The AIIDE-17 Workshop on
Intelligent Narrative Technologies
WS-17-20

Sketching a Generative Model of Intention Management for Characters In Stories: Adding Intention Management to a Belief-Driven Story Planning Algorithm

R. Michael Young
School of Computing and
The Entertainment Arts and Engineering Program
University of Utah
Salt Lake City, UT, 84112

Abstract

Previous work on story planning has shown success in the generation of plans that are both intention-coherent and demonstrate aspects of inter-character conflict. However, the initial models of intention and conflict have been limited, in that they lack methods to generate story plots where characters drop sub-plans to achieve their goals in believably consistent and expressive ways and adopt new sub-plans in the face of plan failure. In current work, we have developed models of failed actions in stories that go hand in hand with erroneous belief models for character. Motivated by characterizations of rational agents' intentions as choice combined with commitment, we provide a framing of the plan generation process that is intended to show how characters form their own plans to achieve their own goals, act upon those plans until they feel that conditions no longer support their plans, and then re-plan in the face of adversity to achieve their goals. We show an example story plan that contains several types of character-based intention dynamics targeted by our approach.

Introduction

In stories, characters commonly adopt their own goals, form their own plans to achieve those goals, revise their plans in the face of unfolding beliefs about the world and drop their goals altogether when they feel its appropriate. For example, when the Impossible Missions Force is working to stop the departure of a transport plane carrying chemical weapons in the opening scene of *Mission Impossible: Rogue Nation* (McQuarrie, C. 2015), a team member hacks the plane's satellite uplink channel, intending to disable the plane's computer systems, preventing it from taking off. His plan involves grounding the plane by remotely shutting down its fuel pump. When he learns that he is locked out from access to the plane's mechanical system, he maintains his goal of stopping the plane, but changes his plan to involve shutting down the plane's electrical system. Finding the electrical system also inaccessible, he switches to a plan to disable the plane's hydraulics. When he learns these controls are encrypted, it seems to the team and the audience that there are no executable plans to achieve the team's goal. Finally, however, team leader Ethan Hunt appears from nowhere, directing a new plan where the hacker remotely opens the plan's cargo door, Hunt boards the plane, retrieves the weapons and saves the day.

There are no analyses of narrative corpora that can conclusively tell us just how frequently characters revise their plans and goals in stories, but this example appears at least weakly representative of a common phenomenon. In examples like this, the dynamics of what a character intends to make true in the story world and the way that she intends to make that condition true aren't simply emergent properties of a complex environment and the rational deliberation of agents operating within it. Rather, the intentional dynamics are designed specifically by authors for narrative effect, e.g., to build tension, to prolong efforts around goal achievement, or to highlight to a viewer the critical roles of key characters trying to achieve their goals within the unfolding story world.

These roles and others played by intentional dynamics within stories are central to many narrative functions. Because of this, the development of principled means to generate intentionally appropriate story advances the broader goal of automatically creating more natural and compelling narratives. In recent years, work on automated story generation has shown success developing planning-based generative methods (e.g. (Young et al. 2013; Porteous and Cavazza 2009; Coman and Munoz-Avila 2012)). Planning-based methods for story generation offer a number of attractive features, including guarantees of soundness and completeness and the natural representational fit between plan structures and the goal-directed activity that characters undertake inside narratives. Increasingly, however, researchers have identified limited expressive capabilities in previously developed plan representations when used to characterize storyline structure. Knowledge representations that are adequate to produce plans that control robot execution fall short in their characterization of a range of features commonly found in stories. Much work that has gone in to plan-based story generation (e.g. (Ware et al. 2014; Bahamón, Barot, and Young 2015; Teutenberg and Porteous 2013)) has sought to retain as many of the benefits of classical planning as possible while also increasing the expressive range of narrative generators.

One point where planning approaches are limited arises from their inability to generate plans containing intentional dynamics – that is, plans where characters adopt new intentions during the course of the story, form plans to achieve those intentions, act on those plans and revise both plans and goals in the face of new beliefs they gain about the world around them. In the work we describe here, we provide the design of an algorithm for story generation that explicitly plans for character intention adoption and character-driven planning, and uses a knowledge representation that provides a context for intention dynamics based on characters' beliefs about the story world around them. We build on a new planning model and knowledge representation called HEADSPACE (Thorne and Young 2017). The HEADSPACE knowledge representation allows for the explicit representation of characters' beliefs about the world and how actions depend upon and change character beliefs during execution. Our extension to the HEADSPACE algorithm produces story structure that has many of the advantageous properties found in previous plan-based approaches and produces intention dynamics that may more closely mirror typical intentional behavior of characters in stories.

