

Adapting IRIS, a Non-Interactive Narrative Generation System, to an Interactive Text Adventure Game

Matthew William Fendt and R. Michael Young

Digital Games Research Center
 Department of Computer Science
 NC State University
 Raleigh, NC 27695-8206

Abstract

The IRIS story generation system is a narrative planner that is designed to create suspenseful stories at the story-action level. The system has been previously shown to be able to create suspense in a non-interactive context. We show that, when a series of transformations on the original IRIS output are applied, the IRIS framework is also effective at creating suspense in the context of an interactive text adventure game.

Introduction

Historically, video game narrative has been written by hand by the authors during the game’s creation. There has been interest in creating adaptive games that change based on player choices. Instead of having authors enumerate and the game designers hard code these different story branches, adaptive computer systems can be used in the creation of game content to alleviate authorial burden.

The IRIS narrative generation system (Fendt and Young 2011) is a tool used to generate story outlines with built-in suspense that can be instantiated in different media. An overview of the IRIS generation process is shown in Figure 1. IRIS was shown to be able to create localized moments of suspense at the story-action level comparable to a human author composing story fragments from the same set of actions in the experimental domain (Fendt, Roberts, and Young 2013). For this paper, the IRIS system was extended to the domain of a text-based adventure game. The original output of IRIS is a story outline in a non-interactive context which, following two transformations, was then instantiated into an interactive game. The two transformations are 1) an in-game method of giving hints of how to achieve the game’s goals and 2) the presence of the suspenseful action sequences from the original story outline and the restriction of player agency during their execution. We hypothesized that both of these transformations were necessary to preserve IRIS’s sense of suspense in an interactive context. In our experiment, participants played one of four versions of a text adventure video game in which one, both, or neither of the transformations were used. The IRIS story outline used in the game had three suspenseful goals. After the game, participants answered survey questions that were created by the

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

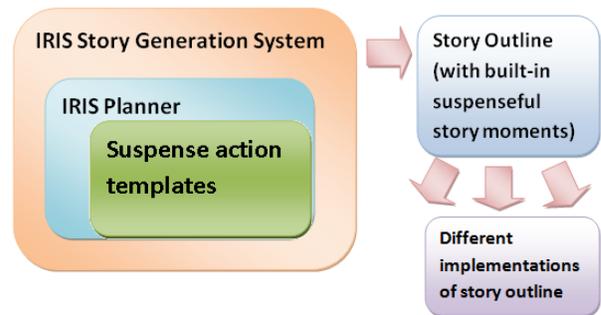


Figure 1: The IRIS story generation system. The generation process creates story outlines containing suspense that can then be instantiated in different media.

operationalization of IRIS’s definition of suspense. These questions measured participants’ sense of suspense during the execution of the three suspenseful goals and several other non-essential, non-suspenseful goals. It was found that the version of the game containing both transformations was rated higher in suspense than each of the three other versions of the game for two of the three suspenseful goals. This validates IRIS as a viable tool in creating suspense in a text adventure game when both of the transformations are applied.

Related Work

Narrative generation systems have been designed to evoke a wide range of narrative phenomenon. One such system controls an interactive narrative where the player is presented with archetypal dilemmas such as Betrayal and Sacrifice (Barber and Kudenko 2008). The system adapts future dilemmas to past player behavior. Another system, CPOCL (Ware and Young 2011), uses planning to generate a story in which character conflict is deliberately introduced into the planning process. Facade (Mateas and Stern 2003) is an interactive drama where the concept of “beats,” or narrative pacing, is explored. TALE-SPIN (Meehan 1977) was an early story generation system that created fables that conveyed a certain moral. These are just some of the examples of different narrative phenomena that have been explored.

There are a few existing systems that have explored sus-

pense in narrative generation. MINSTREL (Turner 1994) is a case-based reasoning system that creates stories that have, among other properties, suspense. It creates suspense at two levels: the story action level and the discourse level. The story action level refers to the selection and ordering of the actions themselves, while the discourse level refers to the presentation and medium of the story actions. MINSTREL has a rigid framework for when to create suspense, namely, in a scene that the author wants to emphasize and by putting a character's life in danger. This is an overly simplistic method of creating suspense, likely because many different narrative phenomena were present in the system and suspense was not the focus of the work.

