Drama Management Evaluation for Interactive Fiction Games

Manu Sharma, Santiago Ontañón, Manish Mehta, and **Ashwin Ram**
Cognitive Computing Lab (CCL)
College of Computing, Georgia Institute of Technology
Atlanta, Georgia, USA
{manus, santi, mehtama1, ashwin}@cc.gatech.edu

Abstract

A growing research community is working towards employing drama management components in interactive story-based games. These components gently guide the story towards a narrative arc that improves the player's experience. However, the success of drama management approaches has not been evaluated using human players in a real game implementation. In this paper, we evaluate our drama management approach deployed in our own implementation of an interactive fiction game Anchorhead. Our approach uses player's feedback as a basis for guiding the personalization of the interaction. The results indicate that our Drama Manager (DM) aids in providing a better overall experience for the players while guiding them through their interaction. Based on this work, we suggest that the strategies employed by the DM should take into account the player's previous playing experience with the current game as well as his general game-playing experience.

Introduction

There has been a growing interest in creating story-based interactive fiction games where the player is considered an active participant in the ongoing narratives. The component in charge of guiding the complete dramatic experience towards a particular narrative arc is called Drama Manager (DM) (Bates 1992) or Director (Magerko *et al.* 2004). The DM employs a set of actions provided at appropriate points in the ongoing game whereby the player is guided towards certain aspects of the story. Previous approaches to drama management have either not been connected to a concrete world (Weyhrauch 1997) or have been evaluated without using real human players interacting with a real game (Nelson *et al.* 2006). For successful future research in drama management, it is imperative that the developed approaches are evaluated in a real game and the results of the evaluations are used for guiding future research.

In this paper, we present an evaluation of our drama management approach (Sharma *et al.* 2007b) in an interactive story game called Anchorhead (Gentry 1998). We employ player experience as the guiding point for our evaluation. Anchorhead, created by Michael S. Gentry, is a game with a complicated story and several subplots, making it amenable for drama management studies. It has been previously used as a test bed for testing different drama management approaches (although with simulated players) (Nelson *et al.* 2006). In particular, our approach uses a case-based player modeling technique to predict the interestingness of game events for the current player. The DM uses this player model to decide interesting narratives for the player. Previously, we have shown that player modeling is a key component for the success of drama management based approaches (Sharma *et al.* 2007a). In order to test our approach, we have developed a concrete implementation for the Anchorhead game where the player interacts through a text-based interface (similar to one of the first interactive game, Zork). We created an intervention where the player is asked to play the game twice, once with our DM included as part of the game and then without the DM. To gather further DM design insights, we observed the player's interaction and discussed their subjective opinion at the end of each evaluation. Evaluation with real players has aided us in obtaining valuable information that we could not have noticed otherwise. Specifically, we acquired the following results and design guidelines for future research:

- The use of a DM provides a better overall playing experience. This is especially true for inexperienced players as the DM helps them progress in the game.
- Player modeling is a key factor for the success of DM approaches when dealing with real players for truly improving their overall experience.
- DM should take into account the player's previous gaming experience. Providing more guidance to an experienced game player makes the game less challenging for him.
- The player model should not only be built using the feedback on interestingness of intermediate game events but also on the player feedback on various strategies used by DM that were visible to the player during the interaction.

The rest of the paper is organized as follows. We will first present a brief introduction to the game we have used as our testbed, Anchorhead. Next, we introduce our technical approach to drama management, followed by the results from the player evaluations. Then we will summarize the related work in the fields of drama management and user modeling.

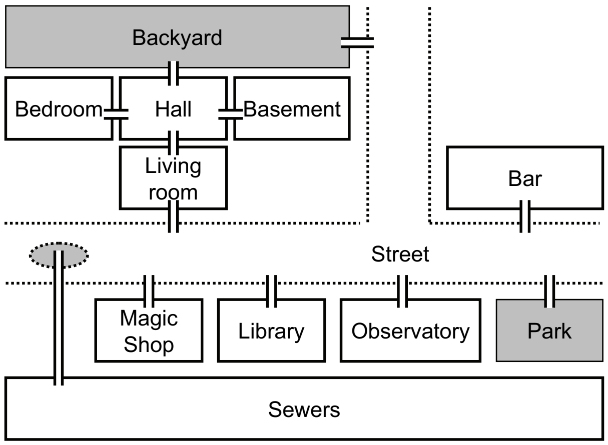

Figure 1: Map of the various locations in Anchorhead.

Finally, we conclude the paper with final thoughts and future directions.

