

A Modular Reinforcement Learning Framework for Interactive Narrative Planning

Jonathan P. Rowe and James C. Lester

Department of Computer Science, North Carolina State University, Raleigh, NC 27695, USA
{jprowe, lester}@ncsu.edu

Abstract

A key functionality provided by interactive narrative systems is narrative adaptation: tailoring story experiences in response to users' actions and needs. We present a data-driven framework for dynamically tailoring events in interactive narratives using modular reinforcement learning. The framework involves decomposing an interactive narrative into multiple concurrent sub-problems, formalized as *adaptable event sequences* (AESs). Each AES is modeled as an independent Markov decision process (MDP). Policies for each MDP are induced using a corpus of user interaction data from an interactive narrative system with exploratory narrative adaptation policies. Rewards are computed based on users' experiential outcomes. Conflicts between multiple policies are handled using arbitration procedures. In addition to introducing the framework, we describe a corpus of user interaction data from a testbed interactive narrative, CRYSTAL ISLAND, for inducing narrative adaptation policies. Empirical findings suggest that the framework can effectively shape users' interactive narrative experiences.

Introduction

Interactive narratives provide opportunities for users to actively participate in rich, engaging story experiences that are dynamically tailored to individual users' preferences and needs. The capacity to dynamically augment and revise narrative plans has shown promise for several applications of interactive narrative, including entertainment (Mateas and Stern 2005; Porteous, Cavazza, and Charles 2010; Thue et al. 2007; Yu and Riedl 2012), education (Lee, Mott, and Lester 2012; Thomas and Young, 2010), and training (Si, Marsella, and Pynadath 2005).

Over the past several years, there has been growing interest in data-driven techniques for interactive narrative planning. For example, the DODM family of drama managers model interactive narrative planning with Markov

decision processes, inducing policies using classical reinforcement learning techniques (Nelson et al. 2006) as well as targeted trajectory distribution Markov decision processes (Roberts et al. 2006). Empirical evaluations of DODM have shown promise, but the data-intensive techniques rely on corpora generated by simulated users. Supervised machine learning techniques have been employed to induce interactive narrative planners from human demonstrations (Lee, Mott, and Lester 2012). Empirical studies have indicated that the approach is effective, but relying on human "wizards" to demonstrate narrative adaptations is time- and labor-intensive. Yu and Riedl (2012) employed prefix-based collaborative filtering to personalize interactive story generation from recurring user self-reports. This technique has also shown promise in empirical studies, but routinely soliciting user self-reports is likely to be disruptive in many interactive narratives. While each of these examples present data-driven frameworks for inducing interactive narrative planners, to date there have been no systems that directly learn narrative adaptation policies from only user interaction and outcome data.

In this paper, we introduce a data-driven framework for devising interactive narrative planning models using modular reinforcement learning. Our framework supports inducing interactive narrative planning policies directly from user interaction data without requiring simulated users, human demonstrations, or recurring user self-reports. The approach is inspired by work on reinforcement learning-based tutorial dialogue management (Chi, VanLehn, and Litman 2010), a related experience management problem that focuses on enhancing user experiences in interactive discourses. Our approach leverages a novel abstraction for decomposing interactive narrative plots called *adaptable event sequences*. Dynamically adapting a narrative event sequence is modeled as a Markov decision process (MDP), with rewards computed from students' experiential outcomes. Distinct policies are independently machine-learned for each AES, and the policies are then re-

combined using arbitration procedures to form a composite interactive narrative planner.

After introducing our framework, we describe the testbed interactive narrative environment, CRYSTAL ISLAND, used to investigate modular reinforcement learning-based interactive narrative planning. Specifically, we use a modified version of CRYSTAL ISLAND to collect a training corpus for inducing interactive narrative planning policies from user interaction data. We then describe an implemented interactive narrative planner for CRYSTAL ISLAND, as well as empirical findings on the implemented planner’s effectiveness at positively shaping user experiences. We conclude by discussing directions for future work.

