Intentionality and Conflict in *The Best Laid Plans* Interactive Narrative Virtual Environment

Stephen G. Ware, *Member, IEEE*, and R. Michael Young, *Senior Member, IEEE*

***Abstract*—In this paper, we present *The Best Laid Plans*, an interactive narrative adventure game, and the planning technologies used to generate and adapt its story in real time. The game leverages computational models of intentionality and conflict when controlling the non-player characters (NPCs) to ensure they act believably and introduce challenge into the automatically generated narratives. We evaluate the game's ability to generate NPC behaviors that human players recognize as intentional and as conflicting with their plans. We demonstrate that players recognize these phenomena significantly more than in a control with no NPC actions and not significantly different from a control in which NPC actions are defined by a human author.**

***Index Terms*—Computational models of narrative, conflict, intentionality, narrative, planning, *The Best Laid Plans*.**

I. Introduction

PLANNING technologies have proven a popular framework for developing interactive narrative experiences [1]–[3]. Plan data structures represent a sequence of causally-linked and goal-directed events, making them an excellent knowledge representation for computational models of narrative phenomena. Plan-based models are also attractive because they can be generated and adapted automatically by planning algorithms, allowing virtual reality systems to rewrite stories at run time to produce interactive narrative experiences. This article presents one such experience which uses plan-based models of intentionality and conflict to generate stories with believable character behavior and obstacles for the protagonist to overcome.

The Best Laid Plans is an interactive narrative point-and-click adventure game inspired in part by Fantasy Flight's storytelling card game *Aye! Dark Overlord*. The player takes on the role of a hapless goblin minion who has been sent on a menial quest by his master, an evil skeletal wizard. The game uses fast narrative planning techniques to tell, in real time, the story of how that quest goes horribly wrong. The player alternates between acting out a plan for the goblin to complete his mission and watching that plan get thwarted by non-player characters (NPCs) as they attempt to achieve their own goals.

The game's story is generated entirely from scratch by the Glaive narrative planner [4] based on the player's intended plan for the goblin. Riedl and Young's [5] model of agent intentionality ensures that all the actions in the story can be explained in terms of the motivations and goals of the agents who take them. Ware, Young, Harrison, and Roberts' [6] model of conflict allows Glaive to reason about failed and thwarted plans and how conflicts can be introduced to add obstacles for the protagonist to overcome.

We evaluate *The Best Laid Plans* in a controlled playtest to demonstrate that the stories it generates have recognizable intentionality and conflict. The system is tested relative to two controls, one in which NPCs do not act and one in which NPC behaviors are scripted by a human author. Section IV presents the results of this evaluation, which shows that players recognize intentionality and conflict more than in the control and not significantly differently from in the human-authored version. These results provide evidence that plan-based interactive narrative techniques can reason about narratological phenomena in real time to generate interactive stories in a virtual environment.

This article also serves as a description of the game, how it is designed, and how it can be used by other researchers for interactive narrative research. Section II describes the game world, how players interact with it, and the architecture of the system. Section III briefly describes those aspects of the narrative planner that are relevant to the evaluation.

II. Game Design

The Best Laid Plans was conceived as both a tool for evaluating plan-based computational models of narrative and a research prototype to demonstrate the viability of fast narrative planning for interactive virtual environments. The design of the system is partially in response to a demand in the intelligent narrative technologies community for more fully-realized systems which are modular and can be reused and modified by other researchers. Our choice of theme—a goblin minion on a doomed quest—is designed to enable stories fraught with conflict and follows Horswill's observation [7] that nascent interactive narrative techniques may be better suited to farce than high drama.

A. Gameplay

Players alternate between acting out a plan and watching it get thwarted by NPCs. The game begins in *Make Your Plan* mode, where the player is given control of the goblin and asked to make a plan for retrieving a bottle of hair tonic from the local village. The Dark Overlord, whom the goblin serves, watches telepathically from the top left of the screen and occasionally provides information or feedback. NPCs are minimally reactive in this mode, meaning they will not take actions of their own volition and will only participate in interactions initiated by the

Manuscript received December 11, 2014; revised May 19, 2015 and July 29, 2015; accepted September 17, 2015. Date of publication October 09, 2015; date of current version December 13, 2016.

S. G. Ware is with the Department of Computer Science, University of New Orleans, New Orleans, LA 70148 USA (e-mail: sgware@uno.edu).

R. M. Young is with the Department of Computer Science, North Carolina State University, Raleigh, NC 27695 USA (e-mail: young@csc.ncsu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCIAIG.2015.2489159

Fig. 1. When the player clicks on an item, a context-sensitive menu describes the available actions using icons and text.

goblin (e.g., accepting trades). As the player acts, the goblin's plan is described in text on the left side of the screen. This mode ends once the player has formed a complete plan that ends in returning to the Dark Overlord's tower with the hair tonic in hand. All actions taken during *Make Your Plan* mode are visualized with animations, sound, and possibly a textual message from the Dark Overlord.

After forming a plan, play changes to *Watch Your Story Unfold* mode in which the player watches the goblin execute the plan. Actions are visualized the same way.[1] In this mode, NPCs may act, and they will act to thwart the goblin's plan if possible. At any time, the player can choose to return to *Make Your Plan* mode and modify the goblin's plan starting from that moment in the story. If the goblin dies in this mode, the Dark Overlord rewinds time to the moment before the goblin's death and returns the game to *Make Your Plan* mode. Play continues in this fashion, with the player alternately acting out a plan and watching that plan get thwarted, until the goblin finally returns the tower with the hair tonic.

