early information gathering behavior [15].

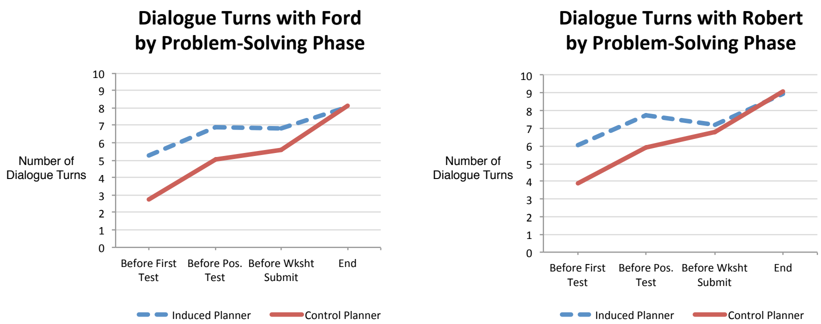

Figure 2. Students' dialogue behaviors by problem-solving phase, which include 1) before running the first lab test, 2) before running a positive lab test, 3) before first submitting the diagnosis worksheet, and 4) before solving the science mystery.

We first analyzed students' hypothesis testing behaviors. In CRYSTAL ISLAND, students test hypotheses about potential sources of the outbreak in the camp's virtual laboratory. A two-tailed t-test indicated that students in the Induced Planner condition (M = 13.7, SD = 10.9) conducted marginally fewer tests than students in the Control Planner condition (M = 19.5, SD = 14.4), $t(59) = 1.80$, $p < .08$. Additionally, the Induced Planner group ran significantly fewer tests (M = 4.7, SD = 7.7) after identifying the transmission source than students in the Control Planner group (M = 11.0, SD = 12.6), $t(59) = 2.39$, $p < .03$. These findings suggest that students in the Induced Planner condition tested their hypotheses more efficiently.

Next, we examined student behaviors during the early stages of problem solving by investigating how students collected background information on the microbiology curriculum prior to forming, testing, and reporting their hypotheses. First, we investigated students' conversations with virtual characters (Figure 2). In terms of total number of conversations, as well as total number of dialogue turns, no significant differences between conditions were observed. However, there were significant condition effects on student dialogue behavior with specific virtual characters. Students in the Induced Planner condition engaged in significantly more dialogue turns with Ford, the camp's virus specialist, prior to running a laboratory test, $t(59) = -2.31$, $p < .03$. Conversely, Induced Planner students engaged in significantly fewer dialogue turns with Ford after running their first test, $t(59) = 2.25$, $p < .03$.

Similar patterns were observed for students' conversational behaviors with Robert, the camp's bacteria specialist. Students in the Induced Planner condition engaged in more dialogue turns with Robert prior to running a laboratory test, $t(59) = -1.71$, $p = .09$. Students in the Induced Planner condition also engaged in fewer dialogue turns with Robert after running their first laboratory test, $t(59) = 2.21$, $p < .04$. In a related finding, the Induced Planner students engaged in significantly more dialogue turns with the camp nurse Kim—a character who provides general background on pathogens, mutagens, and carcinogens—prior to first submitting their diagnosis worksheet, $t(59) = -2.19$, $p < .04$. These patterns are consistent with strategic information gathering in CRYSTAL ISLAND. The findings suggest that students in the Induced Planner condition collected more background information about microbiology prior to testing their hypotheses in the laboratory, which is a desirable problem-solving approach, whereas students in the Control condition gathered background information afterward, which is consistent with an ad hoc approach.

As a further investigation of students' information gathering strategies, we examined poster-viewing behaviors between experimental conditions. In these analyses, we only considered instances lasting longer than one second in duration. Similar to the character dialogue findings, no significant differences in total poster viewing metrics were observed. However, in an examination of the camp's six disease-focused posters, two-tailed t-tests indicated that students in the Induced Planner condition spent significantly more time reading the Salmonellosis poster prior to submitting their diagnosis worksheet than students in the Control Planner condition, $t(59) = -2.18$, $p < .04$. Similarly, students in the Induced Planner condition viewed the Anthrax poster more times prior to submitting their diagnosis worksheet, $t(59) = -1.67$, $p = .1$. Students in the Induced Planner condition viewed the Botulism poster more times prior to successfully testing the transmission source in the laboratory, $t(59) = -1.73$, $p < .09$. And students in the Induced Planner condition viewed the Ebola poster more times prior to submitting their diagnosis worksheet, $t(59) = -1.96$, $p = .05$. No analogous condition effects were observed for the Influenza or Smallpox posters.

