Proceedings of the Fourteenth Artificial Intelligence and
Interactive Digital Entertainment Conference (AIIDE 2018)

Combining Intentionality and Belief: Revisiting Believable Character Plans

Alireza Shirvani, Rachelyn Farrell, Stephen G. Ware
Narrative Intelligence Lab, Department of Computer Science
University of New Orleans, New Orleans, LA 70148, USA
{ashirvan, rfarrell, sgware}@uno.edu

Abstract

In this paper we present two studies supporting a plan-based model of narrative generation that reasons about both intentionality and belief. First we compare the believability of agent plans taken from the spaces of valid classical plans, intentional plans, and belief plans. We show that the plans that make the most sense to humans are those in the overlapping regions of the intentionality and belief spaces. Second, we validate the model's approach to representing anticipation, where characters form plans that involve actions they expect other characters to take. Using a short interactive scenario we demonstrate that players not only find it believable when NPCs anticipate their actions, but sometimes actively anticipate the actions of NPCs in a way that is consistent with the model.

Introduction

Intelligent interactive narratives have numerous applications in entertainment, education, training, and therapy. One of the distinctive challenges of generating stories in virtual environments is that characters must appear realistic, even when they are being controlled by a centralized experience manager. The story might be generated by a single agent in a fully observable environment, but each character has to seem like an individual agent with its own goals and its own limited, possibly inaccurate beliefs about the world.

Narrative planning has emerged as a common story generation technique well-suited to coordinating a story toward the system designer's goal while providing a formal, generative framework for reasoning about narrative phenomena (Young et al. 2013). This paper focuses on two features central to producing realistic virtual agent behavior: intentionality and belief. Each is individually important, and reasoning about both together creates more believable stories.

We previously (2017) presented a plan-based model incorporating intentionality and belief which modified previous definitions of realistic character behavior. First, when a virtual agent forms a plan and acts on it, an audience will find the agent's behavior realistic as long as the agent *believes* that plan is possible (even if it isn't actually possible). Second, an agent's plan may include actions to be taken by other agents, provided that the first agent believes the second agent could and would take those actions. These definitions permit a wider variety of stories, including ones where characters *anticipate*, meaning one character incorporates the actions of other characters into their plan. We present two studies in support of this model. The first asks people to compare the believability of character plans demonstrating intentionality, belief, both, and neither. The second demonstrates that the anticipation enabled by the model appears believable to a human audience, both because they say it does after reading and because they behave similarly when acting as a character in the narrative.

Related Work

Ample prior work has been done on belief and intentionality, especially for agent modeling, where the Belief Desire Intention framework (Bratman 1987) has been influential. Much of this research is for what Riedl and Bulitko (2013) dub *strong autonomy* systems.

Strong autonomy systems focus on creating realistic agents, typically because the agent must operate in the real world or because designers want realistic agents in a virtual world. Their challenge is ensuring the narrative which emerges from those agents meets the author's constraints.

Strong story systems, like narrative planners, centralize reasoning in a single agent to ensure the author's constraints are met; their challenge is ensuring that agents appear believable in the process. Strong story systems are free to exploit the planner's omniscience and omnipotence in the virtual world so long as agents *appear* realistic. We focus this survey on strong story systems.

Intentionality

Planning is a branch of AI devoted to finding a sequence of actions to achieve a goal. Young (1999) proposed plans as a formal, generative model for stories that can explicitly represent the temporal and causal relationships between events.

Riedl and Young (2010) offered a plan-based model of *intentionality*, the tendency of agents to work toward their goals. Formal definitions are in the next section, but essentially an action appears intentional when it achieves part of an agent's goal or enables an action which achieves an agent's goal. Riedl and Young also introduced a limited notion of anticipation by allowing one agent to delegate its intention to another.

A significant limitation of this model is that every agent's plan must succeed to have the desired causal structure. Ware et al. (2014) extended the model to represent failed plans and conflict. An action appears intentional if *there exists* a plan with the aforementioned structure, even if that plan does not actually get carried out, perhaps because it was thwarted.

Teutenberg and Porteous's IMPRACTical (2013) is a hybrid intentional planner. Like a strong autonomy system, each agent identifies relevant actions to achieve its goals, but search is done by a centralized strong story planner. It also allows failed plans, and it models anticipation, where one agent expects the actions of others. It decides which actions are relevant using a simplified version of the problem it solves, so it is not guaranteed to produce plans with Riedl and Young's causal structure.