Interactive Narrative Planning with Modular Reinforcement Learning

Our data-driven framework for modeling interactive narrative planning is based on modular reinforcement learning. Modular reinforcement learning is a multi-goal extension of classical single-agent reinforcement learning (Bhat, Isbell, and Mateas 2006; Karlsson 1997; Sprague and Ballard 2003). In reinforcement learning, an agent must learn a policy for selecting actions in an uncertain environment, guided by delayed rewards, in order to accomplish a goal (Kaelbling, Littman, and Moore 1996; Sutton and Barto 1998). 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 independent Markov decision processes (MDPs) $M = \{M_i\}_1^N$, where $M_i = (S_i, A_i, P_i, R_i)$, and each MDP corresponds to a sub-problem in the composite reinforcement learning task. The state space of the composite task is defined as the cross product of the state sub-spaces for each individual MDP: $S = S_1 \times S_2 \times \dots \times S_N$. The action set for the composite agent is given by the union of the action subsets for each independent MDP: $A = A_1 \cup A_2 \cup \dots \cup A_N$. Each agent M_i has its own reward model R_i . The solution to a modular reinforcement learning problem is a set of N concurrent policies: $\pi^* = \{\pi_i^*\}_1^N$, where π_i^* is the “optimal” policy for a single 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 an arbitration procedure to be employed to select an appropriate action. It should be noted that the policy obtained for each constituent MDP may not be theoretically guaranteed to be optimal. Guarantees of optimality are predicated on the following three assumptions: state representations are fully Markovian, the environment does not change from learning-time to run-time, and the decision-making agent selects

all future actions according to an optimal policy. Regardless, modular reinforcement learning is expected to yield “good” policies that are effective in practice.

Decomposing Interactive Narratives

In order to model interactive narrative planning as a modular reinforcement-learning problem, we establish an abstraction for decomposing an interactive narrative planning task. Specifically, we define two concepts: *adaptable event sequence* and *narrative adaptation*.

Definition. An adaptable event sequence (AES) is a series of one or more related story events that can unfold in multiple forms within an interactive narrative. Each manifestation of an AES involves inserting, re-ordering, augmenting, or removing story events from a “canonical” narrative sequence. While each manifestation of an AES may have a distinct impact on an interactive narrative, it must be interchangeable with all other manifestations without affecting the narrative’s coherence. An AES can occur one or multiple times during an interactive narrative.

Definition. A narrative adaptation is the concrete sequence of events occurring in a specific manifestation of an AES.

An AES is similar to a beat—an abstraction for an atomic unit of interactive narrative (Mateas and Stern 2005)—but there are notable distinctions between the two. In particular, a beat typically occurs once in an interactive narrative, whereas an AES may occur one or multiple times. Additionally, interactive narrative beats classically focus on jointly coordinated behaviors between multiple virtual characters (Mateas and Stern 2005), whereas AESs do not focus on a particular type of narrative content. The number of story branches encoded by an AES is typically smaller than a beat, but the types of narrative content specified by an AES are more general.

To illustrate the concept of an AES, consider the following example: an event sequence that occurs after a user asks a virtual character about her backstory. The virtual character may respond in several ways. She may provide a detailed explanation about her backstory. Alternatively, she may respond suspiciously and reveal only a few details. Or, she may refuse to respond at all. In this example, each of the three possible responses is an alternate manifestation of a “character backstory” AES. The particular response provided by the virtual character is an example of a single narrative adaptation. When an interactive narrative planner performs a narrative adaptation, its “action” is a decision about which manifestation of the AES should occur. In this example, the AES could occur once during the interactive narrative, or multiple times, depending on how many times the user initiates a conversation with the virtual character. In this manner, an interactive narrative planner can direct

the virtual character’s responses to change from one interaction to the next based on the narrative context. As noted above, AESs are not restricted to virtual character behaviors. For example, an AES could specify the alternate locations of a valuable object at different phases of a narrative, or it could encode alternate ways that a user’s abilities are augmented by a story event.

Using these concepts, interactive narrative planning can be cast as a collection of sequential decision-making problems about AESs impacting users’ interactive narrative experiences. In effect, each AES is a sub-problem within an overarching interactive narrative planning task. The planner’s success in performing narrative adaptations is indicated by users’ experiential and attitudinal outcomes. These outcomes can be measured by several different methods, such as post-hoc questionnaires or metrics calculated from users’ in-game behaviors. Outcome metrics serve as the basis for defining delayed rewards that guide interactive narrative decision-making. This conceptualization—a collection of sequential decision-making tasks with delayed rewards—points to modular reinforcement learning for modeling interactive narrative planning.