As we describe below, the extension to HEADSPACE generates stories where

1. agents may adopt their own goals, form plans for their actions in pursuit of their goals and commit to performing the actions in those plans.
2. agents performing actions observe the dynamics of the world around them, including the consequences of others' actions, and they revise their commitment to their plans and goals in response to changes they observe.

Related Work

Following Bratman (Bratman 1987), Cohen and Levesque (Cohen and Levesque 1990) identify seven specific properties that must be satisfied by a reasonable theory of intention:

1. Intentions normally pose problems an agent, who needs to determine a way to achieve them.
2. Intentions provide a "screen of admissibility" for an agent's adoption of other intentions.
3. An agent tracks the success of its attempts to achieve its intentions.
4. An agent believes its intentions are possible.
5. An agent does not believe it will not bring about its intentions.
6. Under certain conditions, an agent believes it will bring about its intentions.
7. An agent need not intend all the expected side effects of its intentions.

Van der Hoek and his collaborators (van der Hoek, Jamroga, and Wooldridge 2007) set out a logical framework for characterizing intention revision. In their characterization, they link the process of intention revision to the interplay between other mentals states such as character beliefs about the future and about their own actions and abilities. Their approach sets out a broad set of dependencies between elements of an agent's mental state dynamics – beliefs and intentions – and demonstrates the connection between belief update and intention update that we are also advancing. In contrast, our work looks to provide a grounded implementation of a generative system capable of producing storylines that demonstrate a subset of the interconnections between belief and intent.

Initial inclusion of explicit representation of intentions of characrters within a planning problem was first defined by Geib (Geib 1994).

Early work on story generation leveraged AI planning to create plot lines, but were limited to the expressivity afforded in conventional real-world task-planning algorithms. To enrich impoverished plan representations, narrative planning research has incorporated additional constructs into the planning process to support aspects of character intention management. IPOCL (Riedl and Young 2010) added an explicit representation of *intention frames*, groupings of actions performed by a character in furtherance of a single goal. Each frame is established by a motivating action whose sole effect is the intended condition, and all character actions in plans produced by IPOCL belong to at least one intention frame. While all character actions in IPOCL plans can be seen as motivated towards a character's goals, characters in IPOCL plans never drop or adapt their plans, and the rationale for the addition of motivating actions is not based on character beliefs or desires. Extending IPOCL, Ware and Young (Ware and Young 2011a; 2011b) introduce a model of conflict wherein characters may undertake actions which thwart the intentions of other characters operating in the plan. The planning algorithm that they define, called Glaive (Ware and Young 2014), constrained plans so that they might also contain actions across frames of commitment that interfere with one another. These conflicts were maintained by the system and marked as conflicts. Actions that depended on conflict conditions or their downstream context were planned for but not executed. In their approach, intentions and the plans to achieve them are dropped just prior to the point at which conflict would make the plans' execution fail.

In this paper, we build our work on the HEADSPACE planner (Thorne and Young 2017), a new planning system that builds plans based on how characters believe the world to be rather than on how it really is. HEADSPACE works using forward-directed state-space search, at its core being a variant of FastForward (Nebel and Hoffmann 2001). This search starts at some specified initial state and generates successor states by considering all the domain actions that are executable from the state being expanded. The planner varies from conventional state-space planners in two ways. First, state descriptions and domain operators allow references to the *beliefs* of the character performing an action. Preconditions and effects may refer to conditions that are true or false in the physical world or may be believed to be true, believed to be false or unknown to the character performing an action. Second, building on Geib's distinction between precondition satisfaction and executability (Geib 1994), HEADSPACE actions can be included in a plan only when they are *appar-*

ently executable. That is, an action is only attempted by a character when the character *believes* the action's preconditions hold.

Actions whose epistemic and non-epistemic preconditions obtain execute correctly. But actions in a plan with one or more physical preconditions that do not obtain are *attempted*. That is, they are included in the plan, marked as attempted and the resulting world state is updated to reflect that the action's effects were not produced as expected. Greater detail about the plan construction process and the ways that characters become aware of their own action failures can be found in Thorne and Young (2017). Because of space limitations, the examples we provide here make use of HEADSPACE's modeling of character belief but will not exploit HEADSPACE's ability to generate actions that fail when beliefs are incorrect.