Ware and Young's Glaive planner (2014) also considers relevant actions to improve its speed, but (at an additional cost) guarantees Riedl and Young's causal structure.

Both IMPRACTical and Glaive offer considerable speedups over earlier intentional planners. All the models above assume full observability—that all agents accurately perceive the entire environment at all times.

Belief

Realistic agents should have (possibly wrong) *beliefs* about the world and a theory of mind, the ability to reason about the beliefs of others. They should appear to suffer partial observability, updating their beliefs as they observe changes and not updating when they do not observe.

Planners for real world agents have modeled belief (e.g. Pollack 1986, Petrick and Bacchus 2002), but often make assumptions not helpful for storytelling in virtual worlds, like that agents are all cooperative (Grosz and Kraus 1996) or all competitive (De Rosis et al. 2003), or that ignorance is always bad (Bolander and Andersen 2011). Much work has also been done in strong autonomy narrative systems (e.g. Bates, Loyall, and Reilly 1992, Pynadath and Marsella 2005, Brinke, Linssen, and Theune 2014, Ryan et al. 2015, and others).

Several narrative planners model some character or audience beliefs to achieve suspense (Cheong and Young 2015), ideal story structure (Robertson and Young 2018), or deception (Christian and Young 2004), but their microtheories are not intended as general solutions to the belief problem.

Teutenberg and Porteous (2015) extended IMPRACTical to include agent beliefs, but limit theory of mind to 1 layer. It tracks what x believes, but to model what x believes y believes or deeper, it relies on a set of popular beliefs shared by all agents. Headspace (Thorne and Young 2017) models belief and uncertainty to enable stories where what an agent thinks will happen may not be what actually happens, but is also limited to 1 layer.

Ostari (Eger and Martens 2017) is an implementation of dynamic epistemic logic that allows an arbitrarily deep theory of mind. It enables planning with belief, but is designed for games with asymmetric information and can reason about all possible worlds, which may become unnecessarily expensive for strong story generation.

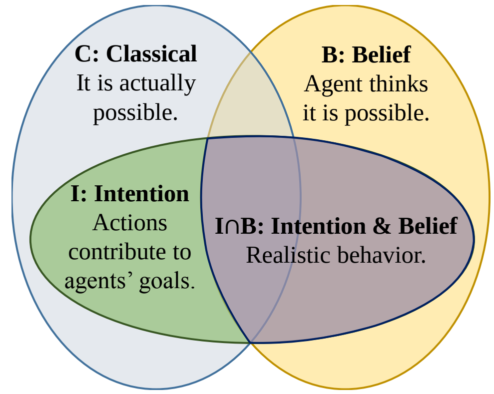

Figure 1: The plans a character can form to achieve a goal can be divided into three overlapping groups: C, I, and B.

We previously (2017) described a model for generating stories that reasons about both intentionality and belief using an arbitrarily deep theory of mind. In Kripke (1963) semantics, it models exactly one epistemically accessible possible world in any state; every agent has a specific (but possibly wrong) commitment to belief about every proposition, and uncertainty is not modeled. This contrasts with other models that allow multiple possible beliefs per character, modeling uncertainty via possible worlds, though at additional cost. Our model is meant for an omniscient, omnipotent planner in a virtual world that only needs to make characters *appear* like they have beliefs—we claim the model is often sufficient to create this illusion, even if it would not be sufficient for an actual agent in a real world setting.

Our model alters previous definitions of which character plans[1] are considered valid by including some plans that are impossible (provided that the character believes them to be possible) as well as some plans that involve actions taken by other characters (dubbed *anticipation*). In this paper we further investigate these definitions and show that the set of character plans enabled by the model, as well as the resultant stories, are believable to a human audience.

Before describing our evaluation, we first consider the space of all possible plans for a narrative planning problem and define which plans demonstrate intentionality, belief, both, or neither.

Narrative Plans

A narrative planning problem designates an initial state, an author goal which must be true by the end of the story, and a goal for each character. We assume a STRIPS-like (Fikes and Nilsson 1972) planning formalism, where actions have preconditions that must be satisfied before they can be taken and effects which modify the world state. The state space of a problem under our model is a graph whose nodes are states and which has two kinds of directed edges. A *temporal edge*

[1]Throughout this paper, we use the term "character plans" similar to Riedl and Young's "intention frames"—sequences of causally linked actions that explain how characters intend to achieve their goals. A character plan is a plan, but may not be the whole story.