Anchorhead

Anchorhead is an interactive story game created by Michael S. Gentry (Gentry 1998). We have developed a subset of the complete game and it includes a text-based interface for interaction with the player. Text-based descriptions of the current scenario are presented to the player, who then enters commands in textual format to interact with the game. These commands are essentially phrases that encode the players intentions, e.g. "enter the mansion", encoding the desire to enter into the a new game location (i.e. mansion).

The original story is divided in five days of action. For this paper we have focused on a subpart of the story, identified by (Nelson *et al.* 2006) as interesting for evaluating drama management approaches. The subpart is based on the second day of the original game. The resulting story has multiple subplots, that are related but can lead to two different endings.

The story takes place in the village of Anchorhead. The main character, i.e. the player, explores the village for clues about a mansion that was inherited and investigates about the odd history of the previous occupants, the Verlac family. Figure 1 shows the map of locations considered in our implementation of Anchorhead. During the game, the player is free to explore each location and interact with several natives, like the bum, the owner of the magic shop or the bartender.

Drama Management Approach

Our approach to drama management consists of three modules (shown in Figure 2), namely: a *game engine*, responsible for actually running the game and interacting with the player; a *player modeling module*, responsible for analyzing the actions of the current player and developing a player model; and a *drama management module*, influencing the development of the game and making it more appealing to the player.

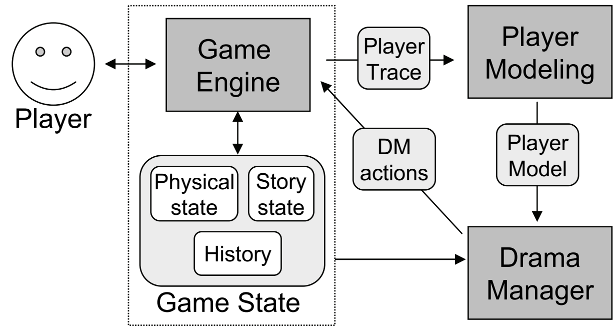

Figure 2: Basic scheme of the three modules that compose a game in our approach.

The game engine is responsible for running the game, maintaining the game state, enforcing rules of the game, interacting with the player, presenting game elements to the player and maintaining a *story state*, represented as a collection of events relevant to the game known as 'plot points' (Weyhrauch 1997).

Player Modeling Module

The player modeling module (PMM) builds and constantly maintains a player model for the current player of the game based on the feedback provided by players at the end of each game. This feedback contains player opinions on the game, including the parts they enjoyed and those that were not interesting from their perspective. The goal is to capture the interestingness rating for the story elements encountered by the player during the game episode. At the end of each interaction, the PMM stores this player feedback along with the corresponding trace of player actions during the game.

In particular, in our implementation, we use a case-based reasoning approach (Aamodt & Plaza 1994; Kolodner 1993) for the PMM module (show in Figure 3) that builds a *player preference model* that models the stories that the player is likely to enjoy. As part of its internal representation, it stores records of previous player interactions in the form of cases. These cases encode a combination of the history (i.e. the player trace) and an associated interestingness rating (explained below) elicited at the end of each game episode.

As a particular player is playing the game, his current trace is compared to the traces within the different cases stored in the case base. In order to facilitate calculating the similarity between these player traces, we have categorized the player actions. For example, the actions to move from one location to another are labeled as *movement*, others that have no effect in the story (e.g. having a nap) are classified as *no_effector*. In order to retrieve cases that are applicable for the current player, his trace is compared with stored player traces. We have used a similarity function that compares the histogram of the different kinds of actions in the corresponding traces.

In our experiments we retrieve the 5 most similar cases and use them as the estimate of the current player's preferences. The interestingness value for each plot point is then computed by aggregating the interestingness values from

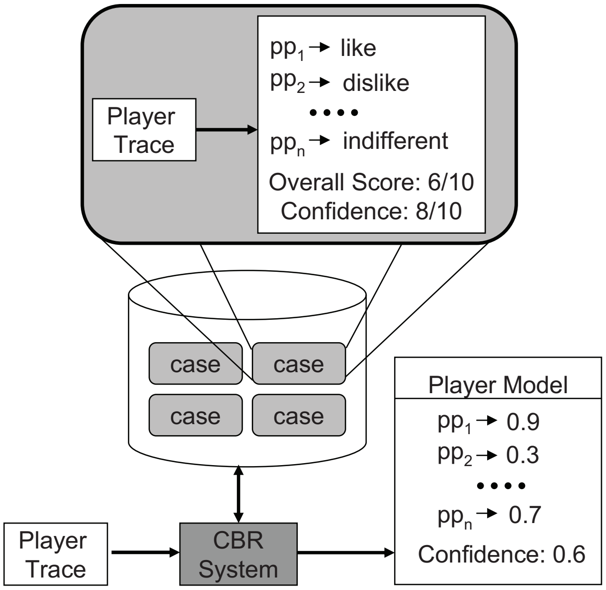

Figure 3: The Player Modeling Module uses a Case Based approach to build the player model.

those three cases by using the similarity metric to weight the individual interestingness values.