Representing Beliefs, Desires and Intentions in HEADSPACEX

In this paper, we extend the knowledge representation used by HEADSPACE to include desires and intentions. We connect HEADSPACE's existing capability to manage character beliefs with this extended representation to manage dynamic intention formation and revision driven by character beliefs. The resulting representation and the algorithm that uses it is called here HEADSPACEX.

Within HEADSPACEX, we represent a character's beliefs about the current state, about conditions that must obtain in the world in order for actions to execute and about beliefs that change as a result of an action's execution. These beliefs are non-nested – beliefs about the current world state only. Adopting terminology from Thorne and Young (2017), we refer to ground first-order literals that describe the physical state of the world (e.g., Holding(Bilbo,Sting)) as *material* literals, and ground modal literals characterizing a character's belief about a material literal (e.g., Bel(Shelob, Holding(Frodo, Sting)) as *epistemic* literals.

In HEADSPACE, a *belief state* characterizes the material literals that a character believes to be true and false, as well as those whose truth values that are unknown to the character.

Definition 1 (Belief State) *A belief state for some character c is a tuple $BS_c = \langle B_c^+, B_c^-, U_c\rangle$ such that B_c^+, B_c^- and U_c together form a partition of the material literals for the given story domain, where B_c^+ designates all the material literals that c believes to be true, B_c^- includes all the material literals that c believes to be false and U_c designates all the material literals that c does not believe to be true and does not believe to be false.*

We write $B_c^+(p)$ just when $p \in B_c^+$ in the relevant belief state for c. We use similar notation and intended meaning for $B_c^-(p)$ and $U_c(p)$. Throughout this paper, we use $\varphi(p)$ as a placeholder when referring to some epistemic literal without needing to specify which form it takes.

Desires

In this work, a character's desires describe ways that they want the world to be. Each desire is defined relative to a given context – when character c believes that condition p holds, then c will translate his or her desire into an intention to make p obtain. Formally, HEADSPACEX represents a desire as a) a partial description of world state that the character prefers, along with b) a contextual annotation indicating when the desire could become an intention. Desires are defined in the belief space of a character. That is, the desired conditions are expressed as desired beliefs, and the context for adopting a desire as a goal is also relative to characters' beliefs rather than material descriptions of a world state.

Definition 2 (Desire) *A desire is a tuple $D = \langle c, M, \varphi(g)\rangle$ where c is a character, M contains a set of mutually consistent epistemic literals called the* motivations *for $\varphi(g)$ and $\varphi(g)$ is a single epistemic literal of the form $B_c^+(g), B_c^-(g)$, or $U_c(g)$.*

HEADSPACEX doesn't require that a character's desires are consistent. A character may have inconsistent motivations for the same condition and/or may desire contradictory conditions. That is, a character may have a desire $\langle c, M_1, B_c^+(g)\rangle$ and another desire $\langle c, M_2, B_c^+(g)\rangle$, where $p \in M_1$ and $\neg p \in M_2$. Further, a character may hold both the desire $\langle c, M_1, B_c^+(g)\rangle$ and the desire $\langle c, M_2, B_c^-(g)\rangle$.

Intentions

Intentions are translations of desires into specific goals and plans to achieve them.

Definition 3 (Intention) *An intention is a tuple $I_{\varphi(g)} = \langle c, w, M, \varphi(g), P_{\varphi(g)}\rangle$ where c is a character, w is the world state where the intention is adopted, M is a set of conditions that motivated the adoption of $I_{\varphi(g)}$, $\varphi(g)$ is a single epistemic literal of the form $B_c^+(g)$, $B_c^-(g)$ or $U_c(g)$ and $P_{\varphi(g)}$ is a plan for achieving $\varphi(g)$ that c believes is executable from w.*

While the structure of the plans held by characters as part of an intention could take many forms, in this paper we adopt a model where these plans are structured as ground partial-order, causal link (POCL) data structures comparable to those produced by typical POCL planners (e.g., (Weld 1994)). [1] Each plan has a set of steps, with an initial step s_0 whose effects are just c's belief state at the world state where the intention was adopted, and a final step s_∞ whose preconditions are just $\varphi_c(g)$. Each plan has a set of ordering constraints defining a partial order on its steps, and a set of causal links of the form $s_i \rightarrow^{\varphi(p)} s_j$, where $\varphi(p)$ indicates that c believes that s_i establishes condition $\varphi(p)$ as an effect needed by s_j as a precondition.