Suspenser (Cheong and Young 2008) may be the closest relative of the IRIS system. Suspenser is a planning system that takes complete story outlines and iteratively refines the story to increase the story's suspense at the discourse level. However, since Suspenser generates suspense at the discourse level, it may not be generating stories that are intrinsically suspenseful, but rather adding suspense to the stories after they have been generated. Suspenser also is not guaranteed to generate high suspense stories, since it can get stuck in a local maximum when adding suspense. The discourse-level suspense created by Suspenser is complimentary to the story-action level suspense the IRIS system creates. In future work, the two approaches could be combined; however, that is beyond the scope of this paper.

Transitioning From a Non-Interactive to an Interactive Context

IRIS (Fendt and Young 2011) is a narrative planner that uses a Drama Manager and belief/desire/intention mental modeling with intention revision to produce suspenseful stories. The definition of suspense that IRIS uses is "the feeling of excitement or anxiety that audience members feel when they are waiting for something to happen and are uncertain about a significant outcome" (Cheong and Young 2008). The suspense-creating components of the planning algorithm were shown to be able to create localized moments of suspense at the story-action level comparable to a human author composing story fragments from the same set of actions in the experimental domain (Fendt, Roberts, and Young 2013). In this paper, we show that the non-interactive narrative output of the IRIS system can be translated into more robust narrative experiences, in particular an interactive text game, while preserving the sense of suspense present in the initial outline. To preserve this suspense in an interactive experience, two transformations need to be made: 1) an in-game method of giving hints of how to achieve the game's goals and 2) the presence of the suspenseful action sequences from the original story outline and the restriction of player agency during their execution.

The translation from a non-interactive story outline to an interactive experience that preserves the story outline's sense of suspense is difficult for at least two reasons. First, the simplifying assumption in IRIS where the audience knows the complete state of the world at all times does not translate to an interactive narrative environment. Second, if given the

choice, players in an interactive environment might knowingly or unknowingly take actions that interfere with the suspenseful situations that the author created in the outline.

However, there is support that such a translation can be made. In non-interactive contexts, suspense is largely created around the protagonist, who is at the forefront of the story. In an interactive game, the player will serve as the protagonist, so any suspense created will be experienced by him or her. Also, the player's role in the creation of an interactive experience is analogous to an author's role in the creation of a non-interactive narrative. Both can create a rich narrative experience by exercising authorial creativity with regard to action ordering, addition of narratively interesting but non-essential actions, and discourse choices while still adhering to restrictions placed on them by a set of rules from an external system.

There are two transformations that need to be made when translating from the non-interactive outline to an interactive game in order to preserve suspense. The first consideration is that when the player's plan fails, they need to have sufficient information about how to replan to achieve their goal. This ensures that players will not abandon plans around which suspense is being created. One important component of a helpful hint requires modeling the player's knowledge. Instead of having complete knowledge of the world at every point in the story, the player's knowledge can be modeled as containing any introductory information given prior to the story, plus the effects of every action that takes place in the player's location. This is a simplifying assumption since it assumes the player knows the effects of every action and information is uniformly salient. Since the player's knowledge is modeled, the system will know what new information it needs to introduce at a given point in the story to allow the player to continue on a new plan. Assumptions include that the player will know how to use the new information and that there is a non-immersion breaking way to introduce the information to the player.

The player needs to have a level of agency in the game to create the interactive experience, but at certain points in the story their level of control needs to be reduced to ensure that he or she experiences the suspense that the system is trying to create. The player should not be able to take actions that will bypass the suspenseful action sequences in the story. Small details, for example the actors or location of the event, could be altered, but the actions themselves and their ordering need to be preserved. This may require the system to prevent the player from performing an action if it would bypass the suspenseful situation or to guide the player to take actions that will trigger the suspense. Also, during the execution of the suspenseful sequence of actions, the player needs to have reduced agency with respect to the actions that can be performed. The player should not be allowed to walk away or take extraneous actions during these suspenseful moments. This restriction will be modeled as a dynamic list of actions available to the player that will be shown in the game interface.