In order to develop cases for the player modeling module, players provide feedback at the end of each game. Specifically, we ask each player to fill out a short form. The player is presented with a sequence of the plot points visited in his game. From the list, the player is asked to select his preference of the plot points based on a 5 point scale classification: *strongly like*, *like*, *indifferent*, *dislike* and *strongly dislike*. In addition to this, he is asked to rank the game as a whole on a 5 point scale. The player also provides a confidence value on a 5 point scale. Notice that in our Anchorhead test system, the player model is a player preference model, and we are only modeling the interest of a particular player for each plot point. From each player feedback form, the system can build a case in the following way:

- Player annotations (like that of interest) for each plot point pp_j are converted to a number $\delta(pp_j)$ using the mapping: strongly dislike = -1, dislike = -0.5, indifferent = 0, like = 0.5 and strongly like = 1.
- The overall score provided by the player is converted to a number $s \in [-1, 1]$ by mapping 0 to -1 and 4 to 1.
- The confidence provided by the player is converted to a number $c \in [0, 1]$ by mapping 0 to 0 and 4 to 1.
- The interestingness of each plot point pp_j is computed as $ppi(pp_j) = \frac{\delta(pp_j)+s}{2}$, i.e. the average between the overall score and the particular annotation for that plot point.
- A new case consists of the player trace, the interestingness values for each plot point $ppi(pp_j)$, and the confidence c.

The output of the PMM is a player model that consists of the predicted interestingness of each plot point for the current player and also a confidence on this player model, as shown in Figure 3.

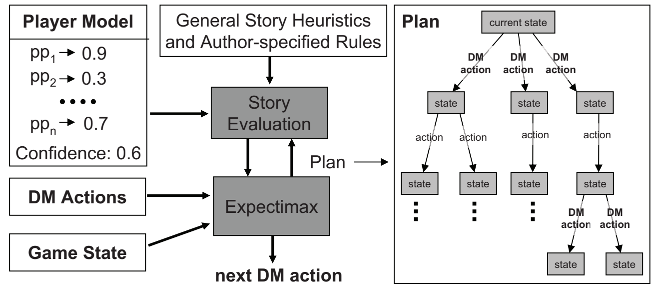

Figure 4: The Drama Manager Module consists of a planner and a story evaluation module.

Drama Management Module

Given the player preference model from the PMM, the current game state, and the author specified story guidelines, the Drama Management Module (DMM) plans story arcs with narrative coherence. Specifically, at every game cycle the DMM uses this information to select, if necessary, a particular action to influence the story.

In the same way that the game author specifies a set of actions that the player might take at each moment of time in the game (player actions), he also specifies a set of *drama manager actions* (DM actions). These actions represent the things that the drama manager can carry out to influence the game, e.g. "prevent the player from entering the library by locking the door" or "make the bar keeper start a conversation with the player about the suspicious person". Basically, the DM actions can be classified in two groups:

- *Causers*, which are designed to lead the player towards a particular direction in the story. Causers can be hints or direct causers, i.e. a hint to the player to go towards a particular direction, or directly make something happen so that the player cannot prevent from going in that particular direction.
- *Deniers*, which are designed to prevent the player to move towards a particular direction in the story. Again, deniers can be hints or direct deniers (the above example about locking the library's door is a direct denier).

To illustrate the kind of DM actions available in Anchorhead, consider the following example:

$$dma_3 = \begin{bmatrix} op = bum\_hints\_crypt\_key\_is\_in\_basement \\ p_l = park \\ p_{pp} = true \\ e_l = \emptyset \\ e_{pp} = \emptyset \\ h_a = \{obtain\_crypt\_key\} \end{bmatrix}$$

This DM action dma_3 causes one of the characters in the game, the bum, to tell the player that there is a key hidden in the basement. It is important for the player to find this key (hidden in the basement) to advance in one of the subplots. Specifically, this action is a *hint*, and h_a represents a set of player actions at which this DM action hints; i.e. after this DM action has been executed during the game, the player is more likely to choose an action from this set. As the game is going on, the DM will choose to execute this action if it realizes that by providing the key to the player, it will potentially cause the player to reach plot points that he will enjoy.

