Proceedings of the Tenth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2014)

Optimizing Player Experience in Interactive Narrative Planning: A Modular Reinforcement Learning Approach

Jonathan P. Rowe, Bradford W. Mott and James C. Lester
Department of Computer Science, North Carolina State University, Raleigh, NC 27695, USA
{jprowe, bwmott, lester}@ncsu.edu

Abstract

Recent years have witnessed growing interest in data-driven approaches to interactive narrative planning and drama management. Reinforcement learning techniques show particular promise because they can automatically induce and refine models for tailoring game events by optimizing reward functions that explicitly encode interactive narrative experiences' quality. Due to the inherently subjective nature of interactive narrative experience, designing effective reward functions is challenging. In this paper, we investigate the impacts of alternate formulations of reward in a reinforcement learning-based interactive narrative planner for the CRYSTAL ISLAND game environment. We formalize interactive narrative planning as a modular reinforcement-learning (MRL) problem. By decomposing interactive narrative planning into multiple independent sub-problems, MRL enables efficient induction of interactive narrative policies directly from a corpus of human players' experience data. Empirical analyses suggest that interactive narrative policies induced with MRL are likely to yield better player outcomes than heuristic or baseline policies. Furthermore, we observe that MRL-based interactive narrative planners are robust to alternate reward discount parameterizations.

Introduction

Interactive narratives provide opportunities for players to participate in rich, engaging story experiences that are dynamically tailored to individual players' preferences and actions. The capacity to dynamically augment and revise narrative plans has shown promise for several applications of interactive narrative, including entertainment (Mateas and Stern 2005; McCoy et al. 2013; Porteous, Cavazza, and Charles 2010; Thue et al. 2007; Yu and Riedl 2012), education (Lee et al. 2014; Thomas and Young 2010), and training (Si, Marsella, and Pynadath 2005).

Over the past decade, there has been growing interest in data-driven techniques for interactive narrative planning (Lee et al. 2014; Orkin and Roy 2012; Roberts et al. 2006; Yu and Riedl 2014). Yu and Riedl (2014) employ prefix-based collaborative filtering to personalize interactive stories based on recurring player self-reports. This technique has shown promise in empirical studies, but recurring self-reports are likely to prove disruptive for many types of interactive narratives. Orkin and Roy (2012) have proposed collective artificial intelligence, an approach that combines crowdsourced game log collection and annotation procedures with automated plan recognition, to control non-player characters in two-player interactive stories. Although a promising direction, collective artificial intelligence requires a substantial amount of corpus annotation to incorporate story structure information.

To devise data-driven interactive narrative planners without extensive annotation or self-report procedures, reinforcement learning is a promising approach (Nelson et al. 2006). Reinforcement learning enables automatic induction of sequential decision-making models that operate under uncertainty, an apt description of interactive narrative planning. Furthermore, a natural way to evaluate interactive narratives is to ask players to retrospectively judge the quality of their experiences. This approach aligns nicely with reinforcement learning's support for optimizing delayed rewards. However, reinforcement-learning problems suffer from the curse of dimensionality, in most cases requiring large amounts of training data. Typically, interactive software applications, such as spoken dialogue systems and drama managers, leverage simulated users in order to acquire adequate training data for reinforcement-learned planners (Chi et al. 2011; Nelson et al. 2006; Roberts et al. 2006; Tetreault and Litman 2008). Unfortunately, simulating human users raises its own challenges, particularly in media characterized by subjective user experiences, a hallmark of interactive narrative.

To address these challenges, we leverage a data-driven framework for inducing interactive narrative planners that uses modular reinforcement learning (MRL). Our approach exploits structural characteristics of interactive narrative to decompose the planning problem, enabling narrative adaptation policies to be induced directly from human players' interaction and outcome data. As a consequence, it is possible to ask players directly about the quality of their experiences post hoc and subsequently derive reward functions that can be optimized by an interactive narrative planner. The approach creates opportunities for investigating the impact of alternate formulations of reward (e.g., engagement-centric measures vs. plot-centric measures) on interactive narrative planning models, and ultimately player experiences. In this paper, we describe an overview of the MRL framework and its implementation in the CRYSTAL ISLAND game environment, and present results from an empirical investigation of alternate reward parameterizations for interactive narrative planning models.