$s_1 \xrightarrow{a} s_2$ exists from state s_1 to state s_2 via action a if the preconditions of a are met in s_1 and s_2 is the state that would result from applying the effects of a to state s_1.[2] The second type of directed edge, epistemic edges, are used to reason about belief. An *epistemic edge* $s_1 \xrightarrow{c} s_2$ exists from state s_1 to state s_2 for character c iff in state s_1 character c believes the current state to be s_2. The model prescribes exactly one outgoing epistemic edge for every character for every node.

Valid Classical Plans

A plan for some goal is a *valid classical plan* in state s iff there exists a path of temporal edges that starts at s and ends at a state where the goal is met. In Figure 1, the blue region labeled *C: Classical* (henceforth, just C) represents all valid classical plans—i.e. plans that would actually achieve a goal. Consider the simple pirate domain described in Figure 2, and the plan for the character Marley to have the treasure. A valid classical plan would be any sequence of actions that is actually possible and ends in Marley having the treasure (plans 1, 2, 3, and 4 in Figure 2). Plans 5 and 6 are not possible because Blackbeard has already dug up the treasure (see the initial actions under the "Story 2" heading).

Valid Intentional Plans

To reason about intentionality, every action must define a set of zero, one, or many *consenting characters* who are responsible for executing the action and who must have a motivation to act. Valid intentional plans must still achieve the author's goal, but may only do so via actions that are *causally explained*, i.e. contribute to the goals of the consenting characters who take them, as defined below.

An earlier action a_1 is *causally linked* to a later action a_2 via proposition p if a_1 has effect p, a_2 has precondition p, and no action occurring after a_1 and before a_2 has effect $\neg p$. A *causal chain* is an alternating sequence of n actions and n propositions $\langle a_1, p_1, ..., a_n, p_n \rangle$ such that for i from 1 to $n-1$, action a_i is causally linked to action a_{i+1} via p_i, and the final action a_n has effect p_n. A causal chain explains how a sequence of causally linked actions eventually achieves p_n. A *causal tree* is formed by any two causal chains for which there exists a number $i \geq 1$ such that p_i is different, but for all $j > i$, a_j and p_j are the same (i.e. the chains have a different beginning but the same end). A causal chain by itself is trivially a causal tree.

An *explanation* for a character goal is a sequence of actions after which the goal will be satisfied, and in which every action is part of a causal tree to achieve one of the propositions in that goal. An action is *explained* iff, for each of its consenting characters, there exists an explanation for that character's goal such that every action in the explanation is also explained. In short, a plan for some goal is a *valid intentional plan* in state s iff there exists a path of temporal edges that starts at s, ends at a state where the goal is met, and every action on that path is explained. In Figure 1, the green region I represents valid intentional plans—i.e. plans where every action causally contributes to achieving the goal. In Figure 2, plans 1, 3, and 5 are valid intentional plans because every action is explained for all of its consenting characters. Plans 2, 4, and 6 are not valid intentional plans because Marley has no reason to steal the treasure map from James.

Valid Belief Plans

Epistemic edges allow a narrative planner to model not only what is actually true but what each character believes to be true, what they believe others believe, and so on to any level of nesting. Given some current state s, let $\beta(c, s)$ denote what character c believes the current state to be. In other words, if the epistemic edge $s_1 \xrightarrow{c} s_2$ exists, then $\beta(c, s_1) = s_2$. A plan for some goal is a *valid belief plan* for character c in state s iff there exists a path of temporal edges that starts at $\beta(c, s)$ and ends at a state where the goal is met. In Figure 1, the yellow region B represents valid belief plans—i.e. plans which an agent would expect to work (even if they won't actually work). In Figure 2, plans 1, 2, 5, and 6 are valid belief plans because Marley believes that every step will be possible. Plans 3 and 4 are not, because Marley did not observe Blackbeard digging up the treasure and sailing to Port Royal, and thus she has no way of knowing that these plans would actually work.