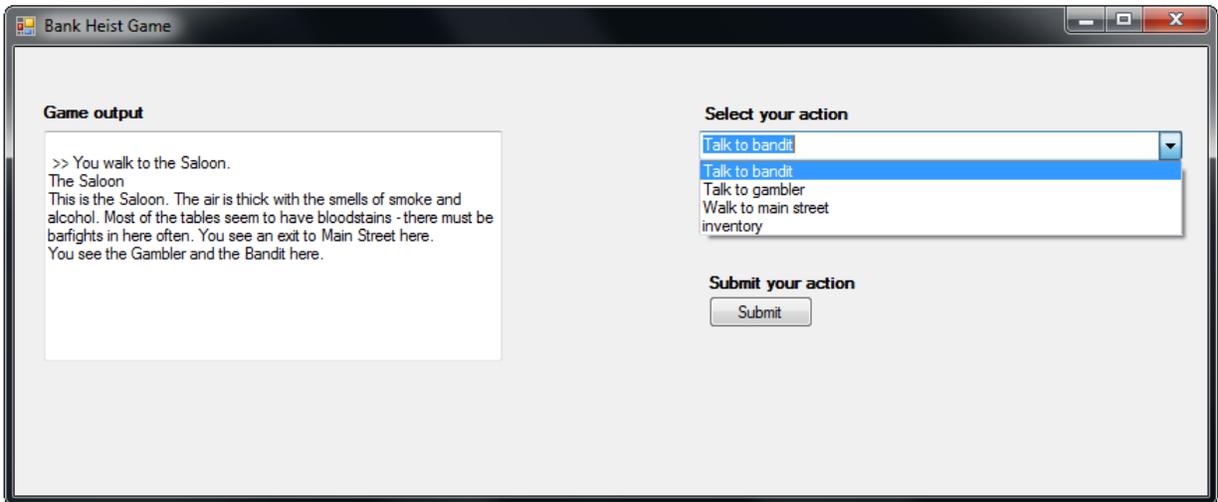


Figure 2: A screenshot of the GUI that the participants used to play the game.

Western Heist Game Domain

We provided the IRIS system with a western heist domain where the protagonist is a bank robber. A visual representation of the game world is shown in Figure 3. Actions allowed in the domain included move, talk, pick up, give, attack, steal, unlock, blow up, and capture. The important items to gather were the key and the dynamite, which allowed the protagonist to make it past the two locked doors on the way to the treasure room. These items were available at several different places in the world. Once in the treasure room, the protagonist could take the loot and try to leave town. There were numerous characters in the world that could aid or hinder the protagonist and some characters present just to provide background flavor for the town.

The protagonist’s objective is to get the loot out of the vault. The IRIS system identified subgoals in the story where suspense should be introduced. The subgoals are 1) getting the dynamite used to blow up the vault door, 2) getting the key used to open the vault safe, and 3) moving the loot from the vault to the train. There were multiple ways to accomplish these subgoals. To get the dynamite, the player could convince the store owner to give him or her a stick or dynamite, or it could be found when exploring the mines. To get the key, the player could steal it from the bank teller or sneak into his room and steal it. When the player gets the loot, another bandit steals it from you. To get it back, the player can attack the bandit or try to negotiate.