Given the set of possible directions in which the story might unfold (as a function of the player's selected actions), the drama manager has to plan a set of DM actions (causers and/or deniers) to guide the player towards story directions that are likely to be more appealing to him, according to the player module and author specified story guidelines.

In order to decide the DM action to executed at each moment of time, the DMM uses an *expectimax* method (Michie 1966). The starting node of the search tree is the current game state (see Figure 4). In the odd plys of the tree, each branch consists of a DM action (including a branch for doing no action). In the even plys of the tree, each branch consists of a player action. In our evaluation, we have kept a fixed depth of 5 plys so that the time required by the DMM to search is not appreciable by the player. For each leaf, l_j, of the tree, we compute an interestingness value $nodei(l_j) = c \times p(l_j) + (1 - c) \times a(l_j)$, where $p(l_j)$ is the average interestingness of the visited plot points in l_j (computed using the player model i.e. using $ppi(pp_k)$ from the cases closest to the state represented by l_j and contains plot points pp_k), $a(l_j)$ is the interestingness obtained using the author defined rules, and c is the confidence suggested by the player model. This essentially aids us in combining the author specified aesthetic values with those of the player model while searching for relevant DM actions.

The interestingness values are propagated up in the tree by selecting the maximum interestingness in the plys with DM actions and the average interestingness in the plys with player actions. Averaging is a good approximation as we do not have a concrete player action selection model. Moreover, if a hint DM action has been executed, then the subtree below it assumes that the hinted actions by that DM action has double the probability of being executed (another approximation due to lack of a player action selection model), and thus that is taken into account when averaging interestingness values in the player action plys. In the end, each of the DM actions in the first ply has an interestingness value (obtained by propagating up the interestingness as described above), and the DMM executes the DM action with maximum interestingness. If the maximum interestingness is given to the branch with no DM action, then no DM action will be executed. The result of this process is that the DMM selects the action that leads the story in a direction that would be more interesting for the player.

Evaluation and Analysis

We recruited twenty two participants (P1 . . . P22) with a range of genders (4 females and 18 males), races, education levels, and ages (from 22 to 37 with an average age of 25.5). Nine of these participants had absolutely no or low gaming experience. Each player was provided with an explanation on Anchorhead and asked to sign a consent form before starting the game. The player filled a background questionnaire to obtain information such as previous gaming experience or types of games they liked to play. The next step of evaluation was conducted in five phases:

Player	**Phase I**		**Phase IV**	
	No DM		**DM (26 cases)**	
	Rating	Confidence	Rating	Confidence
P1	3	4	3	4
P2	3	3	4	3
P3	2	2	1	3
P4	2	3	4	4
P5	2	3	3	4
P6	3	3	3	4
	Phase V		**Phase VI**	
	No DM		**DM (26 cases)**	
	Rating	Confidence	Rating	Confidence
P17	3	4	3	2
P18	3	4	4	4
P19	4	4	4	3
P20	2	3	1	3
P21	3	4	3	4
P22	3	3	2	3
Average	2.83		2.98	

Table 1: Overall game rating from participants in phases I, IV, and V, VI. A weighted average of ratings taking into account the confidence values is also shown. The ratings are on a 5 point Likert scale (0-4).

- In Phase I, P1 to P6 played the game once without the DM. By collecting feedback at the end of each game, we obtained six cases (C1 . . . C6).
- In Phase II, P7 to P11 first played without the DM and then with the DM (that used cases C1 . . . C6).
- In Phase III, P12 to P16 first played with the DM (that used cases C1 . . . C6) and then without the DM. The different orders in phases II and III helps in accounting for any possible discrepancy in results due to the order in which they play with or without the DM.
- In Phase IV, P1 to P6 played the game again with DM that used cases C1 . . . C26 (cases C7 to C26 were collected from phases II and III at the end of each game played).
- In Phase V, P17 to P22 played with the DM (using cases C1 . . . C26).
- In Phase VI, P17 to P22 played the game without the DM. Phases V and VI were conducted to account for any possible discrepancy in results due to the order in which players P1 and P6 played with and without DM in Phases I and IV.

In short, the evaluation has taken care of the following two aspects:

- Account for the order in which the DM was present in the game.
- Make participants play with different sizes of case library i.e. knowledge about the player preferences.