During *Make Your Plan* mode, the game has a simple point-and-click interface. The player can click on characters and objects in the world (including objects in the goblin's inventory, which are displayed at the bottom of the screen) to reveal a context-sensitive menu of available actions, as seen in Fig. 1. For example, when the player clicks on the hair tonic bottle, he or she is presented with three options: the eye icon (to examine the item), the hand icon (to pick up the item), and the handshake icon (to trade for the item, if it is for sale). A text description of the action is also shown to inform the player what the icons mean and to provide additional context. For example, if the hair tonic is for sale and its owner alive, the hand icon's description will read "Steal the hair tonic." A complete list of actions and other elements of the game's interface are shown in Fig. 2.

Two numbers are displayed on the screen: score and mana. Score begins at 0. Every time the game enters *Make Your Plan* mode, mana is reset to 25. When the player takes an action, mana is reduced by 1 (thus limiting the player to 25 actions per plan) and score is increased by 1. The player's final score is the total number of actions taken in *Make Your Plan* mode throughout the game. The total number of actions taken over the course of the game can exceed 25, but each individual time the game enters *Make Your Plan* mode, the player is limited to 25 actions. The goal of the game is to minimize the final score, which encourages players to choose shorter plans over longer ones. To increase replay value, various achievements can be earned that lower the player's final score. Achievements are based on particular play styles. For example, the *Triple Homicide* achievement is earned for killing three or more human NPCs and lowers the player's final score by 3 points, whereas the *Mr. Nice Guy* achievement is earned if the goblin does not kill any human NPCs and lowers the player's final score by 4 points.

An example game transcript is given in Fig. 3. The player starts by making a simple plan where the goblin walks to the town potion shop, buys the hair tonic, and walks home. This plan has 11 steps, so the player's score is 11 at the end of the first *Make Your Plan* mode. When this plan is visualized in *Watch Your Story Unfold* mode, the goblin gets waylaid and killed by a bandit on the way to the tower. Time rewinds to the moment before the bandit kills the goblin, and the game returns to *Make Your Plan* mode. Now the player must find some way around the bandit. The player decides that the goblin should return to town, steal a sword from the market, and use is to fight his way past the bandit. This new plan has 7 steps, so the player's score is now $11 + 7 = 18$. When this plan is visualized, the bandit follows the goblin into town, where the town guard kills first the bandit and (since he is now a criminal) the goblin. Time rewinds to the moment before the guard killed the goblin, and the player must now find a way around this new obstacle. Play continues like this until the goblin finally makes it to the tower with the hair tonic.

B. Game World

The game's 3-D environment provides a modest but nontrivial storytelling space. It contains 15 locations (see the map in Fig. 2), 17 items which can be picked up and used (including 4 kinds of spells to cast), 2 animal NPCs, and 7 human NPCs. The state of the game world—where each character is, what items each character is carrying, etc.—can be described by 745 Boolean predicate literals.

There are 10 kinds actions available to the player. Two actions, *Look At* and *Talk To*, provide information to the player but do not affect the state of the world. The remaining 8

[1]To help players distinguish between modes, the words *Make Your Plan* or *Watch Your Story Unfold* are written at the top of the screen. Additionally, a semitransparent halo effect (see Fig. 2) is placed around the corners of the screen in *Make Your Plan* mode.

Fig. 2. The game's interface (top), world map (bottom left), and the list of all 10 available actions (bottom right). This image shows the goblin in the town square along with the town guard and the weapon merchant. Clickable arrows on the ground indicate other locations the goblin can walk to: down to the junction, left to the tavern, up to the alley, and right to the potion shop. The current mode, *Make Your Plan*, is indicated at the top of the screen and by the frosted corners of the viewport. The Dark Overlord is watching telepathically from the top left and occasionally provides information. The player's plan is listed to the left. The goblin's inventory, mana, and score are at the bottom of the screen, along with the *Stop* button which returns the game to *Make Your Plan* mode (which is currently deactivated since the game is already in *Make Your Plan* mode).

actions can be taken by any human character (player or NPC). There are 848 possible parameterizations of these actions which might occur during the game. Note that this number does not include the thousands of possible parameterizations that can never occur.

One purpose of this game is to test if an audience can perceive intentionality in the actions of NPCs. NPC goals are conveyed to the player both through text and by relying on genre tropes. For example, merchant characters make sales pitches when they speak to the player, and the presence of a *trade* option when examining items for sale is meant to communicate that a character wants to sell certain items. The Dark Overlord will warn the goblin telepathically about the intentions of the bandit and the hungry wild animals. Section IV-B shows that players mostly agreed with the statement "The other characters were following their own goals," so we believe the game successfully communicates NPC goals to the player.

Player's First Plan
1. The goblin walks to the crossroads.
2. The goblin walks to the bridge.
3. The goblin walks to the junction.
4. The goblin walks to the town square.
5. The goblin walks to the potion shop.
6. The goblin trades gold coins for the hair tonic.
7. The goblin walks to the town square.
8. The goblin walks to the junction.
9. The goblin walks to the bridge.
10. The goblin walks to the crossroads.
11. The goblin walks to the tower.

First Story
9. The bandit walk to the junction.
10. The bandit kills the goblin with his sword.
11. The bandit picks up the hair tonic.

Player's Second Plan
10. The goblin walks to the town square.
11. The goblin steals a sword from the merchant's stall (and is now a criminal).
12. The goblin walks to the junction.
13. The goblin kills the bandit with his sword.
14. The goblin walks to the bridge.
15. The goblin walks to the crossroads.
16. The goblin walks to the tower.

Second Story
12. The bandit walks to the town square.
13. The town guard kills the bandit with his sword.
14. The town guard kills the goblin with his sword.

Player's Third Plan
14. The goblin kills the town guard with his sword.
15. The goblin walks to the potion shop.
16. The goblin steals the Teleport spell.
17. The goblin teleports to the bridge.
18. The goblin walks to the crossroads.
19. The goblin walks to the tower.

Third Story
15. The merchant picks up the bandit's sword.
16. The merchant kills the goblin with his sword.