Intentionality ∩ Belief

Earlier definitions by Riedl and Young (2010) and Ware and Young (2014) impose the constraint that a character's plan may only be composed of actions for which they are a consenting character. Ware (2014) noted this was necessary to avoid generating strange stories with an intentional planner that did not reason about belief. Because our model does reason about belief, we relax that constraint, allowing one character to anticipate the actions of another character and incorporate those actions into their own plan. Importantly, it is only reasonable for character *x* to anticipate character *y*'s action when *x* believes that *y* has a reason to take that action. Therefore our model extends the definition of an explanation as follows.

Let c be a character and g be that character's goal. An *explanation* for g is a sequence of actions after which g will be satisfied, in which every action is part of a causal tree to achieve one of the propositions of g, and in which c believes that every action is explained for all its consenting characters. A character *c believes an action is explained for another character d* if that action is part of any explanation for d's goal that c believes is a possible sequence. We say that *c believes an action is explained* iff c believes it is explained for all of its consenting characters. [3]

We can now formally redefine a valid character plan using the combined structures of intentionality and belief. A plan for some goal is a *valid character plan* for character c in

[2] The edge may actually exist even if its preconditions are not met in s_1; we call this type of temporal edge a *surprise edge* because it represents the occurrence of an action that a character believed impossible (2017).

[3] We also expand the definition of a causal link to include propositions about beliefs (2017). This allows character plans to include actions that inform, deceive, and gain the consent of others for future actions.

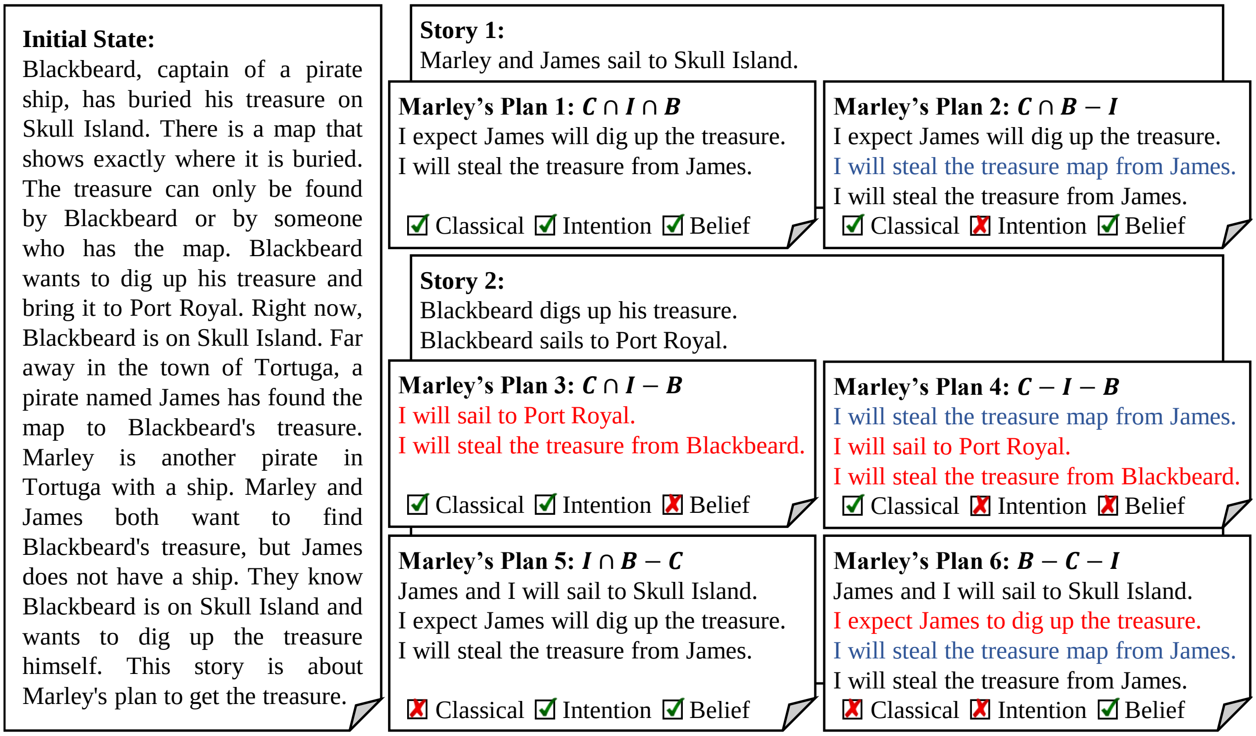

Figure 2: Six examples of character plans that are valid classical, intentional, and belief plans, along with the initial state as described to participants (abbreviated for space). Actions in blue do not contribute to Marley's goal. Actions in red are ones Marley does not believe are possible.

state s iff there exists a path of temporal edges that starts at $\beta(c, s)$, ends at a state when the goal is met, and, for every action on that path, c believes that action is explained.