Player	Phase II			
	No DM		DM (6 cases)	
	Rating	Confidence	Rating	Confidence
P7	3	3	3	4
P8	3	4	3	4
P9	2	3	4	3
P10	3	4	3	4
P11	3	4	4	4
	Phase III			
	No DM		DM (6 cases)	
	Rating	Confidence	Rating	Confidence
P12	3	4	2	3
P13	2	4	3	4
P14	4	4	2	3
P15	3	4	3	4
P16	3	4	3	4
Average	2.92		3.03	

Table 2: Overall game rating from participants in phases II, III. A weighted average of ratings taking into account the confidence values is also shown. The ratings are on a 5 point Likert scale (0-4).

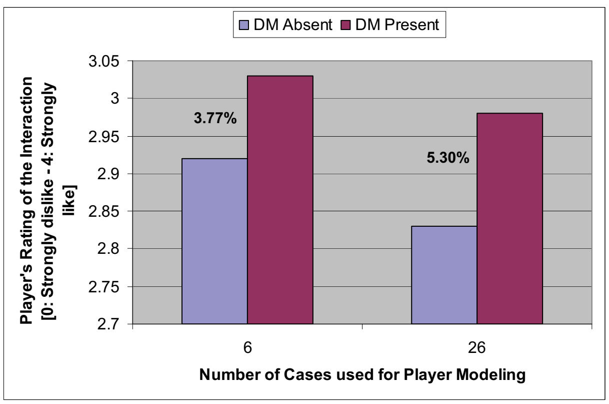

Figure 5: Diagrammatic representation of the increases in the interestingness ratings shown in Table 1 and Table 2

Thus, results have to be grouped like this: a first set of results coming out of merging phases II and II and another set of results coming out of merging phases I, IV, and V, VI.

During each episode, a researcher logged his observation of player actions and any unusual reactions. On an average, the complete player interaction (both game playing episodes) lasted for about 45 minutes each. At the end of each episode, the player was asked to provide an interestingness value and an associated confidence value on a 5 point Likert scale for the overall game experience as well as the intermediate story events that were encountered during the interaction. At the end of both gaming episodes, participants were interviewed about their experience.

We transcribed the player responses from the interviews and observed players' actions during the game episodes. We analyzed the data obtained from both the numerical and the qualitative perspectives (as we report in the following subsections). For the quantitative analysis we report on the score that players gave to the game after playing it (both with and without the DM). Moreover, in order to further understand the obtained results, we performed a qualitative analysis using a well-known qualitative analysis method, Grounded Theory (Strauss & Corbin 1990). Using grounded theory principles, we made notes for each player action during the game play and their responses from the interview, iteratively developing a base set of categories. We used these categories to tag game play observations and interview responses. Tagged phenomena were then organized into higher-level concepts by grouping related categories into a single concept. A subset of these categories are shown in Table 3.

Quantitative Results

Tables 1 and 2 show the results from our quantitative analysis. Specifically, the tables show the rating that each player gave to the interaction after completion of the game accompanied by the confidence in their rating. The first trend that we can observe from the tables is that the average rating increases when playing with the drama manager: 2.98 with DM versus 2.83 without DM (for players from Phases I, IV and V), and 3.03 with DM versus 2.92 without DM without DM (for players from Phases II and III). This shows that the players preferred to play Anchorhead with our drama manager. These results can be better visualized in Figure 5. Also, as the number of cases increase, i.e. the knowledge about player preference increases, from 6 to 26, the percentage increase in the interestingness from no DM to DM increases from 3.77% to 5.3%. In Figure 5, it is important to note that there is no good way to compare the two sets of bars corresponding to the 6 and 26 cases respectively. The ratings provided when the DM included 6 cases belong to players P1 to P6 and the ratings provided when the DM included 26 cases belong to players P7 to P16. Hence, the drop in the ratings from 6 to 26 cases has no relation to the performance of the DM.

We observed that the players enjoy much more when they play the second time. Thus, it was important to perform experiments where some players played first with DM while some other played first without DM. Moreover, this shows that the first time that the participants played the game, they were slightly lost. This is when the DM helped them the most (we elaborate more about this in the qualitative results section). Next section analyzes the results from a qualitative perspective, taking into account the comments given by all the players that participated in the evaluation.

Qualitative Results

From the qualitative analysis, we draw four main conclusions, explained in detail in the following subsections.