Player's Fourth Plan
16. The goblin walks to the potion shop.
17. The goblin steals the Teleport spell.
18. The goblin teleports to the bridge.
19. The goblin walks to the crossroads.
20. The goblin walks to the tower.

Fourth Story
The player has won!
Final score: 29

Fig. 3. Example transcript of a complete game. Steps in gray never occur due to conflict. Steps in red are visualized, but then time is rewound to restore the goblin to life.

C. Architecture

The game has a simple client/server architecture in order to decouple the virtual world from the story planner. This allowed us to test various interactive narrative approaches during the evaluation and will enable future researchers to use *The Best Laid Plans* as a testbed for their own techniques.

The client comprises the 3-D world and game interface. It was created in the Unity game engine, version 4.2. The server was created in Java 1.7 and communicates with the client over a TCP socket. Each time the player finishes acting out a plan in *Make Your Plan* mode, the client sends the intended plan for the goblin to the server. The server responds with a plan which may also include actions by the NPCs and must either result in the player's plan executing successfully or the goblin's death. This plan is then visualized in *Watch Your Story Unfold* mode. The client sends a report each time an action is completed so that the server can update its model of the current world state.

The client and server executables, source code for the server, instruction manual, tutorial video, and a complete formal description of the story world [in the Planning Domain Definition Language (PDDL)] can be downloaded from: http://nil.cs.edu/projects/blp/.

We encourage other researchers to use this game as an environment for testing their own interactive narrative techniques.

III. Computational Models of Intentionality and Conflict

The Best Laid Plans uses a narrative planner called Glaive [4] to generate an interactive story in tandem with the player by controlling the game's NPCs. Glaive leverages two computational models of narrative to ensure that its stories meet the narrative expectations of the audience: Riedl and Young's [5] model of agent intentionality and Ware *et al.*'s [6] model of conflict. The purpose of this article is to demonstrate that human players recognize intentionality and conflict in the stories generated by *The Best Laid Plans*. The details of the Glaive algorithm are outside the scope of this article, but Glaive's definitions of intentionality and conflict are reproduced here.

A. Plan-Based Causal Models

A planner attempts to solve the following problem: given a world in some initial state, a goal, and a set of possible actions, find a sequence of those actions (called steps) which achieves the goal [8]. Each step has preconditions which must be true immediately before it is executed and effects which modify the world state.

The partial order causal link (or POCL) plan model contains a first class representation of causality called causal links. Riedl and Young [5] extended this model to define a kind of intentional planning which ensures that characters appear to act believably while still achieving the author's goal. Each action in an intentional planning problem is annotated with a list of characters who must consent to take that action. We call these *consenting characters*. In a valid plan, each action must be explained in terms of the motivations and goals of its consenting characters. Actions which do not require the consent of any characters represent accidents or forces of nature. Actions which require the consent of multiple characters require that each character have its own reason to take the action and thus represent cooperation.

Ware and Young [9] further extended this model to represent failed plans and conflict. By marking certain actions as *intended but not executed*, a plan can represent actions that characters wanted to take but could not due to causal conflicts with other agents or the environment. For each conflict, a plan represents multiple possible worlds, one in which the agent's plan succeeds and one in which the agent's plan fails.

In short, a narrative planner using both of these models is solving a multi-agent coordination problem. The planner must find a series of steps which achieves the author's goal (in this case, the author's goal is that the goblin die), but each step must be explained in terms of the individual goals of the agents who take them. The planner acts as an invisible puppet master, achieving the author's goal by only taking steps which are clearly motivated and goal-oriented for the characters who take them.

B. Intentionality

The rich knowledge representation of POCL plans makes them appealing for computational models of narrative, but little research has been done on fast POCL planning in recent years, so the current state of the art in that area is often considered too slow for use in real time systems such as *The Best Laid Plans*.[2] Alternatively, forward-chaining state-space heuristic search planners have proven extremely fast [10] but use a very simple knowledge representation. Glaive attempts to combine the strengths of both approaches. Glaive is a forward-chaining state-space heuristic search planner based on Hoffmann and Nebel's [11] fast-forward planner that tracks causal links in addition to the current state of the world. In other words, a Glaive plan tracks not only which propositions are true and false but also which previous steps made them true or false. This allows Glaive to benefit from the speed of fast heuristic planning while still performing the causal reasoning on which its models of intentionality and conflict are based.

Definition 1: A *causal link* $s \xrightarrow{p} t$ exists from step s to step t for proposition p if and only if s occurs before t, s has the effect p, t has a precondition p, and no step occurs between s and t which has the effect $\neg p$. We say s is the *causal parent* of t, and that a step's *causal ancestors* are those steps in the transitive closure of this relationship.

A causal link explains how a fact which is needed by a later step was established by an earlier step. Cognitive scientists [12], [13] have observed that humans reason about causally linked chains of events when experiencing a narrative. These chains can be used to characterize intentionality [14], [2] in terms of goal-directed actions.

A narrative planner must reason about two kinds of goals: the author's goal and the goals of individual characters. Character goals are expressed as modal *intends* propositions in the current state (e.g., the goblin intends to be in possession of the hair tonic). These modal propositions are part of the planning domain which can appear anywhere a proposition might, such as in the initial state of the story world or the effects of a step. For example, the *take* action specifies that if the item being picked up is for sale and the owner is alive (i.e., the item is being stolen) then the town guard intends that the thief be dead. Thus characters can adopt and abandon goals over the course of plan as a result of the actions they or other characters take. Character goals are specified in the planning domain, but the planner decides how the characters will act to achieve them. We call a causally-linked, goal-directed chain of steps taken by a single character an *intentional path* in a plan:

Definition 2: An *intentional path* $\langle s_1, p_1, s_2, p_2, \ldots, s_n, g \rangle$ is an alternating sequence of n steps and n propositions. All the steps in the path must have a common consenting character c who intends some goal proposition g. Beginning in the state immediately before step s_1 and until the state immediately before step s_n, it must be true that c intends g. The final proposition in the path must be g, and the final step s_n must be a step with effect g. There must also exist a causal link $s_i \xrightarrow{p_i} s_{i+1}$ for all $0 < i < n$.