Each of those steps' epistemic preconditions are satisfied by causal links within the plan labeled with the epistemic effects on c's beliefs contributed by steps in the plan. In this regard, each plan P is what we call an *apparent solution* for

[1] In this paper, we make no commitment to the specific planning algorithm that produces these plans. They could be, for example, POP-style planners (Younes and Simmons 2003), hierarchical POCL planners (Young, Moore, and Pollack 1994) or total-order state-space planners (Ware et al. 2014) that construct casual links in a post-processing phase after plan steps have been determined.

g by c. That is, c believes in the initial state of the plan that all the literals that act as source for causal links in P obtain, and for every step in the plan, c's epistemic preconditions for the step are satisfied by some causal link coming from either the initial state or from a condition that c believes is an effect of some prior step in P, linked to the precondition by a causal link. Finally, c believes that P establishes $\varphi(g)$ as the precondition to its final step via some casual link $s_i \rightarrow^{\varphi(g)} s_\infty$.

Finally, because each plan P marks a plan that may have been partially enacted by its character, each step in P is labeled as either executed or pending. For any given world state w where the plan is intended, zero or more pending actions in P are apparently executable for c in w. We call these steps the *apparent frontier* of the plan.

As we discuss below, intentions are always held by characters at particular world states. Just as a belief state holds the beliefs of a character, an intention state is an element of a world state used in HEADSPACEXto collect a character's intentions.

Definition 4 (Intention State) *An intention state IS_c for some character c is a set of intentions of the form $I_{\varphi(g)} = \langle c, w, M, \varphi(g), P_g \rangle$ such that, for each $I_\varphi(g)$, M and $\varphi(g)$ are drawn from a motivating desire $\langle c, M, \varphi(g) \rangle$ and $P_{\varphi(g)}$ is a plan for c with $\varphi(g)$ as its only goal.*

In HEADSPACEX, the characterization of a world state specifies the truth value of all material literals and belief states for each character. HEADSPACEX's world states also include, for each character, the intention state characterizing their current goals and plans to achieve them.

Definition 5 (World State) *Given a world frame $W = \langle GL, C \rangle$, a world state is a tuple $w = \langle T_w, F_w, BS_{c_1}, ...BS_{c_n}, IS_{c_1}, ...IS_{c_n} \rangle$ where T_w and F_w together form a partition of GL, where T_w designates all the ground literals that are true at w, F_w includes all the ground literals that are false at w, each BS_{c_i} designates the belief state for character c_i at w, and each IS_{c_i} designates the intention state for character c_i at w, where $1 \leq i \leq |C|$.*

Following Grosz and Sidner's (1990) terminology, when $\langle c, w, M, \varphi(g), P_{\varphi(g)} \rangle$ is in the intention state for c at world state w, we say that c *intends that* $\varphi(g)$, and that c *intends to* perform each action $a \in P_{\varphi(g)}$. For any effect e of any action a, $a \in P_{\varphi(g)}$, when e is a label on some causal link in $P_{\varphi(g)}$, we say that at w, c *intends that* e. Otherwise, we say that c considers e a *side-effect* of a.

We require a simple notion of logical consistency for characters' intentions. Specifically, for any character c and material literal g, at most one of $I_{B_c^+(g)}$, $I_{B_c^-(g)}$ or $I_{U_c(g)}$ can be in an intention state IS for c. Further, in any given intention state for c, all of c's character plans must be internally consistent. Specifically, if P_I is a plan that merges all of c current character plans by a) creating an initial state based on c's beliefs in the intention state IS at w, b) creating a final step containing the union of all the goals of all the plans in IS and c) creating causal links, orderings and steps from the corresponding components in each of the plans in IS, then P_I must be apparently executable for c at w.

Story generation with intention maintenance

The HEADSPACEX algorithm is shown in Algorithm 1 below. This is a revision to the HEADSPACE algorithm, slightly simplified here for space and updated to reflect the integration of intention management into character plans. HEADSPACEX iterates over a process constructing a plan in a forward-directed state-space search. Upon each iteration, HEADSPACEX adds a step to the plan's current final step. If that step changes a character's beliefs about the world, then those new beliefs may serve to translate one of the desires of that character into an intention. This translation process results in the construction of a plan to be followed by that character in order to achieve the character's new intention. Plans that are part of a character's intentions serve to filter and guide the character in selecting actions to take upon each iteration of the planning algorithm. New beliefs that result from actions may invalidate an intended plan's structure or the motivations for holding an intention, resulting in re-planning or the dropping of intentions.