DM improves overall player experience, especially for inexperienced gamers Apart from the overall numeric rating increase obtained through quantitative analysis (Tables 1 and 2), participants with little or no gaming experience (we call them non-gamers) explicitly commented that the "hints from the DM were particularly helpful" (P7). Analysis of observations made during the game play further showed that

Category	Example Player Comments
DM helped Player	*"Would have never thought (unless told by DM) that picking up the skull would ever be useful"*
With DM then Without DM	*"Understood the help I was getting from DM only after I played the game without it"*
Player liked a plot point	*"I enjoyed solving the puzzle to find a magic lens"*
Liked the game	*"Game allows you to discover game elements at ones own pace"*
Game Endings	*"I felt that I could influence the particular outcome in the game"*
Lack of Strategy	*"I did not follow any specific strategy in playing the game"*

Table 3: A subset of categories used to categorize player responses to the post game questionnaire.

non-gamers found themselves lost in the game on multiple occasions. Sometimes they struggled finding or using existing clues in the game. Non-gamers felt that the hints from the DM helped them in collecting items so that they could use them in other scenarios in the game (see comment by P4 in Table 3 under 'DM helped Player').

Some non-gamers felt frustrated when playing the game without the DM, since without hints they had troubles advancing in the game. For instance, one player (P15) asked if he could quit the game when playing without the DM since he had been wandering for a while without knowing how to further advance in the game.

Finally, some players realized the importance of the DM when they played the game without the DM. For instance, player P18 played first with DM and then without DM. During his first game, he did not realize that some of the hints that he was receiving were actually generated by the DM. He assumed that the hints were part of the game itself (which in itself is good, since points out that the DM was not too intrusive). Later, when playing without the DM, P18 realized that the DM interventions were important when he found himself playing a game without any hints or guidance and found the experience lacking.

DM strategies should depend on game expertise Our current DM strategies do not take into account player's previous game playing expertise and the number of times the players have played Anchorhead before. Experienced gamers' comments on aspect indicate that the hints provided by DM were sometimes "too obvious" (P3) and should be "more subtle" (P5). However, as indicated in the previous subsection, non-gamers felt that hints were really useful to them. These results suggest that the text corresponding to a particular hint should be designed differently depending upon the player's previous game-playing expertise. One intuitive observation from the players' interactions was that they were much more comfortable playing the game the second time. They felt that "playing the first game helped them while playing the second time" (P10) and they had better sense of the context and the goals of the game.

The easiest way to collect the above information is to ask the player, at the start of the game, his gaming experience over some scale and whether this is his first session with Anchorhead. This information, either automatically learned or directly asked to the user, should be included in the player model that the DM uses to plan story arcs.

Player model should be richer After relating the player's interview responses and their feedback on plot points, an important observation is that the player model should be enriched to take into account the player's liking for hints as well. Certain players liked a particular plot point, but not the hints presented to them at those plot points. Consider this scenario from the evaluation study: After encountering a puzzle and having an interesting experience while solving it, a player provides positive feedback for the plot points associated with the puzzle. Later, if another player with similar playing characteristics reaches that puzzle and receives a hint from the DM that eases the solving process, he would not mark the plot points associated with the puzzle as interesting.

If the feedback at the end of the game includes an opinion on the DM actions (hint, in this example), the player model can avoid situations like the one mentioned above. This means that the next player that has similar playing characteristics would indeed be directed towards the story arc containing the puzzle and would not be presented with the particular hint.

Another situation found in our evaluation supporting this was a comment given by player P21. Specifically, in a particular game situation the DM decided to use a denier to prevent P21 to perform an action that will let to an ending that P21 might dislike. The denier was so obvious that P21 felt manipulated, commenting that he would have preferred a hint towards the other ending, instead of a denier. If the player model would have been able to predict that P21 didn't like deniers, it would have been able to provide the hint instead of the denier.

DM strategies should depend on player's strategy We observed that Non-gamers do not follow any particular strategy. They are good followers of the hints from the DM, and thus remain unsure if they actually influenced the ending of the game. They enjoy interactions with non-playing characters relatively more, e.g., bribing the bum, talking with the owner of magic shop. Experienced gamers intelligently pick items that they feel would be useful later in the game.

Gamers admit that a drama manager can be very useful in much larger games where a strict breadth-first search in the game cannot reach the goal of the game (without wasting an impractical amount of time). A typical strategy used by our participants (gamers or non-gamers) was to explore a lot by going to different locations in the map. As a design insight for future, we want to incorporate this observation in DM design. When the DM observes that the player actions are not doing anything logical to reach a subgoal for some finite amount of time, it can provide a hint with the assumption that the player is confused. If the player was indeed trying to simply explore and did not like this hint at that game instance, he would provide a negative feedback for the particular hint (with the mechanism explained in the previous subsection). Later, if another player has similar playing characteristics, he would not be given the same hint during those game events.