Interactive Narrative Planning with 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: $S = S_1 \times S_2 \times \ldots \times S_N$. The action set for the composite agent is given by the union of the action subsets for each independent MDP: $A = A_1 \cup A_2 \cup \ldots \cup A_N$. Each agent M_i has its own probabilistic state transition model P_i and 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 select an appropriate action. It should be noted that, in MRL, the policy obtained for an overall planner is not necessarily guaranteed to be optimal. Theoretical guarantees of optimality are predicated on three assumptions: state representations are 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. These assumptions do not always apply in MRL. However, MRL generally yields "good" policies that are effective in practice (Karlsson 1997).

Adaptable Event Sequences

In order to model interactive narrative planning as a MRL problem, we utilize the concept of an *adaptable event sequence* (AES). AESs provide a structure for decomposing interactive narrative planning tasks. 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: by providing a detailed explanation about her backstory, or by responding suspiciously and revealing only a few details, or by refusing to respond at all. Each of these three types of responses is an alternate manifestation of the *NPC Backstory* event sequence. Each option is coherent within the plot, and all three can be interchanged. For this reason, we refer to the event sequence as *adaptable*. In other words, it is an AES.

When an interactive narrative planner chooses a particular manifestation of an AES to perform, it can be said to be performing a *narrative adaptation*. An AES may entail a single narrative adaptation (e.g., the player speaks to the NPC once, and the NPC responds), or a series of narrative adaptations (e.g., the player speaks to the NPC multiple times, each time requiring a response). In the latter case, an interactive narrative planner can direct the NPC's response to change from one interaction to the next.

Importantly, AESs are not restricted to only controlling virtual character behaviors. For example, an AES could encode the location of an important object at different phases of a narrative, or encode alternate ways that players' abilities are augmented during important story events. Additionally, multiple AESs can be interleaved with one another. Conceptually, AESs encode distinct threads of story events, each potentially involving multiple decision points spanning an entire story. For this reason, AESs can be said to operate *concurrently*.

To illustrate, consider a player who encounters the earlier-mentioned NPC and receives a terse, suspicious response during the associated *NPC Backstory* AES. Later, when the player aims to retrieve an important object for the NPC, a corresponding *Object Location* AES models the position of the object among several candidate locations. It unfolds by positioning the object inside a distant building. When the player returns to the NPC and asks about her backstory a second time, a second decision point for the

NPC Backstory AES is triggered. This time, a detailed response about the NPC's backstory is produced. In this case, the *NPC Backstory* and *Object Location* AESs have operated concurrently, interleaving decisions about narrative adaptations.

The concept of an AES is applicable to a broad range of interactive narratives, including many role-playing games and adventure games, so long as they feature event sequences that 1) can unfold coherently in several possible ways, and 2) recur multiple times during gameplay. Formally, we define an AES as follows:

Definition. An adaptable event sequence (AES) is as a series of one or more related story events that can unfold in multiple ways within an interactive narrative. Each manifestation of an AES involves inserting, re-ordering, augmenting, or removing story events from some other valid manifestation of the narrative sequence. Each manifestation of an AES must be interchangeable with all other manifestations without affecting the narrative's coherence. The events of an AES can occur one or multiple times during an interactive narrative.

Leveraging the concept of an AES, interactive narrative planning can be cast as a collection of sequential decision-making problems about selecting narrative adaptations within an interactive narrative. Devising a data-driven interactive narrative planner is cast as a modular reinforcement-learning problem 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 narrative adaptations 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 measures of players' experiential or attitudinal outcomes, including post-hoc questionnaires or in-game behavior metrics. And a state transition model P_i encodes the probability of transitioning between two specific states during successive decision points for the AES. Leveraging this mapping between AESs and MDPs, it is possible to employ model-based reinforcement learning techniques to induce policies for interactive narrative planning.

Off-Line Reinforcement Learning

We employ off-line techniques for reinforcement learning on a corpus of player interaction and outcome data. When inducing interactive narrative planning policies from player data, off-line learning requires that players interact with a version of the interactive narrative environment that is specifically designed for collecting 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 model-based reinforcement learning based on certainty equivalence (Kaelbling, Littman, and Moore 1996). After obtaining machine-learned policies, they are implemented in a new, deployable version of the interactive narrative environment.

Policy Arbitration

There may be circumstances where decision points for multiple AESs are triggered simultaneously. In this situation, the interactive narrative planner may receive multiple conflicting action recommendations from distinct policies. In this case, arbitration procedures must be employed to choose a single action for the planner. In our work, AES conflicts are rare but they do occur. We utilize a domain-independent arbitration procedure known as greatest mass arbitration (Bhat et al. 2006; Karlsson 1997).