Experimental Design

This story outline was implemented into a single player Multi User Dungeon (MUD) environment called LambdaMOO. In the game, a participant plays as the bank robber with the objectives described above. A GUI was developed in C# to aid in game play (see Figure 2). The GUI presents a restricted subset of available player actions at a given time, as per transformation 2 described above. The order in which

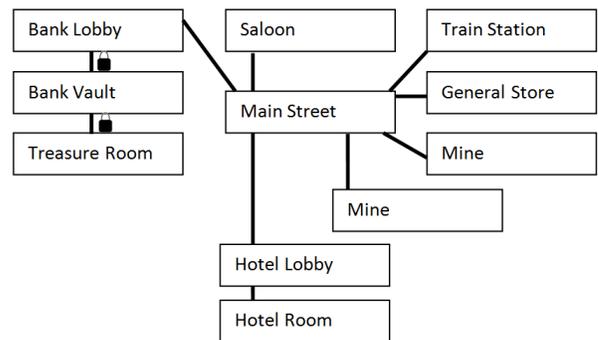


Figure 3: A visual representation of the western heist game world.

the actions were presented was randomized after every action selection to prevent priming bias.

Participants were recruited from Computer Science undergraduate classes. There were activities that the students could perform to receive extra credit in the class, and this experiment was one option. The participants came to the researcher’s lab for the study where they were first given a consent form to read and sign. Next, they played the game. Play time was about 10-15 minutes. After the game was over, they were given a survey to measure their sense of suspense around each available game goal. These goals included the three suspenseful goals as identified by IRIS and several others that were not targeted as suspenseful but were made available to increase the number of things to do in the world. Participants rated their agreement on a five point Likert scale with the five survey questions that were created from the operationalization of IRIS’s definition of suspense. These questions are listed in Figure 4. A screenshot of the survey is shown in Figure 5. For the survey, the order of the questions regarding the game goals and the questions within

Question 1: I felt excited about the outcome of the goal
 Question 2: I felt anxious about the outcome of the goal
 Question 3: I felt like an important story event was about to happen
 Question 4: I was uncertain about the successful completion of the goal
 Question 5: I felt that the outcome would be significant to future events in the story

Figure 4: The survey questions presented to the participants in the post game survey. These questions were generated by an operationalization of IRIS’s definition of suspense.

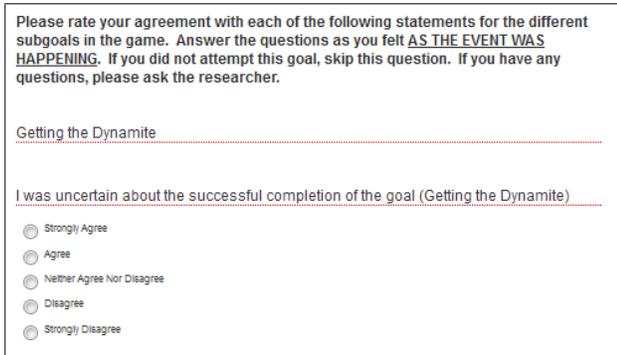


Figure 5: A screenshot showing an example question from the post game survey presented to the participants.

each game goal were randomized for each participant.

A central hypothesis, based on the two proposed transformations described above, was tested:

Hypothesis Both 1) an in-game method of giving hints about the ways to achieve the game’s goals and 2) the presence of the suspenseful action sequences from IRIS’s original story outline and the restriction of player agency during their execution, are necessary to create suspense in an interactive text adventure implemented using IRIS’s narrative output.

Four different versions of the game, using none, one, or both of the transformations, were created. Participants were randomly assigned to the different versions. Version 1 contained no game hints or suspenseful sequences, version 2 contained game hints but no suspenseful sequences, version 3 contained no hints but suspenseful sequences and reduced agency during their execution, and version 4 contained both hints and the suspense sequences. We anticipated that suspense in the version with both transformations would be rated higher in the survey than all of the other three versions.

Experimental Results

64 people played the game and all of them completed the post game survey. 17 people played version 1 (- hints/ - suspense) and version 2 (+ hints / - suspense), and 15 people played version 3 (- hints / + suspense) and version 4 (+ hints/ + suspense). Using a series of one-tailed Mann-Whitney U tests, version 4 was compared to the other three versions.

The other versions were compared among each other using two-tailed Mann-Whitney U tests. This was because no claims about one of these versions being rated higher than the others were made. Mann-Whitney U was used because it tests to see if a measurement from a population of ranked, non-parametric data is significantly different from another. The results of these tests are displayed in Table 1.