When a step appears on an intentional path for some character and some goal, we say it is intended by that character as a means of reaching that goal. Consider an example from the game world. Suppose the goblin has entered the town potion shop and stolen the hair tonic. This causes the guard to adopt the goal that the goblin be dead. The plan which is visualized during *Watch Your Story Unfold* mode could contain the following intentional path for the guard and his goal that the goblin be dead.

1) s_1: The guard enters the potion shop.
2) p_1: The guard is in the potion shop.
3) s_2: The guard kills the goblin in the potion shop.
4) g: The goblin is dead.

The last proposition in the path is the character goal. The last step is an action taken by the guard which achieves the goal. There also exists a causal link $s_1 \xrightarrow{p_1} s_2$ which explains why the guard entered the potion shop—he did so in order to be in the same location as the goblin so that he could kill the goblin. By ensuring that all steps appear on causal chains, Glaive ensures that every action can be explained in terms of the individual goals of the characters who take them.

C. Conflict

Intentionality and conflict are closely related; narratologists have described conflict in terms of thwarted intentional actions [15]. Ware *et al.* [6] demonstrated that when reading plan-based narratives an audiences can recognize when one agent's action undoes the preconditions needed for steps in another agent's plan. Glaive's definition of conflict centers on how one intentional path can thwart another.

A conflict in *The Best Laid Plans* involves two agents. For each agent, Glaive has discovered an intentional path that explains how the agent will achieve one of its goals. We say these intentional paths conflict when one would prevent the other from succeeding. This means that the actual story Glaive generates cannot contain all steps from both intentional paths; once one character's plan has been thwarted, the rest of his plan cannot be included (because it is now impossible to execute). The steps that get left out are referred to as *intended but not executed* steps [6].

Definition 3: A *conflict* exists between agent a, for which there exists an intentional path π_a, and agent b, for which there

[2]Forward-chaining state-space heuristic search planners like Glaive have dominated the biannual International Planning Competition since its inception in 1998 [10]. This does not imply that POCL planners are inherently slower, only that no current POCL algorithms are known which can perform comparably on benchmark problems. As a point of comparison, the original narrative planner, IPOCL [5], was tested on a single benchmark problem. It took over 12 hours to solve it, visited 673,079 nodes and expanded 1,857,373 while using a domain-specific heuristic. By contrast, Glaive takes only 64 milliseconds, visits 12 nodes and expands 189 while using a domain independent heuristic.

exists an intentional path π_b, if there exists some step $s_a \in \pi_a$ with precondition p and some step $s_b \in \pi_b$ with effect $\neg p$ such that, in the story generated by the planner, step s_b occurs, step s_a does not occur, and none of the other steps after s_a from π_a occur.

This formal definition will be more clear when we consider an example from the game world. Suppose the goblin has the hair tonic and is currently in the town. Recall that his goal is to reach the tower with the hair tonic. The steps in his intentional path are as follows.

1) The goblin walks to the junction.
2) s_a: The goblin walks to the bridge.
3) The goblin walks to the crossroads.
4) The goblin walks to the tower.

The bandit wants to steal items from travelers. In this particular situation, the bandit wants to be in possession of the hair tonic. The steps in his intentional path are as follows.

1) The bandit walks to the junction.
2) s_b: The bandit kills the goblin.
3) The bandit takes the hair tonic from the goblin's body.

One precondition of the walk action is that the traveling character be alive. Action s_b (the bandit kills the goblin) makes it impossible for step s_a (the goblin walks to the bridge) to happen. We can say that step s_b thwarts the goblin's intentional path π_a. These two intentional paths can be combined into a story which satisfies the author's goal that the goblin be dead.

1) The goblin walks to the junction.
2) The bandit walks to the junction.
3) The bandit kills the goblin.

All the steps in this plan can be explained in terms of character goals because they lie on intentional paths for those characters. This is true even when some steps from those intentional paths are not included in the actual plan that gets visualized. We know why the goblin walked to the junction (so that he could continue on to the tower) even though we do not actually see it happen.

D. Story Selection

Having stated the general definitions used intentionality and conflict, we now turn to some of the specific constraints placed on the planner specifically for *The Best Laid Plans*.

Every time the player finishes *Make Your Plan* mode, the game client sends the intended plan for the goblin to the server. Glaive then visits the first 5000 nodes in its search space, using the current world state as the root. Due to the limit imposed by the player's mana score, an individual plan for the goblin cannot be more than 25 actions. These limits of 25 actions and 5000 visited nodes were chosen based on performance testing to ensure that the planning phase never took longer than 1 second on the computers used in the evaluation of the game. We consider 1 second or less to be sufficiently fast performance for online interactive narrative generation.

If no plan is found which results in the goblin's death, the player's plan is returned as-is to the client, otherwise the best solution is returned. When multiple plans that results in the goblin's death are found, we rank them according to the following domain-independent metrics.

1) Prefer the story in which the player witnesses the highest number of goals achieved by other characters.
2) In the event of a tie, prefer the story in which 75% of the player's original plan is executed before the goblin dies (or as close to 75% as possible). This threshold of 75% was chosen based on feedback from players of an early version of the game, but we make no claim that this threshold is necessarily optimal.
3) In the event of a further tie, prefer the shorter story.
4) In the event of a further tie, prefer the story in which the player witnesses the highest number of goals acted on but not achieved by other characters.

These metrics are motivated by two design principles: The player should see as much of their own plan as possible, and the player should see as many NPC actions as possible.