Modeling interactive narrative planning as a modular reinforcement learning problem involves four representational considerations. First, the overall task must be decomposed, which involves identifying the set of AESs and goals for the narrative-centered tutorial planner. Each AES corresponds to a sub-problem in modular reinforcement learning; in our framework, each AES is modeled as a distinct, independent MDP. Second, techniques for devising appropriate state representations for each sub-problem must be identified. This includes identifying categories of state features, as well as utilizing principled techniques for feature selection. Third, the set of concrete event sequences that comprise each narrative adaptation must be formally specified. These sequences constitute the action sets for each MDP. Fourth, the task environment for each MDP must be precisely characterized, including when narrative adaptations are performed, sources of uncertainty in the environment’s dynamics, and the duration of narrative experiences. These criteria shape the state transition model and reward model definitions for each AES.

Policy Induction

In order to induce interactive narrative planning policies from user interaction data, we employ off-line techniques for reinforcement learning. Off-line learning refers to a class of procedures that separate the data-collection and model-operation stages. They are contrasted with on-line learning techniques, such as temporal difference methods, that interleave data collection, policy learning, and policy execution stages. When inducing interactive narrative planning policies from human user data, off-line learning requires that users interact with a version of the interactive narrative environment that is specifically designed for col-

lecting a training corpus. This environment should be identical to the final system—including the set of AESs that are supported—with the exception of the policies used to drive narrative adaptation decisions. The data collection system should perform narrative adaptations in a manner that is *exploratory*, such as a random policy, rather than a manner that seeks to maximize accumulated reward. This enables a broad sampling of the state space, producing data that can be used to calculate an approximate environment model. The environment model, which encodes each MDP’s state transition and reward dynamics, are calculated by counting state-transition frequencies in the training corpus. From the state transition and reward models, dynamic programming techniques such as value iteration or policy iteration can be employed to induce a set of “optimal” narrative adaptation policies for each AES (Sutton and Barto 1998). This approach is a form of certainty equivalent learning (Kaelbling, Littman, and Moore 1996). The resulting policies are then implemented in a new, deployable version of the interactive narrative environment.

Policy Arbitration

There may be circumstances where multiple decision points are triggered simultaneously. In this situation, the interactive narrative planner may receive multiple simultaneous action recommendations from distinct policies. If all of the action recommendations agree, no arbitration is necessary. However, if the recommended actions differ, arbitration techniques must be employed in order to choose a single action for the composite planner.

Previous work on modular reinforcement learning has investigated domain-independent arbitration procedures that exclusively consider the state-action values obtained for each module. Example arbitration procedures of this type include greatest mass and nearest neighbor strategies (Karlsson 1997). These arbitration procedures require that rewards for all MDPs be comparable. Bhat and colleagues (2006) have argued that domain-independent arbitration procedures are inadequate for settings where modules’ rewards are not comparable. In these cases, interactive narrative planning requires an arbitrator endowed with domain-specific knowledge about the relations and tradeoffs between individual modules. However, in this work we focus on domain-independent arbitration procedures for MDPs with comparable rewards. Specifically, we adopt the greatest mass arbitration procedure in an implemented planner, which is described in the following two sections.

Exploratory Policy Corpus

In order to investigate modular reinforcement learning-based interactive narrative planning, we used an educational interactive narrative environment, CRYSTAL ISLAND. CRYSTAL ISLAND (Figure 1) is built on Valve Software’s Source™ engine, the 3D game platform for Half-Life 2.

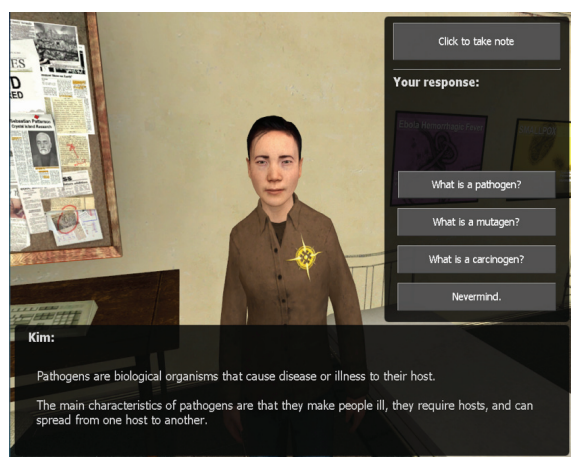


Figure 1. Dialogue interaction in CRYSTAL ISLAND.