Related Work

Bates (1992) first proposed the idea of treating drama management as an optimization problem. The approach termed Search Based Drama Management (SBDM) was based on the fact that the drama manager chooses its best available action with expectation-calculating nodes and the player is typically assumed to act according to some probabilistic model. Peter Weyhrauch (1997) in his dissertation work further developed the idea of a SBDM with a game tree based search that used an author specified evaluation function to measure the interestingness value for a particular story. However the DM employed was not connected to a concrete story world and the techniques were tested using simulated players rather than real human players. The approach furthermore ignored a player preference model capturing the players preference for a particular story. In another approach, Nelson et. al. (2006) define a Declarative Optimization based approach to Drama Management (DODM). The central premise of their technique is to provide the author with the ability to specify what constitutes a good story and use a reinforcement learning approach to optimize DM actions in response to player actions, given a set of plot points and the evaluation function. This approach also uses a simulated player model to predict the next player action. Furthermore, the approach ignores a player preference model to measure the interestingness of a story from the player's perspective.

In our approach to drama management, we construct a player preference model through real human player interaction with the game. Whereas previous approaches employed only an author based evaluation function for story interestingness, our approach essentially constructs an evaluation function that is a combination of both player and author.

Another approach to drama management is that of Façade (Mateas & Stern 2003), which employs a beat based management system suited towards tighter story structures where all the activity contributes towards the story. The Mimesis architecture (Young *et al.* 2004) proposes a story planning based approach for real-time virtual worlds. In this approach, the story plans are tagged with causal structure and the system handles player actions that might threaten the causal links through either replanning the story or disallowing the player the opportunity to carry out his action. In such an approach to drama management, however, only the author specifies the concrete goals that the planner should achieve; the approach doesn't incorporate a player preference model to guide the player experience during the game episode.

Another related area to our research is player modeling. Player modeling is generally accepted as a prerequisite towards achieving adaptiveness in games (Houlette 2004; Charles & Black 2004). Different approaches towards player modeling can be classified in two groups, namely:

- *Direct-measurement* approaches, that employ physiological measures to directly monitor player's emotional state during the game playing episode.
- *Indirect-measurement* approaches, that try to infer (in opposition to directly monitoring player's emotional state) information about the current player (e.g., skill level, and preferences) by computing a set of features from the playing pattern during the interaction in the game episode.

An example of the former, is used in the approach of (Prendinger, Mori, & Ishizuka 2005) where sensed data is used to modify the behavior of an empathic virtual agent situated in a job environment. In our approach towards player modeling, indirect measurements were better suited as the game is currently a text-based interaction, where emotional reactions might not be as rich as in a 3D environment. Further, we were interested in modeling the player from the data that can be derived from the actions taken in the game. Previous work on indirect-measurement techniques for player modeling focuses on modeling the player's skill for automatic adjustment of game level. (Cowley *et al.* 2006) present a decision theoretic framework to model the choices that players make in the well known *pacman* game. They observe the player's deviation from the optimal path and use that to model the player's skill level. One of the results from the evaluation of our DM, presented later, suggested that the skill level of the player is an important measure for a better playing experience. However, our current approach models player preferences and not skill level.

(Togelius, Nardi, & Lucas 2006) present a player modeling technique applied to a racing game. The player models captures the behavior of a player for a given track. Instead of generating hand-made tracks, they use the learned models to automatically generate new tracks that exhibit similar characteristics (speed achieved, difficulty, etc.) when the learned player models are used to drive in the tracks. In their work, they build *player-action* models (i.e., modeling the behavior of the players) whereas we focus on modeling the preferences of the player to provide a better playing experience.

(Yannakakis & Maragoudakis 2005) present results on the usefulness of having a player model. Specifically, they define a set of rules by which the game *pacman* can be considered interesting, and an evolutionary algorithm for the behavior of the enemies in the game that is able to make the game interesting. In that framework, they show the usefulness of a player model to help the evolutionary algorithm to achieve more interesting games (based on the predefined set of rules). In our work, we focus on obtaining the interestingness from the player feedback instead of defining a set of rules. The feedback is in form of an overall game rating, confidence on the rating and a measure of liking for intermediate game events encountered during the interaction.