Interactive Narrative Planning Corpus

In order to investigate MRL-based interactive narrative planning, we use 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. The environment's educational focus is middle school microbiology, and it features a science mystery in which players discover the identity and source of an infectious disease that is plaguing a research team stationed on an island. Players adopt 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 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 AESs. We selected thirteen AESs (rather than one or two) in order to incorporate a broad range of narrative adaptations for shaping players' experiences. The AESs ranged in form and content, and included manipulations to character dialogue, decisions about deliver-

Figure 1. Dialogue interaction in Crystal Island.

ing hints to the player, and augmentations to the player's in-game abilities. Space limitations preclude a detailed description of every AES, but a more extensive discussion is available in (Rowe 2013).

To illustrate how AESs unfold during a player interaction with Crystal Island, consider the following scenario. When a player begins the narrative, the *Mystery's Solution* AES immediately occurs behind the scenes, 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 player, but the selection dictates which symptoms and medical history the sick characters report. As the player explores the camp and learns about the outbreak, she initiates a conversation with a sick scientist named Teresa. When the player asks about Teresa's symptoms, the *Details of Teresa's Symptoms* AES is triggered, which controls the degree of information that Teresa provides 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. If the player chooses to ask Teresa about her symptoms again later in the narrative, the planner may choose a different response. 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 player to record her recent findings. The interactive narrative continues in this manner, driven by the player's actions and periodically triggering narrative adaptations that shape how the story unfolds.

After modifying Crystal Island to incorporate adaptable event sequences, we conducted a pair of human subject studies to collect training data for inducing an interactive narrative planner. The first study involved 300 students from a North Carolina middle school. The second study involved 153 students from a different North Carolina middle school. Every participant used the same version of Crystal Island endowed with thirteen AESs. Participants 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 elapsed, whichever occurred first.

While using Crystal Island, participants unknowingly encountered AESs several times. At each AES, the environment selected a narrative adaptation according to a random policy, uniformly sampling the planning space. By logging these narrative adaptations, as well as participants' 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, participants' interactions with Crystal Island. These questionnaires provided data about participants' individual characteristics, curricular knowledge, and engagement with the environment.

The data collected during both studies were combined into a single corpus. The corpus consisted of two parts: players' interaction logs, and players' pre/post questionnaire results. After removing data from participants with incomplete or inconsistent records, there were 402 participants remaining in the data set. The resulting data set consists of 315,407 observations of narrative events. In addition to player actions, there are 10,057 instances of narrative adaptations in the corpus, which correspond to approximately 25 narrative adaptations per player. More details about the corpus are available in (Rowe 2013).

Implemented Interactive Narrative Planner

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).

State and Action Representation

All of the MDPs comprising the narrative planner shared the same state representation, which 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.[1] 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 participants' individual differences. The first feature was computed from a median split on participants' *content*

[1] State spaces of this size are common in RL-based planning for interactive software, such as spoken dialogue systems (Tetreault et al. 2008).

knowledge pre-test scores, and the second feature was computed from a median split on *game-playing frequency* reports. The final two state features were computed from participants' gameplay behaviors. Specifically, we computed running median splits on the frequency of participants' *laboratory testing behaviors* and in-game *book reading behaviors.*

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. If the entire planning task were modeled as a single MDP, it would require encoding approximately 398,000 parameters to populate the entire state transition model (128 states × 24 distinct actions × 129 states, including the terminal state), although not all state transitions were possible.

Reward Models

Three separate reward functions were compared—each yielding a separate planner—to investigate their effects on interactive narrative planning policies. The first reward function was based on players' normalized learning gains (NLG). NLG is the normalized difference between participants' pre- and post-study knowledge test scores. To determine reward values, NLG was first calculated for each participant, and then a median split was performed. Participants 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.

The second reward function was based on players' self-reported perceptions of presence, as measured by the Presence Questionnaire (PQ; Witmer and Singer 1998). Presence refers to a participant's perception of transportation into a virtual environment. Participants completed the PQ after using CRYSTAL ISLAND. The presence reward function utilized the same median split, +100/-100 reward scheme as the NLG reward function.

The third reward function was based on whether players successfully solved the mystery or not. Players who solved the mystery were awarded +100 points, and participants who did not solve the mystery were awarded -100 points. Approximately 31% of participants in the corpus successfully solved the mystery in the allotted time.