HEADSPACEX's definition includes the following seven changes to HEADSPACE:

- **Specifying characters' desires** [lines 1 and 9]. Desires in HEADSPACEX are a part of a planning problem specification. They're used when adopting intentions, as desribed in the next bullet item.
- **Having characters adopt intentions** [lines 9 and 19]. Following the work of Riedl and Young (Riedl and Young 2010), intentions in HEADSPACEX are adopted as the result of a *motivating action*, a step in the plan that represents the mental act of adopting a character goal and the plan to achieve it. In HEADSPACEX, we extend the definition of enabled actions to include not only apparently executable actions, but also motivating actions that mark the translation of a desire into an intention. At the point where HEADSPACEX considers all enabled actions for inclusion as the next step in a plan under construction, all desires for each character are checked. For each desire whose motivations obtain in the current world state, the planner considers the desire *motivated*.

  For each motivated desire, the planner creates a new planning problem for a micro-planner responsible for creating a plan to achieve the desired outcome. [2] The planning problem starts with a world state based on the character's current beliefs and their commitment to the execution of all of their character plans. The goal state of the planning problem is just the desired outcome. If the micro-planner is unable to find a solution plan for this problem, then the motivating action is not considered enabled. If, however, a solution plan is found, then the motivating action is created and considered enabled along with any physical actions that characters believe they can execute in the current state.

[2] In this paper, we limit a micro-planner to constructing plans where the character is the sole agent performing actions, though a more general approach would have the character constructing plans that contract out parts or rely on anticipated actions by other characters.

- **Having characters respond to achieving their intentions** [lines 23-24]. Whenever a character c changes its beliefs in a new world state, the new belief state is compared to the goal of each intention held by c. If the new belief state supports the goal of an intention, then that intention is dropped (removed from the world state).
- **Restricting characters to act only in pursuit of their goal** [line 9]. In HEADSPACE, actions that are candidates for inclusion in a plan at a world state w are called *enabled* at w. As mentioned above, in HEADSPACE, only actions that are apparently executable by a character are considered enabled. In HEADSPACEX, we modify the definition of *enabled* actions to include just those apparently executable actions that are at the apparent frontier of one or more of a character's plans present in the current intention state.
- **Having characters drop unmotivated intentions** [lines 23-24]. Whenever a character c changes its beliefs in a new world state, the new belief state is compared to the motivations for each of c's intention. If the new belief state doesn't support an intention's motivations, then that intention is dropped from the world state.
- **Having characters revise their plans when they realize they're broken** [lines 26-27 and 31-32]. Whenever a character c changes its beliefs in a new world state, each of c's character plans are checked. If a plan has a causal link that spans the current world state (i.e., a link whose source step has already executed and whose destination step hasn't) *and* the link's label is inconsistent with the new belief state, that indicates that the character believes the link is threatened, making the step at the destination of the link apparently unexecutable. At this point, the system uses the character's micro-planner to create a plan that achieves the intention but contains no threatened causal links. If no such plan can be found, the system drops the intention from the world state.
- **Updating the success criteria for plan search** [line 3]. We modify the definition of a solution to a planning problem so that a plan is a solution just when the plan's final step establishes a world state in which all the planning problem's goals obtain *and* that there are no intentions in any of the intention states in that world state. Effectively, a plan is a solution just when it establishes the author's goals as well as all the goals of all the characters.

A Brief Example

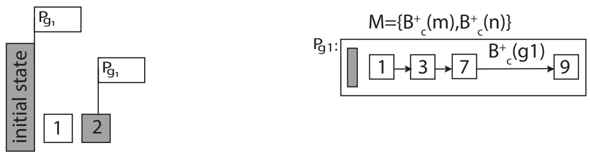
To demonstrate the range of intention dynamics and the interaction between belief, intention and execution in HEADSPACEX, we sketch a plan construction process for an abstracted plan in Figure 1.