These findings suggest that students in the Induced Planner condition examined disease-specific posters more frequently before testing hypothesized diagnoses, particularly for posters about bacterial diseases. The findings raise questions about whether similar patterns were observed for students reading virtual books, which provide similar information for diagnosing the illness. However, an investigation of virtual book-reading behaviors failed to find evidence of significant condition effects. Furthermore, significant condition effects were not observed for students' diagnosis worksheet behaviors, another key problem-solving feature in CRYSTAL ISLAND.

7 Conclusions and Future Work

We have found that a narrative-centered tutorial planner, induced using modular reinforcement learning, significantly improves students' problem-solving processes in the CRYSTAL ISLAND learning environment. We trained the tutorial planner directly upon a corpus of data from students who used CRYSTAL ISLAND in their science classrooms, producing data-driven tutorial planning models capable of adaptive scaffolding. We evaluated the planner's impact in a controlled experiment conducted with 75 middle school students. Results indicated that students in the Induced Planner condition demonstrated greater efficiency at hypothesis testing, as well as greater evidence of strategic information gathering, during their investigations. These findings provide evidence that narrative-centered tutorial planners, induced using modular reinforcement learning, can have positive effects on students' problem solving behaviors. Building on these findings, in future work it will be important to investigate the impacts of alternate MDP state representations on induced tutorial planning policies. In addition, it will be informative to investigate the framework's generalizability by applying it to different types of learning environments.

Acknowledgments. This material is based upon work supported by the National Science Foundation under grants IIS-1344803, REC-0632450, IIS-0812291, and DRL-0822200. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Johnson, W. L.: Serious Use of a Serious Game for Language Learning. International Journal of Artificial Intelligence in Education, 20(2), 175–195 (2010)
2. Vannini, N., Enz, S., Sapouna, M., Wolke, D., Watson, S., Woods, S., Dautenhahn, K., Hall, L., Paiva, A., Andre, E., Aylett, R., Schneider, W.: "FearNot!": A Computer-Based Anti-Bullying-Programme Designed to Foster Peer Intervention. European Journal of Psychology of Education, 26(1), 21–44 (2011)
3. Alvarez, N., Sanchez-Ruiz, A., Cavazza, M., Shigematsu, M., Prendinger, H.: Narrative Balance Management in an Intelligent Biosafety Training Application for Improving User Performance. Intl Journal of Artificial Intelligence in Education, 25(1), 35–59 (2015)
4. Nelson, B. C., Kim, Y., Foshee, C., Slack, K.: Visual Signaling in Virtual World-Based Assessments: The SAVE Science Project. Information Sciences, 264, 32–40 (2014)
5. Graesser, A., Halpern, D., Hakel, M.: *25 Principles of Learning*. Task Force on Lifelong Learning at Work and at Home. Washington, DC (2008)
6. Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., Wainess, R.: Narrative Games for Learning: Testing the Discovery and Narrative Hypotheses. Journal of Educational Psychology, 104(1), 235–249 (2012)
7. Chi, M., Vanlehn, K., Litman, D., Jordan, P.: Empirically Evaluating the Application of Reinforcement Learning to the Induction of Effective and Adaptive Pedagogical Strategies. User Modeling and User-Adapted Interaction, 21, 137–180 (2011)
8. Bhat, S., Isbell, C. L., Mateas, M.: On the Difficulty of Modular Reinforcement Learning for Real-World Partial Programming. In: Proceedings of the 21st National Conference on Artificial Intelligence, pp. 318–323. AAAI Press, Menlo Park, CA (2006)
9. Karlsson, J.: Learning to Solve Multiple Goals. Ph.D. diss., Dept. of Comp. Sci., University of Rochester (1997)
10. Sutton, R., Barto, A.: Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA (1998)
11. Rowe, J., Mott, B., Lester, J.: Optimizing Player Experience in Interactive Narrative Planning: A Modular Reinforcement Learning Approach. In: Proceedings of the 10th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, pp. 160–166. AAAI Press, Menlo Park, CA (2014)
12. Lester, J., Ha, E., Lee, S., Mott, B., Rowe, J., Sabourin, J.: Serious Games Get Smart: Intelligent Game-Based Learning Environments. AI Magazine, 34(4), 31–45 (2013)
13. Rowe, J.: *Narrative-Centered Tutorial Planning with Concurrent Markov Decision Processes*. Ph.D. diss., Dept. of Comp. Sci., North Carolina State University (2013).
14. Spires, H., Rowe, J., Mott, B., Lester, J.: Problem Solving and Game-Based Learning: Effects of Middle Grade Students' Hypothesis Testing Strategies on Science Learning Outcomes. Journal of Educational Computing Research, 44(4), 453–472 (2011)
15. Sabourin, J., Rowe, J., Mott, B., Lester, J.: Exploring Inquiry-Based Problem-Solving Strategies in Game-Based Learning Environments. In: Proceedings of the 11th International Conference on Intelligent Tutoring Systems, pp. 470–475. Springer, Berlin (2012)