The environment’s educational focus is middle school microbiology, and it features a science mystery in which students discover the identity and source of an infectious disease that is plaguing a research team stationed on the island. Students play the role of a visitor who recently arrived on the island and must save the research team from the outbreak. Over the past several years, CRYSTAL ISLAND has been the subject of extensive empirical investigation, and has been found to provide substantial learning and motivational benefits (Rowe et al. 2011).

To investigate interactive narrative planning in CRYSTAL ISLAND, we developed a modified version of the system that includes thirteen adaptable event sequences. Specifically, we identified plot points in CRYSTAL ISLAND’s narrative that could be transformed into adaptable event sequences, in effect retrofitting the existing interactive narrative structure with AESs. The process for identifying adaptable event sequences was driven by several factors: 1) we selected events that would embody three distinct categories of narrative adaptations: plot adaptations, discourse adaptations, and user tailoring (Rowe, Shores, Mott, and Lester 2010b); 2) we chose narrative elements that had previously been found to significantly relate to students’ learning and problem-solving outcomes (Spire, Rowe, Mott, and Lester 2011; Rowe, Shores, Mott, and Lester 2010); and 3) we selected events that were critical to how students received, and processed, information required to solve the mystery. We elected to incorporate thirteen distinct adaptable event sequences (rather than one or two) in order to have a broad range of mechanisms to shape and manipulate users’ interactive narrative experiences. The adaptable event sequences were as follows: *Diagnosis Worksheet Feedback*, *Quentin’s Revelation*, *Bryce’s Revelation*, *Off-Task Behavior Discouragement*, *Record Findings Reminders*, *Increased Urgency*, *Knowledge Quizzes*, *Next Goal Prompt*, *Details of Teresa’s Symptoms*, *Details of Bryce’s Symptoms*, *Mystery’s Solution*, *Reflection*

Prompt, and *Initial Lab Test Count*. While space limitations preclude a detailed description of each AES, a more extensive discussion is available in (Rowe, 2013).

In order to illustrate how the AESs might unfold during a student interaction with CRYSTAL ISLAND, consider the following scenario. When a student begins the narrative, the *Mystery’s Solution* AES immediately occurs behind the scenes, which involves selecting one of six possible “solutions” to the mystery. The narrative planner selects *salmonellosis* as the mystery disease and *contaminated milk* as the disease’s transmission source. This AES is invisible to the student and it occurs only once in the narrative, but the selection dictates which symptoms and medical history the sick characters report during the narrative. As the student explores the camp and learns about the outbreak, she initiates a conversation with a sick scientist named Teresa. When the student asks about Teresa’s symptoms, the *Details of Teresa’s Symptoms* AES is triggered, which controls the degree of information that Teresa will provide in her response. The planner chooses a narrative adaptation where Teresa provides minimal information, leading Teresa to groan and explain that she has a fever. This narrative adaptation is an alternative to moderate or highly detailed descriptions of Teresa’s symptoms. After the conversation, the *Record Findings Reminder* AES is triggered because the player just received relevant information for diagnosing the illness. During this AES, the narrative planner chooses whether to provide a hint to the student to record her recent finding (Teresa’s fever) in an in-game diagnosis worksheet. The planner elects to deliver the hint, causing the student’s in-game smartphone to ring with a text message from the camp nurse containing the reminder. This AES will occur several more times during the narrative, or anytime that the student acquires useful information from a virtual character, in-game book, or virtual laboratory test. The interactive narrative continues in this manner, largely driven by the student’s actions, which periodically trigger AESs that control how events unfold in the storyworld.

After modifying CRYSTAL ISLAND to incorporate adaptable event sequences, we conducted a pair of human subject studies to collect training data for inducing the interactive narrative planning models. The first study involved 300 eighth-grade students from a North Carolina middle school. The second study involved 153 eighth-grade students from a different North Carolina middle school. Every student across both studies used the same version of CRYSTAL ISLAND endowed with thirteen adaptable event sequences. Students in both studies followed identical study procedures, and used CRYSTAL ISLAND individually. Students interacted with the interactive narrative until they solved the mystery, or 55 minutes of interaction time elapsed, whichever occurred first.