We claim that when a virtual agent forms a plan to achieve a goal, the plans that will produce the most realistic behavior are those in the overlap of regions I and B, which we denote $I \cap B$. In Figure 2, plans 1 and 5 are both in this overlap because, regardless of whether the plan will actually work, Marley believes it will work and believes that every action is explained.

Evaluation

In this paper we present two separate evaluations. The first of these compares the believability of plans given by the spaces described in the previous section and illustrated in Figure 1. We claim that our definition for the region $I \cap B$ not only *can* generate plans that are believable, but also that the space of plans defined is, *as a whole*, more believable. We support this claim by comparing plans that have been randomly sampled from the entire set of plans considered valid in each group for a particular narrative problem, and showing that humans consider plans from the $I \cap B$ region more believable than plans from the other regions. The second evaluation (discussed in a later section) focuses on anticipation, specifically demonstrating that the anticipation enabled by our definitions is believable to a human audience.

Evaluation of Subsets in the Solution Space

We build on the results of Riedl and Young (2010) and Ware et al. (2014), who used specific stories generated by their models to evaluate believability. We extend that work to include beliefs and, instead of considering specific stories, consider all possible stories in the entire solution space of a planning domain.

We focus on *character plans* rather than stories in general because we are particularly interested in whether or not an audience finds a character's *behavior* believable. We assume that believable behavior will result from believable character plans, and that such behavior leads to more believable stories. In this evaluation, we explicitly prompt users to think of plans from the perspective of the given character by asking questions in the form: "Which of these plans makes more sense for Marley to be considering, given her goals and her current beliefs?" We show that human readers prefer those plans which demonstrate both intentionality and belief over those with only one of these features, or with neither.

Using the pirate domain shown in Figure 2, we generated all possible sequences of actions (up to length 7) that satisfy the goal of Marley having the treasure. For each of these sequences, we then consider every possible breakdown of that sequence into a *(story, plan)* pair, such that the first $n \geq 1$ steps of the sequence represent the "story so far", and the remaining steps represent Marley's plan to achieve her goal of having the treasure, after the first n steps have happened. For example, consider the 3-step sequence: 1) Marley and

James sail to Skull Island. 2) James digs up the treasure. 3) Marley steals the treasure from James.

This sequence will be broken down into two different pairs. In one pair, step 1 is the story, and steps 2 and 3 represent Marley's plan after the first step has occurred. This is the pair shown in Figure 2 as "Story 1" and "Marley's Plan 1", respectively. The same sequence also produces another pair, in which steps 1 and 2 are the story, and step 3 represents Marley's plan after the first two steps have occurred.

After generating all such sequences and pairs, we tagged each pair according to whether Marley's *plan*, in the state after the *story* has happened, represents a valid classical, intentional, or belief plan, as in Figure 2. We divided the set into four areas of interest: plans with both intentionality and belief ($I \cap B$), all plans with intentionality (I), all plans with belief (B), and plans with neither intentionality nor belief ($C - (I \cup B)$). These four sets included 3532, 4920, 4445, and 500 plans respectively. For each participant, we randomly sampled two pairs that had the same initial *story*, and presented them for comparison as follows.

First we translate all actions into short sentences using a simple natural language template. Actions that are part of the *story* are translated in present tense. Those that are part of the *plan* are translated into future tense; when these are taken by a character other than Marley, we use the form "I expect..." similar to the sentences in Figure 2.

Participants on Amazon's Mechanical Turk viewed an interactive web page which first displays the initial state (as shown in Figure 2) along with a picture of the character Marley and a clear indication that her goal is to have the treasure. Next, the interface displays each of the translated sentences from the *story* shared by the two sampled pairs. After reading these sentences, participants were asked three comprehension questions.

It then presented the two different *plans* and asked the participant to choose which plan made more sense for Marley to be considering. Because these plans were randomly sampled from two different regions of the solution space, we are effectively asking participants to compare the quality of those two spaces.