In this example, the planning problem for HEADSPACEX includes an intention for character c to believe that $g1$ obtains. The intention is shown in the initial state and the character plan for the intention is given in the breakout labeled P_{g1} to the right of the figure. In Figure 1a, the plan is shown with its first step (the box labeled #1) and world state (the box labeled #2) already added to the

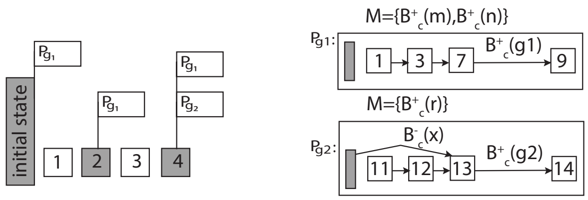
Algorithm 1 HEADSPACEX algorithm. For planning problem consisting of a set of desires D, an initial state w_0, a set of ground literal goals G, a set of ground operators GO and a plan heuristic ranking function H, call *HeadSpaceX*$(D, H, \langle\langle\perp, w_0\rangle\rangle, G, GO)$.

1: *HeadSpaceX*$(D, H, Plans, G, GO)$
2: Using heuristic ranking function H, rank all plans in Plans. Let $\mathcal{P}$ be the highest ranked plan in Plans.
3: **if** $\mathcal{P}$ is a solution **then**
4:   **return** $\mathcal{P}$
5: **else**
6:   Let w be w_k, the world state in kth (final) tuple in the current plan $\mathcal{P}$
7:   Let EnabledActions = $\emptyset$
8:   **for all** characters c **do**
9:     Let EnabledActions = EnabledActions $\cup$ *enabled*(c, w, GO, D).
10:   **end for**
11:   **for all** $a \in$ EnabledActions **do**
12:     **if** a is a physical action **then**
13:       **if** a is executable by c in w **then**
14:         Let w' be the world state resulting from c executing action a in world state w
15:       **else**
16:         Let w' be the world state resulting from c attempting but failing to perform action a in world state w.
17:       **end if**
18:     **else** $\{a$ is a motivating step for intention $I_{\varphi(g)} = \langle c, M, \varphi(g), P_{\varphi(g)}\rangle\}$
19:       Let w' be the world state resulting from adding $I_{\varphi(g)}$ to the intention state of w.
20:     **end if**
21:     Append $\langle a, w'\rangle$ to the end of plan $\mathcal{P}$
22:     **for all** $I_{\varphi(g)} = \langle c, M, \varphi(g), P_{\varphi(g)}\rangle$ in w' **do**
23:       **if** $\varphi(g)$ obtains in c's belief state in w' or some element of M no longer obtains in c's belief state in w' **then**
24:         remove $I_{\varphi(g)}$ from w'
25:       **else**
26:         **if** some causal link in $P_{\varphi(g)}$ spanning w' is threatened **then**
27:           replace $P_{\varphi(g)}$ in $I_{\varphi(g)}$ with a new plan $P'_{\varphi(g)}$ to achieve $\varphi(g)$ from w'. If no such plan can be constructed, remove $I_{\varphi(g)}$ from w'.
28:         **end if**
29:       **end if**
30:     **end for**
31:     Let $Plans = Plans \cup \mathcal{P}$
32:   **end for**
33:   Call *HeadSpaceX*$(D, H, Plans, G, GO)$
34: **end if**

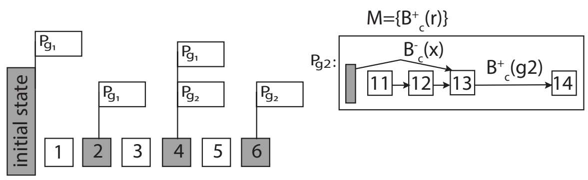
plan. In Figure 1b, the second step (step #3) from plan P_{g1} has been added to the plan, this step has the effect intended by c for P_{g1}, and also unintentionally adds beliefs to c's



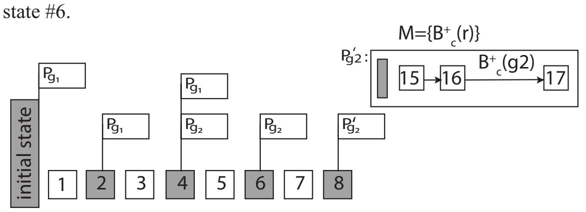
1a. In the initial state, c intends that $B^+_c(g_1)$ obtain, and has formed a plan, P_{g1} as part of that intention. Step #1 is the first step taken in this plan and the resulting world state #2 continues to support the intention.



1b. Here, the next action in c's plan P_{g1}, step #3, has been added to the plan. This step has the side-effect of changing c's belief in condition r, which results in the motivation for c's desire for $B^+_c(g2)$. As a result, world state #4 shows a new intention to achieve $B^+_c(g2)$ via plan P_{g2}.