IV. Evaluation

We claim that the interactive stories generated by *The Best Laid Plans* have recognizable intentionality and conflict. We expect an audience to recognize that these phenomena are lacking when NPCs do not act and present when NPCs act according to their expectations. The ability to generate interactive stories with these phenomena was evaluated in an empirical human trial of *The Best Laid Plans*. Three different versions of the server side of the game were created for this trial.

- Control: NPCs do not act, except when cooperating with the goblin (e.g., trading). In this version, the player's plan is simply returned by the server with no modifications and thus will always succeed.[3]
- Glaive: NPCs are controlled by the Glaive planner as described in the previous section.
- Scripted: NPCs are controlled by triggers written by a human author using a declarative scripting language. In an attempt to remain objective, we asked another scientist familiar with planning technologies, but with no direct experience in the development of this game, to write these triggers. The complete list of triggers is given in Fig. 4. When fully grounded, there are a total of 44,733 possible instances.

This *Scripted* version of the game is meant to approximate the video game industry's current approach to interactive narrative: hand-authored scripts that must anticipate every important narrative situation at design time. Levine [16] has identified the need for procedural narrative generation in games at run time. Such narratives have the potential to increase replay value and to adapt to the needs and desires of individual players. To realize these long-term goals, we must first ensure that we can reliably generate stories with essential narrative properties like intentionality and conflict. Human-authored scripts must be authored individually for each new virtual environment and may vary in quality based on the skill of the author and the author's

[3]The other control we considered is one in which the NPCs are coordinated by a classical planner (i.e., one with no particular model of narrative) that attempts to prevent the goblin from completing his mission. In this version, NPCs would thwart the goblin with no consideration of their individual goals. For example, every NPC capable of attacking the goblin would converge on him and attack immediately with no motivation or provocation. We decided not to use this control because it would be extremely tedious—probably requiring dozens of plans per game—and we feared that too few players would finish the game to provide a sufficient control group.

Human-Authored Triggers

1. The wolf, crocodile, and troll are beasts.
2. If the goblin is at the same location as a beast, the beast attacks the goblin.
3. The bandit is a criminal.
4. If the goblin steals an item, he is a criminal.
5. If the goblin attacks a non-criminal, he is a criminal.
6. If there is a criminal in a location adjacent to the guard's location, the guard will walk to that location.
7. If the guard is at the same location as a criminal, he will attack the criminal.
8. If the bandit and any other character are both at the camp, the bandit will attack that character.
9. If a character is carrying an item, is at a location adjacent to the bandit's location, and the guard is not also at that location, the bandit will walk to that location.
10. If the bandit is at the same location as the guard, the bandit will walk to a different location.
11. If a character is carrying an item and is at the same location as the bandit, the bandit will attack that character.
12. If the merchant is at the same location as a criminal and there is a weapon nearby, the merchant will pick up the weapon.

Fig. 4. Triggers uses in the *Scripted* version of the game, translated from the declarative scripting language into natural language.

ability to foresee every possible situation. The advantage of algorithms like Glaive is their ability to reliably generate stories with these properties in real time using formal computational models.

We make no claim that the particular set of triggers used in the *Scripted* version represents the way that humans in general would script such behaviors, only that this set of scripts is one possible set which succeeds in generating stories with intentionality and conflict. This claim is supported by the results in Section IV-B, which show that people who played this version of the game mostly agreed with the statements, "The other characters had good reasons for their actions," "The other characters were following their own goals," "The other characters were reacting to the things I did," and, "Some characters were trying to prevent me from accomplishing my goals."

A. Experimental Design

Subjects were first shown a tutorial video that explained how to play the game. They were also given a printed instruction manual to consult at any time during play, which included the map of the world and list of available actions shown in Fig. 2. Subjects were asked to play the game only once and then to fill out a survey about their experience. Subjects were alone in the room during play and while filling out the survey. Each subject played only one of the three versions of the game.

We recruited participants from among graduate and undergraduate Computer Science students at North Carolina State University via e-mail, social networking, and posted fliers. Many students are offered extra credit in their classes for participating in research studies.

A total of 75 subjects played the game. From those, 4 were removed from consideration because the game crashed. Five were removed because they played the game multiple times. 1 was removed for not watching the tutorial video and expressing significant confusion about the interface and goal of the game. 1 was removed for having significant knowledge of the Glaive system before participating. The remaining 64 subjects were split evenly between the three treatments.

- Control: 21 subjects, 16 M/4 F/1 T, mean age 22.
- Glaive: 21 subjects, 17 M/4 F, mean age 21.
- Scripted: 22 participants, 16 M/6 F, mean age 22.

Most subjects had prior experience playing video games. 55 of 64 had played at least one role playing game in the last five years; 31 had played 5 or more; 19 had played 10 or more. 44 of 64 had played at least one point-and-click adventure game in the last fives years; 14 had played 5 or more; 6 had played 10 or more.

The post survey was composed of statements about intentionality, conflict, and other factors. Subjects were asked to report to what extent they agreed with those statements on a 7 point Likert scale ranging from 1, "Strongly Disagree" to 7, "Strongly Agree." The statements from the post survey relevant to this evaluation are given in Table I.

B. Results

The purpose of this trial was to test whether or not subjects recognized intentionality and conflict in interactive stories generated by *The Best Laid Plans*. In general, we hypothesized that the Glaive and Scripted versions would outperform the Control, and that there would be no significant differences between the Glaive and Scripted versions. Specifically, we expected subjects who played the Glaive and Scripted versions to agree significantly more with all statements except Statement 4 ("Some characters were trying to help me accomplish my goals."). We expected no significant difference in agreement across treatments for Statement 4 because the game contains only neutral and hostile NPCs and no friendly NPCs.