The participants were offered \$0.10 for completing the experiment and a \$0.40 bonus for answering the comprehension questions correctly. We discarded results for participants who did not answer the comprehension questions correctly. To encourage participants to consider both plans carefully, we told them they would only receive the bonus if they chose the "correct" plan; but in truth every participant who answered the comprehension questions correctly received the bonus, regardless of which plan they chose.

After filtering from the comprehension questions, we collected 60 answers for each pair of regions that we wanted to compare. We use $\bar{A}$ to denote the compliment of A, i.e. the set of all plans not in A. In each case we performed the binomial exact test (Howell 2012) to determine whether participants significantly preferred one set over the other. The results are presented in Table 1, using $>$ to denote *preferred*.

Results Plans meeting the definition of a valid belief plan (i.e. the plans in B) are significantly more believable than those not meeting that definition ($B > \overline{B}$). The same was true for intentionality, though the result ($I > \overline{I}$) was only marginally significant. Nevertheless, plans in I are significantly preferred over ones with neither intentionality nor belief ($I > C - (I \cup B)$).

Table 1: Evaluation Results

Comparison	p-value
$(I \cap B) > \overline{(I \cap B)}$	0.011
$B > \overline{B}$	0.024
$I > \overline{I}$	0.059
$I > C - (I \cup B)$	0.046
$(I \cap B) > (I \cup B) - (I \cap B)$	0.085

Participants did *not* show any significant preference for plans with only one feature ($(I - (I \cap B))$ or $(B - (I \cap B))$) over ones with neither ($C - (I \cup B)$). In other words, plans with valid intentionality but invalid belief are not significantly better than plans that are invalid for both. Plans with valid belief but invalid intentionality are not significantly better than plans that are invalid for both.

There was no significant preference for belief (B) over intentionality (I) or vice versa, but there was a marginally significant preference for plans with *both* over ones including only one of the two ($I \cap B > (I \cup B) - (I \cap B)$).

Finally, as we hypothesized, character plans with both belief and intentionality ($I \cap B$) were significantly preferred over the rest ($(I \cap B) > \overline{(I \cap B)}$). Under the assumption that more believable character plans result in more believable stories, we can conclude that our definitions result in the most believable stories for the solution space of the narrative problem defined in Figure 1.

Evaluation of Anticipation

The previous evaluation supports the new definitions of realistic character plans given by our model, which included stories demonstrating anticipation. Although we have demonstrated that humans find these stories believable, we may have prompted users to believe the stories simply by narrating them. The following experiment was designed to elicit stories from humans (via their actions in an interactive scenario) that include anticipation with minimal or no prompting. Here we test two claims: First, that humans will find it believable when virtual agents exhibit anticipation in accordance with our model; and second, that humans will anticipate the actions of virtual agents when acting as one of the characters themselves.

We built a text-adventure game using the Twine interactive fiction engine. The narrative is based on a simple planning domain, and all actions of the NPC are consistent with our model. We recruited 100 participants using Amazon's Mechanical Turk platform and paid them each \$0.25 to complete the scenario. Our experiment requires the NPC to have correct knowledge of the player's goal, so we structured the scenario as a game with a specific win condition rather than an open-ended narrative. We offered participants a \$1.00 bonus for winning the game to ensure that players actually possessed the goals we assign them. The following is a summary of the scenario we presented.

The player has entered a fishing competition against a virtual agent called "the crook". Each of the two competitors has a bag, initially empty, and may choose from a variety of actions that allow them to add fish to their bag. There are two locations in which to take these actions; the lake and the store. Competitors only observe each others' actions when they are in the same location. The competitors always know the number of fish in their own bag (the player's fish count is always displayed), but may have incomplete knowledge about the number of fish in the other's bag. The crook's belief about the number of fish in the player's bag is tracked using our model. The player's beliefs are not tracked; we enforce the partial observability constraint on the player by only narrating the crook's actions when the two are in the same location.

The competitors alternate taking actions, and present their bags to the judge when they are finished acquiring fish. The goal is to have more fish in one's bag than the other competitor when the judge reveals the results, but there is a twist! Upon handing in one's bag, either competitor (or both) may secretly switch the bags. If the bags are switched, the player's total will count as the crook's total, and vice-versa. If both competitors switch the bags, they will be switched twice. This action is always done in secret; it is never observed by the other competitor. This information is clearly stated at the beginning of the competition along with the other rules of the domain. It is also made clear to the player that: 1) the crook is willing to switch the bags if he thinks it will help him win, and 2) the crook believes that the player is *not* willing to switch the bags.[4] Before starting the competition, participants were asked comprehension questions about the rules. They repeated the process if necessary until they were able to answer all the questions correctly.