The first two player goals could be completed in either order. To see if the order of the completion affected player responses in Version 4, a two-tailed Mann-Whitney U test was performed comparing the responses for the first and second goals that the players performed. There was no significant difference in responses for survey questions 2-5, but players rated the second goal that they performed as significantly more suspenseful ($p = 0.04$) for question 1: “I felt excited about the outcome of the goal.”

Experimental Discussion

Figure 6 shows a visualization of the Mann-Whitney U tests for the key and loot goals. There are two ways to view the results: compare responses across versions or across questions. When viewing across versions, Version 4 had the most improvement over the versions in this order: 1, 3, 2. This is interesting because it suggests that the presence of the suspense sequences (transformation 2) alone is not sufficient to evoke a high level of suspense without also providing game hints (transformation 1).

When viewing across questions, Version 4 had the most improvement over the other versions in questions 1, 2, and 5. This is encouraging because questions 1 and 2 both directly measure the player’s sense of emotion while playing the game, which the full treatment seemed to increase. The lower improvements in questions 3 and 4 may have been caused by player genre expectations. Question 3 was “I felt like an important story event was about to happen.” Since players were given these three goals explicitly, they may have felt that, across versions, completing these goals would advance the story. Question 4 was “I was uncertain about the successful completion of the goal.” Again, across versions, players may have had the genre expectation that if they are being asked to complete a goal, than there will be some way to complete it.

There was not a significant improvement in Version 4 with regards to suspense for the goal of obtaining the dynamite. This may be because the actions needed to complete the goal might not have been amenable to plan failure, the method IRIS uses to create suspense. The steps to get the dynamite may have seemed to contain less risk than the steps required to get the key or the loot, thus decreasing the overall sense of suspense when the player’s plan was thwarted. This suggests that in addition to applying the IRIS formula and the transformations needed to make the experience interactive, the goals that are targeted as suspenseful need to contain high risk actions that will concern the player when plan failure is introduced by the system.

Version		4v1	4v2	4v3	3v2	3v1	2v1
Key	Q1	0.006	0.01	0.003	0.66	1	0.69
	Q2	0.34	(0.06)	0.03	0.87	0.12	0.19
	Q3	(0.09)	0.17	(0.07)	0.58	0.94	0.66
	Q4	(0.08)	0.21	0.54	0.35	0.12	0.47
	Q5	0.05	1	0.03	(0.06)	0.88	(0.08)
Dynamite	Q1	0.05	0.14	0.14	0.75	0.81	0.46
	Q2	0.28	0.85	0.16	0.009	0.68	0.05
	Q3	0.03	0.97	0.96	1	0.001	0.0007
	Q4	0.23	0.49	0.23	0.48	0.95	0.48
	Q5	0.44	0.96	0.95	1	(0.10)	(0.08)
Loot	Q1	0.007	(0.09)	0.15	0.71	(0.08)	0.13
	Q2	(0.10)	(0.06)	(0.06)	1	0.67	0.67
	Q3	0.32	0.18	0.84	0.12	0.21	0.64
	Q4	(0.07)	0.19	0.002	0.04	0.19	0.56
	Q5	0.04	0.02	0.15	0.28	0.48	0.59

Table 1: The results of the Mann-Whitney U tests comparing different versions of the game for each of the three potentially suspenseful goals. The one-tailed Mann-Whitney U test was used in the comparisons involving Version 4 and the two-tailed test for comparisons not involving Version 4. The values in bold indicate significant difference with regards to reported suspense for that survey question, $p = 0.05$, and the values in parentheses indicate marginal significance, $p = 0.10$.