While using CRYSTAL ISLAND, students unknowingly encountered AESs several times. At each AES, the envi-

ronment selected a narrative adaptation according to a random policy, uniformly sampling the planning space. By logging these narrative adaptations, as well as students' subsequent responses, the environment broadly sampled the space of policies for controlling adaptable event sequences. In addition, several questionnaires were administered prior to, and immediately after, students' interactions with CRYSTAL ISLAND. These questionnaires provided data about students' individual characteristics, curricular knowledge, and engagement with the environment.

The data collected during both studies were combined into a single corpus for inducing interactive narrative planning policies. The corpus consisted of two parts: students' interaction logs, and students' pre/post questionnaire results. After removing data from students with incomplete or inconsistent records, there were 402 students remaining in the data set. The resulting data set consists of 315,407 observations of narrative events. In addition to student actions, there are 10,057 instances of narrative adaptations in the corpus, which correspond to approximately 25 narrative adaptations per student. The most commonly occurring adaptable event sequences were the following: *record findings reminder*, *next goal prompt*, *knowledge quiz*, and *diagnosis worksheet feedback*. More details about the corpus are available in (Rowe, 2013).

Implemented Interactive Narrative Planner

In order to evaluate our interactive narrative planning framework, we modeled each of the adaptable event sequences in CRYSTAL ISLAND as an independent MDP. Using the training corpus described in the prior section, we induced "optimal" policies for each MDP to control CRYSTAL ISLAND's run-time narrative adaptation behavior, with the exception of one AES for which we had insufficient training data (off-task behavior discouragement).

In order to define the set of MDPs, we identified a state representation, action set, and reward function for each AES. We used the training corpus described in the prior section to calculate each MDP's state transition model and reward model, and we utilized value iteration to optimize the planning policies. Additionally, we selected the greatest-mass arbitration procedures for AESs in which concurrent policies prescribed conflicting actions.

State and Action Representation

In the interactive narrative planning model, all of the concurrent MDPs shared the same state representation. The state representation consisted of eight binary features drawn from three categories: narrative features, individual difference features, and gameplay features. We limited the state representation to eight binary features to reduce potential data sparsity. The first four features were narrative-focused. Each feature was associated with a specific sub-

goal from CRYSTAL ISLAND's narrative and indicated whether the sub-goal had been completed in the narrative thus far. The four sub-goals were: 1) camp nurse presents mystery's main objective, 2) student tests disease's transmission source in the laboratory, 3) student submits a diagnosis to the camp nurse, and 4) student submits a correct, complete diagnosis and solves the mystery. These sub-goals were chosen because they distinguish between salient phases of the narrative problem-solving scenario.

The next two state features were based on students' individual differences. The first feature was computed from a median split on students' *content knowledge pre-test scores*, and the second feature was computed from a median split on students' *game-playing frequency* reports.

The final two state features were computed from students' gameplay behaviors. Specifically, we computed running median splits on students' *laboratory testing behaviors* and in-game *book reading behaviors*. The running median splits were necessary to account for these features' values changing during students' interactions.

The action sets for the 12 MDPs corresponded to the narrative adaptations for the associated AESs. The action sets' cardinalities ranged from binary to 6-way decisions.

Reward Models

A single reward function was used for all of the MDPs, which was based on students' normalized learning gains. Normalized learning gains are the normalized difference between students' pre- and post-study knowledge test scores. In order to determine the actual reward values, normalized learning gains were first calculated for each student, and then a median split was performed. Students who had normalized learning gains that were greater than or equal to the median were awarded +100 points at the conclusions of their episodes. Students with normalized learning gains that were less than the median were awarded -100 points. All rewards were associated with the final observed state transitions of each episode. In computing solutions to the MDPs, we sought to maximize the expected discounted reward obtained during an episode.

The probability values in the state transition models P_i were calculated in terms of state values at successive encounters for their respective AESs. In other words, state transitions did not characterize the immediate effects of narrative adaptations, but rather the aggregate impacts of narrative adaptations between sequential occurrences (i.e., decision points) of a given AES. If a particular AES in an episode possessed no successive decision point, the episode's final state was used instead.