After completing the scenario, but before displaying the results of the competition, we asked users to explain their actions in free text. The hope was that users would clearly report, without any prompting, that they expected the crook to take certain actions they had not witnessed, and had thus incorporated those actions into their plan to win the game. (This question was also used to filter out participants who made it clear that they did not understand the rules.) Next we asked how many fish the participant believes are in the crook's bag. This can further highlight anticipation without prompting; if this number is different than the number of fish the crook acquired while the player was observing, then the player must have anticipated him taking some other action. Finally, we explicitly prompted the user to select which actions they believe the crook has taken, via a multiple choice question. We discarded the data of 18 participants who clearly demonstrated that they did not understand the rules. The tests below were conducted using data from the remaining 82 participants.

[4]This second constraint was added to prevent infinite nesting of anticipation: without it, the player could expect the crook to switch the bags, but could also expect the crook to expect the player to switch the bags and thus *not* switch the bags.

Results for Player Anticipation The scenario was designed such that winning the game could not be accomplished without anticipating some NPC action. Only 28 participants won the game, so we cannot claim that a significant number of people actively anticipated the NPC's actions correctly (in accordance with our model). To determine whether a significant number of participants demonstrated *any* anticipation, we combined the number who won the game (28), the number who did not win the game but did report the NPC having a different number of fish than what they observed him take (11), and the number who did neither of these, but when prompted, reported that they believed the NPC had taken some actions that they did not observe (24). These involve increasing amounts of prompting, but each demonstrate that the user found it reasonable to expect certain actions based on the NPC's goals. In total 63 out of 82 demonstrated some form of anticipation ($p < 0.0001$ using the binomial test).

Results for Believability of NPC Anticipation After revealing the results of the competition, we showed participants the list of actions the NPC had actually taken. In all cases, this included at least one action that was motivated by an anticipation of some player action. For example, the crook anticipates that the player will catch a fish, and therefore decides not to catch any fish and switch the bags at the end. We asked participants whether or not this story made sense for the crook. Out of 82 participants, 72 stated that the story made sense to them ($p < 0.0001$ using the binomial test). This provides further support for our claim that the model's simulation of anticipation is believable to human readers.

Discussion and Future Work

This paper evaluates a new definition of believable character plans for strong story systems that reason about intentionality and belief. We define when a character's plan is valid in the classical sense, for intentionality, for belief, and for both together. This new definition expands the space of valid plans to include ones that demonstrate anticipation. We showed first that the space of plans demonstrating both intentionality and belief is more believable than that of plans demonstrating only one of these features, or neither. Second, we showed that anticipation is believable to human readers when exhibited by virtual characters, and that humans themselves sometimes exhibit similar anticipation when acting as a character in the narrative.

Our model reasons about two of the elements of the Belief Desire Intention framework, so one obvious direction for future work will be *desire*. Currently, agents have no way of prioritizing some goals over others, which can lead to strange behavior when they have conflicting goals. Therefore, we should first review desire in the context of strong story systems and then adapt it to narrative planners. More importantly, the simulated concept of *desire* must be compatible with the current definitions of belief and intentionality.

Related to this is the need for *proactive anticipation*. Currently, when an agent achieves its goals, it stops planning, and thus reasoning about what other agents will do. Agents should continue to anticipate the actions of others, anticipating threats to their goals and proactively protecting them.

Acknowledgements

This research was supported by NSF award IIS-1647427.

References

Bates, J.; Loyall, A. B.; and Reilly, W. S. 1992. An architecture for action, emotion, and social behavior. In *in Proceedings of the European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, 55–68.

Bolander, T., and Andersen, M. B. 2011. Epistemic planning for single and multi-agent systems. *Journal of Applied Non-Classical Logics* 21(1):9–34.

Bratman, M. 1987. *Intentions, Plans, and Practical Reason*. Harvard University Press.

Brinke, H. t.; Linssen, J.; and Theune, M. 2014. Hide and Sneak: story generation with characters that perceive and assume. In *Proceedings of the International Conference on Artificial Intelligence and Interactive Digital Entertainment*, 174–180.