1c. The planner adds step #5 next, a step executed by some character other than c. It has an effect that changes c's belief about condition m, one of the motivating conditions for c's intention for $B^+_c(g_1)$. As a result, the intention is no longer motivated and is removed in the resulting world state #6.



1d. Step #7 is also a step performed by some other character. It has an effect that changes c's belief about $B^-_c(x)$, which threatens a causal link in c's plan P_{g2} to establish $B^+_c(g2)$. Consequently, c's micro-planner creates a new plan for c's intention towards $B^+_c(g2)$ and substitutes that plan in the intention in world state #8.

Figure 1: A schematic representation of the planning process for a problem in HEADSPACEX. In the figures above, squares with white backgrounds indicate actions in the plan while squares with gray backgrounds indicate world states in the plan. Rectangles indicate character plans that are held as part of intentions in some world states. Insets show additional plan structure for character plans labeled P_{g1}, P_{g2} and P'_{g2}

belief state so that a new desire is motivated (all of the desire's motivation conditions are supported in c's belief state in world state #4). As a result, an intention (for c to achieve $B^+_c(g2)$) is created; world state #4 then shows both intentions for c: intending $B^+_c(g1)$ and $B^+_c(g2)$.

In Figure 1c, step #5 has been added to the plan. Here step #5 is a step performed by some character other than c. Step #5's execution changes c's beliefs in such a way that the motivations for c's intention to achieve $B^+_c(g1)$ is no longer supported. As a result, that intention is removed and no longer appears in the resulting world state #6.

In Figure 1d, step #7 is added to the plan, also a step executed by some other character. Step #7's execution changes c's beliefs so that c believes their plan to achieve $B^+_c(g2)$ is no longer valid. Specifically, an effect of step #7 asserts $B^+_c(x)$ threatening a causal link in P_{g2}. As a result, P_{g2} is no longer apparently executable for c in world state #7.

As a result, the micro-planner for c searches for and finds a new plan $P'_{g2}$to achieve $B^+_c(g2)$ from world state #7. This new plan is part of a replacement intention for c at world state #7 to achieve $B^+_c(g2)$.

Discussion and Future Work

The HEADSPACEX algorithm generates plans that have a number of desirable properties. HEADSPACEX plans align well with many of the properties put forward by Cohen and Levesque:

- When intentions arise, characters work to form plans to achieve them though the invocation of a character-specific micro-planner. The plans are formed in the belief space of the intending character, so that each character believes that their intentions are achievable. When an intention is motivated but no character plan can be formed for it, that intention is never adopted.
- Character plans are formed in the context of the character's existing plans. Character plans formed for new intentions are a) consistent with the character's beliefs and intentions and b) can leverage existing commitments to action to achieve the new goals.
- If characters act at all, they work to follow the plans that they form.
- The conditions under which characters believe they will bring about their intentions are explicitly represented in their plans, through the motivating conditions for the intentions and the causal structure of the character plans.
- The distinction between conditions that characters intend and those that are side effects is made clear in the plan structures formed by characters' micro-planners.
- When characters detect that some aspect of their plans has failed, they review those plans and are able to either a) revise their plans in light of their new beliefs or b) drop the intention altogether.

Work by Rao and Georgeff (1992) provided categories of agent commitment characterizing the contexts in which types of agents would drop intentions. They identify three types of commitment: blind commitment, single-minded commitment and open-minded commitment. In blind commitment, an agent (a) only considers new plans that are consistent with its current goals, (b) doesn't adopt new beliefs that conflict with its current intentions and (c) doesn't entertain any changes to goals that are inconsistent with its current intentions. In single-minded commitment, agents operate with comparable commitment to blind commitment agents except that they allow new beliefs that might invalidate the plans they already hold to achieve their goals. When adopting these new beliefs, single-minded agents drop their now inconsistent plans but maintain their old goals. In open-minded commitment, characters may also adopt new goals that are inconsistent with their old goals, causing them to drop both their old goals and the plans that were intended to achieve them. In HEADSPACEX , characters currently use single-minded committment. That is, they only consider new plans that are consistent with their current goals and don't adopt goals that are inconsistent with already adopted goals. However, characters in HEADSPACEXalways adopt new beliefs, overriding any inconsistent old beliefs and causing plan revision or goal dropping when characters current plans are invalidated by new beliefs. The single-mindedness of HEADSPACEX characters is not meant to be the final and defining characterization for all characters in the stories it produces. This approach is the initial behavior defined here, but clearly stories see a wide range of goal and plan commitment in characters, and our system will require extension to be able to support that range of behavior on a per character basis.