Policy Induction

In order to induce the interactive narrative planning policies, we used the value iteration algorithm (Sutton and Barto 1998). The 12 MDPs, one for each AES in CRYSTAL IS-

LAND, were implemented with a reinforcement-learning library written in Python by the first author. A discount factor of 0.9 was used. Distinct policies were induced for each AES. The resulting policies were encoded as direct mappings between state values and planner actions.

Empirical Findings

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

To evaluate CRYSTAL ISLAND’s induced narrative planner, we conducted a controlled experiment with middle school students, which compared the induced model to a baseline narrative adaptation approach. Participants were drawn from a different school district than the the corpus collection studies. A total of 75 eighth-grade students participated. Among these students, fourteen were removed from the dataset due to incomplete or inconsistent data.

The study had two conditions: a Baseline planner condition and an Induced planner condition. Students in both conditions played CRYSTAL ISLAND, but the conditions differed in terms of the narrative adaptation policies employed by each system’s planner. The Baseline Planner employed a uniform random policy, where narrative adaptations were selected randomly whenever the planner encountered a decision point. This was equivalent to the exploratory planner used during the corpus collection studies. The Induced Planner followed policies obtained by solving the Markov decision processes associated with each AES.

Students were randomly assigned to the two conditions when they entered the experiment room to play CRYSTAL ISLAND. Among students with complete data, random assignment resulted in 33 participants in the Induced Planner condition, and 28 participants in the Baseline Planner condition. Students played until they solved the mystery or the interaction time expired, whichever occurred first. The study procedure and pre/post questionnaires were otherwise identical to the corpus collection studies.

Analyses of students’ narrative experiences found that in the Induced Planner condition, 61% of students solved the mystery, compared to 46% of students in the Baseline Planner condition. This finding is notable because students in the Induced Planner condition solved CRYSTAL ISLAND’s mystery at a higher rate than in any prior study with comparable versions of the CRYSTAL ISLAND software, although a Pearson’s chi-square test failed to find evidence of statistical significance. To follow up on the analysis, we examined condition effects on problem-solving efficiency. We examined the time taken to solve CRYSTAL

ISLAND’s mystery in each experimental condition. Because several students in each condition did not complete the scenario in the allotted time, we devised a Projected Completion Time metric, which is shown in Equation 1.

$$Proj. Time = Session Time + (300 \times Unsolved Goal Cnt) \quad (1)$$

On average, students in the Induced Planner condition had a Projected Completion Time of 3607 seconds ($SD = 1286$), and students in the Baseline Planner condition had a Projected Completion Time of 4119 seconds ($SD = 988$). A two-tailed t-test indicated that the difference between the two conditions was marginally statistically significant, $t(59) = 1.72$, $p = .09$. These findings provide initial evidence that students interacting with the Induced Planner solved CRYSTAL ISLAND’s educational mystery more efficiently than students interacting with the Baseline Planner.

Conclusions and Future Work

We have presented a novel framework for interactive narrative planning that decomposes the task into multiple independent sub-problems, formalized as adaptable event sequences. These adaptable event sequences are modeled as concurrent Markov decision processes, with rewards based on users’ experiential outcomes. Policies for solving the MDPs are obtained through reinforcement learning techniques, such as certainty equivalent learning. Training data for learning the narrative adaptation policies is collected through human subject studies with an exploratory version of the interactive narrative planner. Arbitration techniques are employed to resolve conflicts between multiple competing AESs. In addition to presenting the framework, we described a sizable training corpus, consisting of user interaction data, for inducing interactive narrative planning policies. We presented details of an implemented interactive narrative planner for the CRYSTAL ISLAND environment, and empirically demonstrated its capacity to positively shape users’ interactive narrative experiences. Building on these findings, in future work it will be important to systematically investigate the impacts of alternate reward models that characterize the quality of users’ narrative experiences, as well as alternate state representations, on induced interactive narrative planning policies.

Acknowledgments

This research was supported by the National Science Foundation under grants REC-0632450, IIS-0812291, and DRL-0822200. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Additional support was provided by the Bill and Melinda Gates Foundation, the William and Flora Hewlett Foundation, and EDUCAUSE.