Cheong, Y.-G., and Young, R. M. 2015. Suspenser: a story generation system for suspense. *IEEE Transactions on Computational Intelligence and Artificial Intelligence in Games* 7(1):39–52.

Christian, D., and Young, R. M. 2004. Strategic deception in agents. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems*, 218–226.

De Rosis, F.; Carofiglio, V.; Grassano, G.; and Castelfranchi, C. 2003. Can computers deliberately deceive? a simulation tool and its application to Turing's imitation game. *Computational Intelligence* 19(3):235–263.

Eger, M., and Martens, C. 2017. Practical specification of belief manipulation in games. In *Proceedings of the 13th AAAI International Conference on Artificial Intelligence and Interactive Digital Entertainment*, 30–36.

Fikes, R. E., and Nilsson, N. J. 1972. STRIPS: a new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2(3):189–208.

Grosz, B. J., and Kraus, S. 1996. Collaborative plans for complex group action. *Artificial Intelligence* 86(2):269–357.

Howell, D. C. 2012. *Statistical methods for psychology*. Cengage Learning.

Kripke, S. A. 1963. Semantical considerations on modal logic. *Mathematical Logic Quarterly* 9(5-6):67–96.

Petrick, R. P. A., and Bacchus, F. 2002. A knowledge-based approach to planning with incomplete information and sensing. In *Proceedings of the International Conference on Artificial Intelligence Planning and Scheduling*, 212–222.

Pollack, M. E. 1986. A model of plan inference that distinguishes between the beliefs of actors and observers. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, 207–214.

Pynadath, D. V., and Marsella, S. C. 2005. Psychsim: Modeling theory of mind with decision-theoretic agents. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 1181–1186.

Riedl, M. O., and Bulitko, V. 2013. Interactive narrative: an intelligent systems approach. *AI Magazine* 34(1):67–77.

Riedl, M. O., and Young, R. M. 2010. Narrative planning: balancing plot and character. *Journal of Artificial Intelligence Research* 39(1):217–268.

Robertson, J., and Young, R. M. 2018. Perceptual experience management. *IEEE Transactions on Games*. (forthcoming).

Ryan, J. O.; Summerville, A.; Mateas, M.; and Wardrip-Fruin, N. 2015. Toward characters who observe, tell, misremember, and lie. *Proceedings of the workshop on Experimental AI in Games*.

Shirvani, A.; Ware, S. G.; and Farrell, R. 2017. A possible worlds model of belief for state-space narrative planning. In *Proceedings of the 13th AAAI International Conference on Artificial Intelligence and Interactive Digital Entertainment*, 101–107.

Teutenberg, J., and Porteous, J. 2013. Efficient intent-based narrative generation using multiple planning agents. In *Proceedings of the 12th international conference on Autonomous Agents and Multi-Agent Systems*, 603–610.

Teutenberg, J., and Porteous, J. 2015. Incorporating global and local knowledge in intentional narrative planning. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, 1539–1546.

Thorne, B. R., and Young, R. M. 2017. Generating stories that include failed actions by modeling false character beliefs. In *Proceedings of the 10th workshop on Intelligent Narrative Technologies at the 13th AAAI International Conference on Artificial Intelligence and Interactive Digital Entertainment*, 244–251.

Ware, S. G., and Young, R. M. 2014. Glaive: a state-space narrative planner supporting intentionality and conflict. In *Proceedings of the 10th AAAI International Conference on Artificial Intelligence and Interactive Digital Entertainment*, 80–86.

Ware, S. G.; Young, R. M.; Harrison, B.; and Roberts, D. L. 2014. A computational model of narrative conflict at the fabula level. *IEEE Transactions on Computational Intelligence and Artificial Intelligence in Games* 6(3):271–288.

Ware, S. G. 2014. *A plan-based model of conflict for narrative reasoning and generation*. Ph.D. Dissertation, North Carolina State University.

Young, R. M.; Ware, S. G.; Cassell, B. A.; and Robertson, J. 2013. Plans and planning in narrative generation: a review of plan-based approaches to the generation of story, discourse and interactivity in narratives. *Sprache und Datenverarbeitung, Special Issue on Formal and Computational Models of Narrative* 37(1-2):41–64.

Young, R. M. 1999. Notes on the use of plan structures in the creation of interactive plot. In *Proceedings of the AAAI Fall Symposium on Narrative Intelligence*, 164–167.