The Wilcoxon sum-rank test [17] was used to compare treatments. This is a non-parametric test for comparing independent samples of ordinal data. It not only allowed us to reject the null hypothesis—that agreement was about the same across treatments—but also to confirm an alternative hypothesis—that one treatment showed higher agreement than another. We compared each of the three treatments pairwise to one another. The results of those comparisons are given in Table I. A p value is given for each hypothesis, which indicates the probability that the observed differences are due to chance. These values were adjusted using Benjamini and Hochberg [18] correction for multiple comparisons. Common language effect size is also given, which indicates the proportion of samples which supported the hypothesis. A green check mark indicates support for our hypotheses, whereas a red X indicates no support. Three statements on the post survey showed a significant difference ($p < 0.05$) in agreement across treatments, and the alternative hypotheses for these are highlighted in bold.

1) Intentionality: Fig. 5 demonstrates that subjects who played the Scripted version mostly agreed with question 1, 2, and 3, which were designed to measure the perception of

TABLE I
PAIRWISE COMPARISONS OF AGREEMENT WITH STATEMENTS ON THE POST SURVEY

Statement	Hypothesis	p Value	Effect Size	Support?
1. The other characters had good reasons for their actions.	Control>Glaive	>0.999	30%	✓
	Glaive>Control	0.490	52%	✗
	Control>Scripted	>0.999	25%	✓
	Scripted>Control	0.281	59%	✗
	Glaive>Scripted	>0.999	34%	✓
	Scripted>Glaive	0.686	46%	✓
2. The other characters were following their own goals.	Control>Glaive	>0.999	16%	✓
	Glaive>Control	**0.042**	65%	✓
	Control>Scripted	>0.999	8%	✓
	Scripted>Control	**0.002**	75%	✓
	Glaive>Scripted	>0.999	32%	✓
	Scripted>Glaive	0.681	44%	✓
3. The other characters were reacting to the things I did.	Control>Glaive	>0.999	17%	✓
	Glaive>Control	**0.042**	66%	✓
	Control>Scripted	>0.999	18%	✓
	Scripted>Control	**0.042**	66%	✓
	Glaive>Scripted	0.990	34%	✓
	Scripted>Glaive	>0.999	33%	✓
4. Some characters were trying to help me accomplish my goals.	Control>Glaive	0.677	50%	✓
	Glaive>Control	>0.999	37%	✓
	Control>Scripted	0.947	45%	✓
	Scripted>Control	>0.999	42%	✓
	Glaive>Scripted	>0.999	35%	✓
	Scripted>Glaive	0.632	52%	✓
5. Some characters were trying to prevent me from accomplishing my goals.	Control>Glaive	>0.999	6%	✓
	Glaive>Control	**<0.001**	86%	✓
	Control>Scripted	>0.999	2%	✓
	Scripted>Control	**<0.001**	92%	✓
	Glaive>Scripted	>0.999	20%	✓
	Scripted>Glaive	0.640	34%	✓

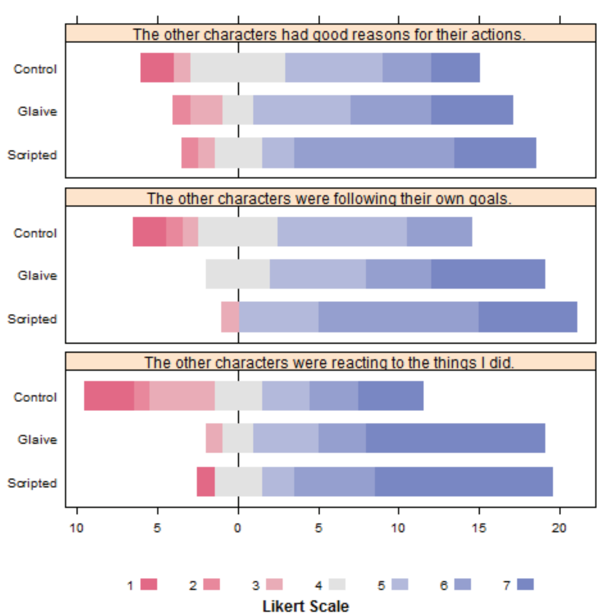

Fig. 5. Diverging stacked bar charts showing agreement with three statements on the post survey measuring the perception of intentionality in NPCs.

Fig. 6. Diverging stacked bar charts showing agreement with two statements on the post survey measuring the perception of conflict between the player and NPCs.

intentionality. This implies that the human author's scripts successfully conveyed intentionality through the actions of the NPCs.

Subjects who played the Glaive and Scripted versions agreed more with the statements "The other characters were following their own goals," and "The other characters were reacting to the things I did," than subjects who played the Control. There was no significant difference in agreement between the Glaive and Scripted versions for these statements. This supports our hypotheses for those two statements and demonstrates that subjects recognized intentionality in the NPCs for the Glaive and Scripted versions of the game.

Most subjects agreed with the statement "The other characters had good reasons for their actions," across all treatments. This is encouraging for the Glaive and Scripted versions of the game, but strange for the Control since the NPCs did not act other than to react to the goblin during actions like trades. We speculate that subjects who played the Control and observed no actions by NPCs were not sure how to answer this question.

2) Conflict: There were no significant differences in agreement across treatments for the statement "Some characters were trying to help me accomplish my goals." This supports our hypothesis and was expected because the game contains no friendly NPCs.

Subjects who played the Glaive and Scripted versions agreed more with the statement "Some characters were trying to prevent me from accomplishing my goals," than subjects who played the Control. This supports our hypothesis and demonstrates that subjects did recognize conflict with NPCs in the Glaive and Scripted versions. This is also consistent with previous experiments by Ware *et al.* [6] which demonstrated that thwarted plans are recognizable to audiences reading short text stories.