Policy Induction

To induce the interactive narrative planning policies, we used value iteration (Sutton and Barto 1998). 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 discount rates ranging between 0 and 1. The discount rate parameter governs how reward is attributed to planner actions during reinforcement learning. Policies were encoded as direct mappings between state and planner actions.

Empirical Findings

To investigate the impact of alternate reward parameterizations on MRL-based interactive narrative planners, we compared policies induced with different combinations of reward function and discount rate. Specifically, we investigated two dimensions of MRL-based interactive narrative planning: 1) the effectiveness of planners induced with MRL relative to competing techniques, and 2) MRL-based planners' sensitivity to alternate reward parameterizations.

Expected Cumulative Reward of MRL-Based Planners

To investigate the anticipated effectiveness of MRL-based interactive narrative planners, we computed expected cumulative reward (ECR) for three induced policies: one that maximized the NLG reward function, one that maximized the Presence reward function, and one that maximized the Solve-Mystery reward function. ECR is a measure of the average anticipated reward produced by a policy across all possible narrative paths and start states. In addition to computing ECR for the aforementioned policies, we computed ECR for three alternate policies that were not induced using MRL. Two of the policies were heuristic-based. The first policy uniformly sought to maximize assistance to the player at every opportunity (e.g., every time the policy had the opportunity to provide a hint, it did so). The second heuristic policy did the opposite; it sought to minimize assistance to the player under all circumstances. The third comparison policy chose actions randomly.

To calculate ECR values for each policy, the expected value for each state was determined using policy evaluation (Sutton and Barto, 1998), and then a weighted average of state values was calculated based on each state's probability of occurring as a start state. ECR values were calculated for each of the three reward functions and four types of policies. Average ECR values for all of the adaptable event sequences are shown in Table 1. For each of the reward functions, policies induced with MRL are found to yield higher ECR than heuristic or random policies. Notably, these differences are observed to be statistically significant for both the NLG and Solve-Mystery reward functions using Wilcoxon signed-rank tests, $p < .001$. This finding has two notable implications. First, it provides empirical evidence from a real-world game environment that AESs and interactive narrative planners can be devised that provide sufficient variation in player experiences to measurably impact player outcomes. Second, the results suggest that interactive narrative planners induced with MRL are likely to be more effective than heuristic-based planners for optimizing player experiences.

Sensitivity to Discount Factor and Reward

To investigate the impact of alternate reward parameterize-

Table 1. Average ECR by reward function and policy.

Reward	Value Iteration	Maximize Heuristic	Minimize Heuristic	Random
NLG	25.4 (15.5)**	6.6 (12.2)	2.4 (12.7)	3.6 (6.03)
Presence	68.4 (26.4)	56.8 (26.1)	57.9 (26.8)	56.2 (25.2)
Solve Mystery	-3.1 (8.44)**	-19.7 (10.2)	-22.6 (12.8)	-22.4 (10.3)

Note: ** signifies $p < .001$. Standard deviations shown in parentheses.

tions on MRL-based interactive narrative planners, we induced a range of policies using different combinations of reward function and discount rate across all AESs. For each AES, we compared every *induced policy* to two distinct *reference policies*.

The first reference policy was induced using a discount rate of 0.9 and the same reward function as the induced policy. In other words, an NLG induced policy was compared to an NLG reference policy, and a Presence induced policy was compared to a Presence reference policy. The second reference policy was induced using a discount rate of 0.9 and the NLG reward function. This particular reference policy was used because an interactive narrative planner induced with the same combination of reward and discount rate was found to be effective at shaping players' problem-solving processes during user studies with CRYSTAL ISLAND (Rowe 2013). To compare induced policies and reference policies, we computed Cohen's kappa for each pairing. Cohen's kappa measures the agreement between a series of judgments provided by two sources—in our case, the state-action pairs comprising each policy—adjusted for chance agreement. A kappa of 1.0 indicates perfect agreement, whereas 0.0 indicates chance agreement and -1.0 indicates no agreement.

Figure 2 shows a plot of the kappa values for each combination of induced policy, reference policy, and discount rate, averaged across all AESs. As can be seen at the top of the figure, comparing induced policies with reference policies that share the same reward function reveals high degrees of similarity. This similarity degrades only slightly across different values of discount rate. For example, a policy induced with the NLG reward function and a discount rate of 0.1 has a kappa value of 0.94 when compared to the reference NLG policy, which shares the same reward function and uses a discount of 0.9. This suggests that interactive narrative planners induced with MRL are robust to different discount rate settings.