References

- Bhat, S., Isbell, C. L., and Mateas, M. 2006. On the difficulty of modular reinforcement learning for real-world partial programming. In *Proceedings of the 21st National Conference on Artificial Intelligence*, 318–323. Menlo Park, CA: AAAI Press.
- Chi, M., VanLehn, K., and Litman, D. 2010. Do micro-level tutorial decisions matter: Applying reinforcement learning to induce pedagogical tutorial tactics. In *Proceedings of the 10th International Conference on Intelligent Tutoring Systems*, 224–234. New York, NY: Springer.
- Johnson, W. L. 2010. Serious use of a serious game for language learning. *International Journal of Artificial Intelligence in Education*, 20(2): 175–195.
- Kaelbling, L. P., Littman, M. L., and Moore, A. W. 1996. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4: 237–285.
- Karlsson, J. 1997. Learning to solve multiple goals. Ph.D. diss., Department of Computer Science, University of Rochester, Rochester, NY.
- Lee, S. Y., Mott, B. W., and Lester, J. C. 2012. Real-time narrative-centered tutorial planning for story-based learning. In *Proceedings of the 11th International Conference on Intelligent Tutoring Systems*, 476–481. New York, NY: Springer.
- Mateas, M., and Stern, A. 2005. Structuring content in the Façade interactive drama architecture. In *Proceedings of the 1st International Conference on Artificial Intelligence and Interactive Digital Entertainment*, 93–98. Menlo Park, CA: AAAI Press.
- Nelson, M. J., Roberts, D. L., Isbell, C. L., and Mateas, M. 2006. Reinforcement learning for declarative optimization-based drama management. In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems*, 775–782. New York, NY: ACM Press.
- Porteous, J., Cavazza, M., & Charles, F. 2010. Narrative generation through characters’ point of view. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, 1297–1304. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.
- Roberts, D. L., Nelson, M. J., Isbell, C. L., Mateas, M., and Littman, M. L. 2006. Targeting specific distributions of trajectories in MDPs. In *Proceedings of the 21st National Conference on Artificial Intelligence*, 1213–1218. Menlo Park, CA: AAAI Press.
- Rowe, J. P., Shores, L. R., Mott, B. W., and Lester, J. C. 2010. Individual differences in gameplay and learning: A narrative-centered learning perspective. In *Proceedings of the 5th International Conference on Foundations of Digital Games*, 171–178. New York, NY: ACM Press.
- Rowe, J. P., Shores, L. R., Mott, B. W., and Lester, J. C. 2010b. A framework for narrative adaptation in interactive story-based learning environments. In *Proceedings of the FDG ’10 Workshop on Intelligent Narrative Technologies III*.
- Rowe, J. P., Shores, L. R., Mott, B. W., and Lester, J. C. 2011. Integrating learning, problem solving, and engagement in narrative-centered learning environments. *International Journal of Artificial Intelligence in Education*, 21(2): 115–133.
- Rowe, J. P. 2013. Narrative-centered tutorial planning with concurrent Markov decision processes. Ph.D. diss., Department of Computer Science, North Carolina State University, Raleigh, NC.
- Si, M., Marsella, S. C., and Pynadath, D. V. 2005. Thespian: An architecture for interactive pedagogical drama. In *Proceedings of the 12th International Conference on Artificial Intelligence in Education*, 595–602. Amsterdam, Netherlands: IOS Press
- Spires, H. A., Rowe, J. P., Mott, B. W., and Lester, J. C. 2011. 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.
- Sprague, N., and Ballard, D. 2003. Multiple-goal reinforcement learning with modular sarsa (0). In *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, 1445–1447. San Francisco, CA: Morgan Kaufmann
- Sutton, R. S., and Barto, A. G. 1998. *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Thomas, J. M., and Young, R. M. 2010. Annie: Automated generation of adaptive learner guidance for fun serious games. *IEEE Transactions on Learning Technologies*, 3(4): 329–343.
- Thue, D., Bulitko, V., Spetch, M., and Wasylshen, E. 2007. Interactive storytelling: A player modelling approach. In *Proceedings of the 3rd Artificial Intelligence and Interactive Digital Entertainment Conference*, 43–48. Menlo Park, CA: AAAI Press.
- Watson, S., Vannini, N., Woods, S., Dautenhahn, K., Sapouna, M., Enz, S., Schneider, W., et al. 2010. Inter-cultural differences in response to a computer-based anti-bullying intervention. *Educational Research*, 52(1): 61–80.
- Yu, H., and Riedl, M. O. 2012. A sequential recommendation approach for interactive personalized story generation. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, 71–78. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.