V. Conclusion

The results collected from these 64 playtesters support our hypotheses (to the extent discussed above). We have demonstrated two critical results. The first is that players can recognize when intentionality and conflict are present to different extents. They recognize these phenomena more clearly in stories generated by Glaive and Scripted than in stories with no NPC actions. The second result is that they agree these phenomena are present in stories generated by Glaive and Scripted, and they do not recognize significant differences in these phenomena between the two versions. We make no claim that these models of intentionality and conflict are the only way to express these qualities, nor that they are the simplest possible models. We only claim that Glaive succeeds in generating stories which an audience does not recognize as significantly different from those scripted by a human author with regards to intentionality and conflict in this virtual environment.

Previous studies by Riedl and Young [5] and by Ware *et al.* [6] evaluated these same computational models of intentionality and conflict in static text stories. This evaluation provides further support for the validity of these models by testing them in a new context, a dynamic 3-D virtual environment whose narrative was potentially different for each player. We consider both kinds of studies valuable and complimentary for the evaluation of these models because each controls for different aspects of the experience.

VI. Mis-Spun Tales and Future Work

Two notable emergent properties arose during the design and testing of *The Best Laid Plans* which suggest interesting directions for future work. Firstly, it was our intention to design a virtual world in which it was always possible to thwart the player's first plan. However, while the game was being exhibited at a conference, one dedicated player discovered a plan which could not be thwarted. This demonstrates that, even in relatively small virtual environments such as this one, it is difficult to anticipate the entire space of play at design time. As Pizzi *et al.* [19] have demonstrated, the same plan-based tools which control the NPCs can be used to simulate players during the design phase to generate a representative sample of the play space and to detect situations such as this one.

Another interesting behavior arose when the town guard had to decide between justice and hunger. In the Glaive version, the guard has two motivations: to kill criminals and to find something to eat. Food is available in the tavern, but the guard has no money. We anticipated numerous interactions based on these motivations: the guard can thwart the player if the goblin commits a crime, the guard will go to the tavern to eat if given money, and the guard can be circumvented by giving him poisoned food. Instead, at the start of every story, the guard would walk to the tavern, steal food from the innkeeper, eat it, and then (because he was now a criminal) kill himself. This behavior is a valid intentional plan, and the model correctly represents it as an internal conflict between the guard and himself. However, because this removed the guard from the game before the player had a chance to interact with him, we imposed the constraint that the guard cannot enter the tavern. This fixed the problem, but a more robust solution is needed. Perhaps the model of conflict needs to be extended with some mechanism for resolving internal conflicts based on a character's priorities. For example, we could specify that the guard's desire for justice is greater than his desire for food, so when presented with a conflict between these two desires he would choose not to act on the plan to steal the food.

The Best Laid Plans attempts to validate the plan-based computational model of conflict proposed by Ware *et al.* in an interactive setting. The player needs to experience conflict for the evaluation to be effective, so it is essential that the game be able to predict the player's plan accurately, but this task is exceedingly difficult [20]. The distinction between *Make Your Plan* mode and *Watch Your Story Unfold* mode was implemented to avoid the need for plan recognition; by asking the player to report a plan the game can be sure to thwart it if possible. However, this distinction breaks the immersion of the game and was reported to be tedious. Having demonstrated that players recognize conflict when the system has an accurate means of plan recognition, we intend to relax this constraint in future versions of the game by eliminating the distinction between the two modes and allowing NPCs to act at the same time the player is acting.

Controlling NPCs in real time will present new interactive narrative challenges to ensure that the player experiences conflict, challenge, and suspense. Accurate plan recognition techniques [21] can be leveraged to achieve similar results without first asking the player to report a complete plan. We also see great potential for extending the possible worlds reasoning used in Glaive's model of conflict to cover a wider variety of phenomena because numerous narrative theorists have developed analytical systems based on possible worlds [22], [23]. Building on work by Gerrig and Bernardo [24], Cheong and Young [25] demonstrated that as readers perceived fewer plans for the protagonist to escape peril their experience of suspense increased. By generating and updating a network of possible worlds, fast narrative planning algorithms should be able to reason about how many possible solutions exist for the player in *The Best Laid Plans*. This system would coordinate NPCs to minimize the number of available solutions while still ensuring that at least one exists, thus generating conflict and increasing the user's experience of suspense.

Finally, one important direction for future work will be to develop larger, more complex interactive narrative virtual environments. While we were encouraged to see no significant differences between the Glaive and Scripted versions, the true value of fast planning-based interactive narrative techniques lies in controlling narratives whose space is too large to be anticipated by human authors. We believe that as narrative planning techniques scale up to handle larger environments, they will eventually be able to outperform some human-scripted interactive stories.

ACKNOWLEDGMENT

The authors would like to acknowledge C. Stith and P. Wright for their exceptional work in programming the Unity client for *The Best Laid Plans*, and the NCSU Department of Computer Science Research Experiences for Undergraduates program for supporting their effort. They would also like to thank R. E. Cardona Rivera for writing the triggers used in the Scripted version of the game. Finally, they wish to thank B. Harrison and all the members of the Liquid Narrative Group at North Carolina State for their help in testing and designing the game.

REFERENCES

[1] R. M. Young, "Notes on the use of plan structures in the creation of interactive plot," in *Proc. AAAI Fall Symp. Narrative Intell.*, 1999, pp. 164–167.

[2] F. Charles, M. Lozano, S. J. Mead, A. F. Bisquerra, and M. Cavazza, "Planning formalisms and authoring in interactive storytelling," in *Proc. Technol. Interactive Digital Storytelling Entertain.*, 2003.

[3] R. M. Young, S. G. Ware, B. A. Cassell, and J. Robertson, "Plans and planning in narrative generation: A review of plan-based approaches to the generation of story, discourse and interactivity in narratives," *SDV. Sprache und Datenverarbeitung*, vol. 37, no. 1–2, pp. 41–64, 2013.

[4] S. G. Ware and R. M. Young, "Glaive: A state-space narrative planner supporting intentionality and conflict," in *Proc. 10th AAAI Int. Conf. Artif. Intell. Interactive Digital Entertain.*, 2014, pp. 80–86.