Conversely, comparing induced policies with reference policies with distinct reward functions reveals relatively low degrees of similarity, with kappa values ranging between 0.22 and 0.35. This suggests that interactive narrative planning policies induced with MRL are likely to be sensitive to how reward functions are formulated. This is

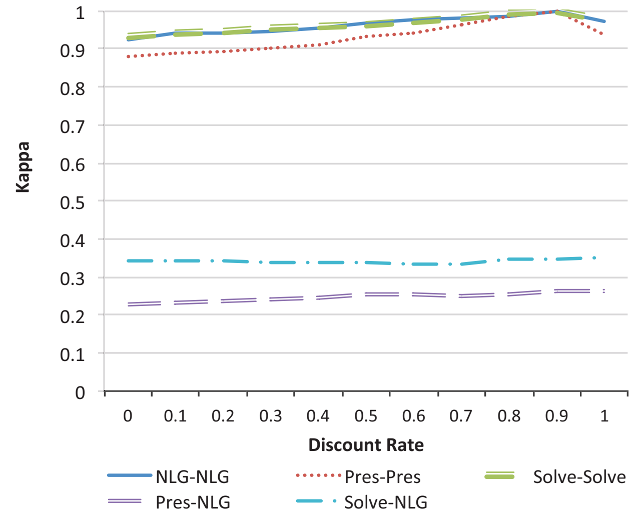

Figure 2. Cohen's kappa values showing inter-policy agreement by discount rate.

particularly notable, since many different reward functions can be engineered to encode the myriad dimensions of participants' interactive narrative experiences.

Conclusions and Future Work

We have presented a framework for interactive narrative planning that uses modular reinforcement learning to induce planners from player interaction data. The framework decomposes interactive narrative planning into multiple independent sub-problems, which are abstracted as adaptable event sequences. These AESs are modeled as concurrent Markov decision processes, with rewards based on players' experiential outcomes. Policies for solving the MDPs are obtained using model-based reinforcement learning. An empirical investigation of alternate parameterizations of reward found that interactive narrative planners induced with MRL are likely to be more effective than heuristic and baseline planners, and they are robust to alternate parameterizations of discount rate. Building on these findings, in future work we plan to systematically investigate the effects of alternate state representations on induced interactive narrative planners, as well as corpus size requirements for the framework. Furthermore, we plan to compare our MRL-based approach to other data-driven computational frameworks, including centralized planners induced with classical reinforcement learning techniques.

Acknowledgments

This research was supported by the National Science Foundation under grant IIS-1344803. 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.

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., Litman, D., and Jordan, P. 2011. Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Modeling and User-Adapted Interaction*, 21(1-2): 137–180.

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., Rowe, J. P., Mott, B. W., and Lester, J. C. 2014. A supervised learning framework for modeling director agent strategies in educational interactive narrative. *IEEE Transactions on Computational Intelligence and Artificial Intelligence in Games*, 6(2), 203–215.

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.

McCoy, J., Treanor, M., Samuel, B., Reed, A., Mateas, M., and Wardrip-Fruin, N. 2014. Social story worlds with Comme il Faut. *IEEE Transactions on Computational Intelligence and Artificial Intelligence in Games*, 6(2), 97–112.

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.

Orkin, J., and Roy, D. 2012. Understanding speech in interactive narratives with crowdsourced data. In *Proceedings of the 8th International Conference on Artificial Intelligence and Interactive Digital Entertainment*, 57-62. Menlo Park, CA: AAAI 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. 2013. Narrative-centered tutorial planning with concurrent Markov decision processes. Ph.D. diss., Department of Computer Science, North Carolina State University, Raleigh, NC.

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.

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

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.

Tetreault, J. and Litman, D. 2008. A reinforcement learning approach to evaluating state representations in spoken dialogue systems. Journal of Speech Communication, 50(8-9): 683–696.

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 Wasylishen, 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.

Witmer, B., and Singer, J. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments,* 7(3), 225–240.

Yu, H., and Riedl, M. O. 2014. Personalized Interactive Narratives via Sequential Recommendation of Plot Points. *IEEE Transactions on Computational Intelligence and Artificial Intelligence in Games*, 6(2), 174–187.