Conclusion

Previous approaches to drama management have not been evaluated using real game players; also they do not build player models during real game play. In this paper we have presented an evaluation of a drama management approach using real players interacting with a text-based interactive fiction game Anchorhead. Our results show that having a drama manager makes the player, especially ones with no or low gaming experience, have a better playing experience. Specifically, we found that player modeling is a key factor in drama management. Our system also had author-specified story interestingness rules, but the experiments show that for drama management to be successful with real users, good player modeling was required.

Moreover, the evaluation with real users provided us with several insights extremely useful to guide our future research. We found that the strategies used by the DM to guide the player should take into account player's previous game playing experience and number of previous interactions with the given game. Further, the hints used by the DM should be more subtle, rather than explicitly directing the players towards a particular narrative arc. As a future step, we plan to incorporate these design insights into our drama management technique as well as applying our approach to a more complex game with a three-dimensional immersive interface.

Acknowledgments

The authors would like to thank all the twenty two participants of our Anchorhead drama management evaluation.

References

Aamodt, A., and Plaza, E. 1994. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Communications* 7(1):39–59.

Bates, J. 1992. Virtual reality art and entertainment. *The Journal of Teleoperators and Virtual Environments* 2(1):133–138.

Charles, D., and Black, M. 2004. Dynamic player modeling: A framework for player-centered digital games. In *Proc. of the International Conference on Computer Games: Artificial Intelligence, Design and Education*, 29–35.

Cowley, B.; Charles, D.; Black, M.; and Hickey, R. 2006. Using decision theory for player analysis in pacman. In *SAB'06 Workshop on Adaptive Approaches for Optimizing Player Satisfaction in Computer and Physical Games*, 41–50.

Gentry, M. S. 1998. Anchorhead. available online at http://www.wurb.com/if/game/17.html.

Houlette, R. 2004. Player modeling for adaptive games. In *AI Game Programming Wisdom II*. Charles River Media. chapter 10.1, 557–566.

Kolodner, J. 1993. *Case-Based Reasoning*. San Mateo, CA: Morgan Kaufmann.

Magerko, B.; Laird, J.; Assanie, M.; Kerfoot, A.; and Stokes, D. 2004. AI characters and directors for interactive computer games. In *Proceedings of the 2004 Innovative Applications of Artificial Intelligence Confercence*.

Mateas, M., and Stern, A. 2003. Integrating plot, character, and natural language processing in the interactive drama faade. In *Proceedings of the 1st International Conference on Technologies for Interactive Digical Storytelling and Entertainment*.

Michie, J. 1966. Game-playing and game-learning automata. In Fox, L., ed., *Advances in Programming and Non-Numerical Computation*, 183–200.

Nelson, M.; Mateas, M.; Roberts, D.; and Isbell, C. 2006. Declarative optimization-based drama management in interactive fiction. *IEEE Computer Graphics and Applications* 26(3):33–41.

Prendinger, H.; Mori, J.; and Ishizuka, M. 2005. Recognizing, modeling, and responding to users affective states. In *User Modeling 2005*, number 3538 in Lecture Notes in Computer Science. Springer-Verlag. 60–69.

Sharma, M.; Mehta, M.; Ontañón, S.; and Ram, A. 2007a. Player modeling evaluation for interactive fiction. In *Third Artificial Intelligence for Interactive Digital Entertainment Conference (AIIDE), Workshop on Optimizing Player Satisfaction*. AAAI Press.

Sharma, M.; Ontañón, S.; Strong, C.; Mehta, M.; and Ram, A. 2007b. Towards player preference modeling for drama management in interactive stories. In *Twentieth International Florida Artificial Intelligence Research Society Conference (FLAIRS)*, 571–576. AAAI Press.

Strauss, A., and Corbin, J. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage.

Togelius, J.; Nardi, R. D.; and Lucas, S. M. 2006. Making racing fun through player modeling and track evolution. In *SAB'06 Workshop on Adaptive Approaches for Optimizing Player Satisfaction in Computer and Physical Games*, 61–70.

Weyhrauch, P. 1997. *Guiding Interactive Drama*. Ph.D. Dissertation, Carnagie Mellon University.

Yannakakis, G. N., and Maragoudakis, M. 2005. Player modeling impact on players entertainment in computer games. In *User Modeling 2005*, number 3538 in Lecture Notes in Computer Science. Springer-Verlag. 74–78.

Young, R.; Riedl, M.; Branly, M.; Jhala, A.; Martin, R.; and Sagretto, C. 2004. An architecture for integrating plan-based behavior generation with interactive game environments. *Journal of Game Development* 1(1).