[5] M. O. Riedl and R. M. Young, "Narrative planning: Balancing plot and character," *J. Artif. Intell. Res.*, vol. 39, no. 1, pp. 217–268, 2010.

[6] S. G. Ware, R. M. Young, B. Harrison, and D. L. Roberts, "A computational model of narrative conflict at the fabula level," *IEEE Trans. IEEE Trans. Comput. Intell. AI Games*, vol. 6, no. 3, pp. 271–288, 2014.

[7] I. Horswill, "Punch and judy ai playset: A generative farce manifesto, or, the tragical comedy or comical tragedy of predicate calculus," in *Proc. 5th Workshop Interactive Narrative Technol./8th AAAI Conf. Artif. Intell. Interactive Digital Entertain.*, 2012, pp. 14–19.

[8] A. Newell, J. C. Shaw, and H. A. Simon, "Report on a general problem-solving program," in *Proc. Int. Conf. Inf. Process.*, 1959, pp. 256–264.

[9] S. G. Ware and R. M. Young, "CPOCL: A narrative planner supporting conflict," in *Proc. 7th AAAI Int. Conf. Artif. Intell. Interactive Digital Entertain.*, 2011, pp. 97–102.

[10] A. J. Coles *et al.*, "A survey of the seventh international planning competition," *AI Mag.*, 2011.

[11] J. Hoffmann and B. Nebel, "The FF planning system: Fast plan generation through heuristic search," *J. Artif. Intell. Res.*, vol. 14, pp. 253–302, 2001.

[12] T. Trabasso and L. L. Sperry, "Causal relatedness and importance of story events," *J. Memory Lang.*, vol. 24, no. 5, pp. 595–611, 1985.

[13] T. Trabasso and P. Van Den Broek, "Causal thinking and the representation of narrative events," *J. Memory Lang.*, vol. 24, no. 5, pp. 612–630, 1985.

[14] A. B. Loyall, "Believable agents: Building interactive personalities," Ph.D. dissertation, Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, 1997.

[15] D. Herman, *Story Logic: Problems and Possibilities of Narrative*. Lincoln, NB, USA: Univ. Nebraska Press, 2004.

[16] K. Levine, "Narrative legos," in *Game Developers Conf.*, 2014 [Online]. Available: http://www.gamespot.com/videos/narrative-legos-with-ken-levine-gdc-2014/2300-6417876/

[17] E. L. Lehmann and H. J. M. D'Abrera, *Nonparametrics: Statistical Methods Based on Ranks*. New York, NY, USA: Springer-Verlag, 2006.

[18] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Stat. Soc.*, ser. B (Methodological), pp. 289–300, 1995.

[19] D. Pizzi, J.-L. Lugrin, A. Whittaker, and M. Cavazza, "Automatic generation of game level solutions as storyboards," *IEEE Trans. Comput. Intell. AI Games*, vol. 2, no. 3, pp. 149–161, 2010.

[20] B. Harrison, S. G. Ware, M. W. Fendt, and D. L. Roberts, "A survey and analysis of techniques for player behavior prediction in massively multiplayer online games," *IEEE Trans. Emerg. Top Comput.*, vol. 3, no. 2, pp. 260–274, Jun. 2015.

[21] E. Y. Ha, J. Rowe, B. Mott, and J. Lester, "Recognizing player goals in open-ended digital games with markov logic networks," in *Plan, Activity and Intent Recognition: Theory and Practice*, G. Sukthankar, C. Geib, H. H. Bui, D. Pynadath, and R. P. Goldman, Eds. New York, NY, USA: Morgan Kauffman, 2014, pp. 289–311.

[22] M.-L. Ryan, *Possible Worlds, Artificial Intelligence, and Narrative Theory*. Bloomington, IN, USA: Indiana Univ. Press, 1991.

[23] J. S. Bruner, *Actual Minds, Possible Worlds*. Cambridge, MA, USA: Harvard Univ. Press, 1986 [Online]. Available: http://books.google.com/books?isbn=0674003667

[24] R. J. Gerrig and A. B. I. Bernardo, "Readers as problem-solvers in the experience of suspense," *Poetics*, vol. 22, no. 6, pp. 459–472, 1994.

[25] Y.-G. Cheong and R. M. Young, "Narrative generation for suspense: Modeling and evaluation," in *Proc. 1st Int. Conf. Interactive Digital Storytelling*, 2008, pp. 144–155.

Stephen G. Ware (M'12) received the B.S. degree in computer science and philosophy from Loyola University, New Orleans, LA, in 2008, the M.S. degree in computer science from North Carolina State University, Raleigh, NC, in 2011, and the Ph.D. degree in computer science from North Carolina State University in 2014.

Currently, he is an Assistant Professor of Computer Science and Director of the Narrative Intelligence Lab at the University of New Orleans. His research interests are in computer narrative, especially automated planning, knowledge representation, narratology, cognitive science, and artificial intelligence in games. He is a member of the Association for the Advancement of Artificial Intelligence, the International Game Developers Association, and the Association for Computing Machinery.

R. Michael Young (M'98–SM'07) received the B.S. degree in computer science from the California State University at Sacramento, Sacramento, in 1984, the M.S. degree in computer science with a concentration in symbolic systems from Stanford University, Stanford, CA, in 1988, and the Ph.D. degree in intelligent systems from the University of Pittsburgh, Pittsburgh, PA, in 1998.

He was a Postdoctoral Fellow at the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA. He is currently a Professor of Computer Science, University Faculty Scholar and Executive Director of the Digital Games Research Center, North Carolina State University, Raleigh. His research interests span the computational modeling of narrative, the use of artificial intelligence techniques within computer games, and virtual worlds.

Dr. Young is an Association for Computing Machinery Distinguished Scientist, as well as a Senior Member of the Association for the Advancement of Artificial Intelligence.