The structures produced by HEADSPACEX have clear parallels to the frames of commitment used in the IPOCL planner (Riedl and Young 2010) and to the intention frames in the Glaive planner (Ware et al. 2014) (minus, of course, the representation of conflict). These parallels suggest that people might recognize the same intention-driven features from these planners' plans in HEADSPACEXplans; future work will seek to explore those responses empirically.

Additionally, future work will seek to build more effective heuristics for forward-directed search in story domains, policies for when characters should drop, adapt and retain commitments, and methods for constructing character-specific plans that address narrative needs (e.g., (Bahamón, Barot, and Young 2015)) rather than just causal soundness.

Acknowledgements

This material is based upon work supported in part with funding from the Laboratory for Analytic Sciences (LAS). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the LAS and/or any agency or entity of the United States Government. Further, this material is based upon work supported by the National Science Foundation under Grant No. 1654651. I'm appreciative for this support from both these sources. I'd also like to thank Brandon Thorne for his contribution to the HEADSPACEinitial specification.

References

Bahamón, J. C.; Barot, C.; and Young, R. M. 2015. A goal-based model of personality for planning-based narrative generation. In *AAAI*, 4142–4143.

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

Cohen, P., and Levesque, H. 1990. Intention is choice with committment. *Artificial Intelligence* 42.

Coman, A., and Munoz-Avila, H. 2012. Creating diverse storylines by reusing plan-based stories. In *Working Notes of the ICCBR-12 Workshop on TRUE and Story Cases: Traces for Reusing Users' Experiences - Cases, Episodes, and Stories*.

Geib, C. W. 1994. The intentional planning system: ItPlanS. In *Proceedings of the International Conference on AI Planning Systems*, 55–60.

Grosz, B., and Sidner, C. 1990. Plans for discourse. In Cohen, P.; Morgan, J.; and Pollack, M., eds., *Intentions in Communication*. Cambdrige, MA: MIT Press. 417 – 444.

McQuarrie, C. 2015. *Mission Impossible: Rogue Nation*. Cruise, T. and Abrams, J.J. and Burk, B. and Ellison, D. and Goldberg, D. and Granger, D.: Paramount Pictures.

Nebel, B., and Hoffmann, J. 2001. The ff planning system: Fast plan generation through heuristic search. *Journal of Artificial Intelligence Research* 14:253–302.

Porteous, J., and Cavazza, M. 2009. Controlling narrative generation with planning trajectories: the role of constraint. In *Proceedings of Second International Conference on Interactive Digital Storytelling (ICIDS)*.

Rao, A. S., and Georgeff, M. P. 1992. An abstract architecture for rational agents. *KR* 92:439–449.

Riedl, M. O., and Young, R. M. 2005. Open-world planning for story generation. In *IJCAI*, 1719–1720.

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

Teutenberg, J., and Porteous, J. 2013. Efficient intent-based narrative generation using multiple planning agents. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, 603–610. International Foundation for Autonomous Agents and Multiagent Systems.

Thomas, J., and Young, R. M. 2010. Annie: Automated generation of adaptive learner guidance for fun serious games. *IEEE Transactions on Learning Technologies* 3(4):329–343.

Thorne, B., and Young, R. M. 2017. Generating stories that include failed actions by modeling false character beliefs. In *Working Notes of the AIIDE Workshop on Intelligent Narrative Technologies*.

van der Hoek, W.; Jamroga, W.; and Wooldridge, M. 2007. Towards a theory of intention revision. *Synthese* 155(2):265–290.

Ware, S., and Young, R. M. 2011a. CPOCL: A narrative planner supporting conflict. In *Proceedings, Seventh International Conference on Artificial Intelligence and Interactive Digital Entertainmen*.

Ware, S. G., and Young, R. M. 2011b. CPOCL: A narrative planner supporting conflict. In *Proceedings of the Conference on AI and Interactive Digital Entertainment*.

Ware, S. G., and Young, R. M. 2014. Glaive: a state-space narrative planner supporting intentionality and conflict. In *Tenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

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.

Weld, D. 1994. An introduction to least committment planning. *AI Magazine* 15(4):27–61.

Younes, H. L. S., and Simmons, R. G. 2003. VHPOP: versatile heuristic partial order planner. *J. Artif. Intell. Res. (JAIR)* 20:405–430.

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.; Moore, J. D.; and Pollack, M. E. 1994. Towards a principled representation for discourse plans. In *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, 946–951.