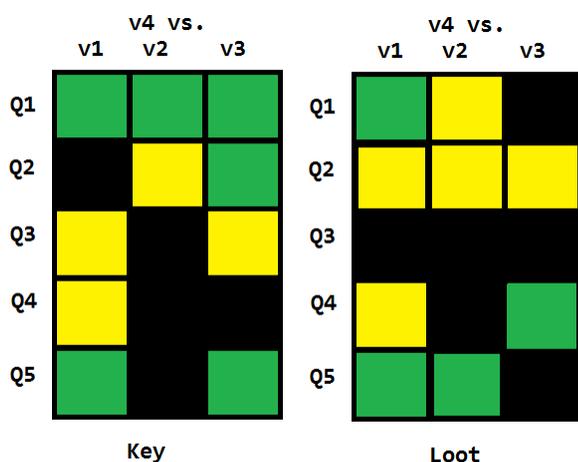


Figure 6: A visualization of the results of the one-tailed Mann-Whitney U tests comparing Key and Loot goals in Version 4 to those goals in the other three versions. A green box indicates that Version 4 was rated significantly higher with regards to suspense for that survey question, $p = 0.05$, yellow indicates marginal significance, $p = 0.10$, and black indicates no significant difference.

Future Work

There are several items that we would look to expand upon in future work. The IRIS narrative generation system has a model in which it chooses goals as suitable for the creation of suspense due to plan failure. Future work could consider the sequence of actions that lead up to the completion of the goal. In an interactive context, the players may feel less suspense when acting towards a goal if the actions taken in

service of the goal are not themselves high risk. The IRIS system could be extended to consider the constituent actions of suspenseful goals when generating stories for interactive contexts.

Another item concerns the survey presented to the participants. The survey questions were generated from an operationalization of IRIS's definition of suspense. However, it would be interesting to ask additional questions in another study such as this one. Questions that more explicitly measure emotion and questions that measure game and genre expectations could be added. This could determine if the application of the IRIS system in a game allows for the increase in player emotion that was theorized in Section 6. These questions could also help measure if IRIS is sufficient to overcome player game and genre expectations.

If the study was rerun, several more player metrics could be monitored. The success of the hints could be recorded to measure if they were indeed effective at getting the players back on track towards the game goals. Participants could play multiple versions of the game so we could have inter-version comparisons with the same participants. Finally, the game could be translated into a fully graphical version to see if suspense can be preserved in another interactive game environment.

Conclusion

The IRIS narrative generation system is a tool used to generate story outlines with built-in suspense that can be instantiated in different media. It was hypothesized that IRIS can be extended into the domain of a text-based adventure game when two transformations were applied. These transformations are 1) an in-game method of giving hints of how to achieve the game's goals and 2) the presence of the suspenseful action sequences from the original story outline and

the restriction of player agency during their execution. Participants played one of four versions of the game where one, both, or neither of the transformations were present. The results of this experiment show that for two of the three goals targeted to be suspenseful, the game version with both transformations was rated as significantly more suspenseful than the other versions. This shows that the IRIS system is effective at creating suspense in a text-based adventure game when both transformations are applied.

References

- Barber, H., and Kudenko, D. 2008. Generation of dilemma-based interactive narratives with a changeable story goal. In *Proceedings of the 2nd International Conference on Intelligent Technologies for Interactive Entertainment*, 6. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
- Cheong, Y., and Young, R. 2008. Narrative generation for suspense: Modeling and evaluation. *Interactive Storytelling* 144–155.
- Fendt, M. W., and Young, R. M. 2011. The Case for Intention Revision in Stories and its Incorporation into IRIS, a Story-Based Planning System. In *Intelligent Narrative Technologies*.
- Fendt, M. W.; Roberts, D.; and Young, R. M. 2013. Evaluating the use of action templates to create suspense in narrative planning. In *Intelligent Narrative Technologies*.
- Mateas, M., and Stern, A. 2003. Façade: An experiment in building a fully-realized interactive drama. In *Game Developers Conference, Game Design Track*, volume 2, 82.
- Meehan, J. 1977. TALE-SPIN, an interactive program that writes stories. In *In Proceedings of the Fifth International Joint Conference on Artificial Intelligence*.
- Turner, S. 1994. *The creative process: A computer model of storytelling and creativity*. Lawrence Erlbaum.
- Ware, S. G., and Young, R. M. 2011. CPOCL: A Narrative Planner Supporting Conflict. In *AIIDE*.