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

Experience Management with Beliefs, Desires, and Intentions for Virtual Agents

Rachelyn Farrell
Narrative Intelligence Lab, Department of Computer Science
University of New Orleans, New Orleans, LA 70148, USA
rfarrell@uno.edu

Introduction

Intelligent interactive narrative systems often use an *experience manager* to govern the behavior of non-player characters in a way that guides the story towards its author's agenda, which may be for entertainment, education, training, or other purposes. For such systems, a central challenge is creating *believable* virtual characters. The Belief Desire Intention framework (Bratman 1987) is often cited as a goal for researchers in this field; for characters to seem realistic, a human audience should attribute beliefs, desires, and intentions to them. Much of my prior work has focused on belief; my goal for the future is to finish the work on belief, and to implement a new model of desire and intention that explicitly reasons about characters' commitment to certain plans of action.

Experience management has been framed as a plot graph traversal problem which is jointly solved by the player and an AI experience manager (Bates 1992; Weyhrauch 1997). Some researchers (Nelson et al. 2006; Roberts et al. 2007; Thue and Bulitko 2012) have built on this work using MDP-based methods to find plot graph traversal policies. A particular challenge that I want to address, however, is the non-Markovian nature of stories. Different sequences of actions that lead to the same "state" may suggest very different paths for the future (Farrell and Ware 2016). A narrative representation that ignores history is therefore bound to suffer some limitations. For example, an NPC may walk back and forth between two locations because he is constantly switching between two active goals. Humans would consider this unrealistic because he never makes any progress toward either of his goals; but without tracking history, a Markovian experience manager would not recognize the problem.

Narrative planning is a common approach to formal reasoning about narratives and the causal relationships between narrative events (Young et al. 2013). Using a STRIPS-like (Fikes and Nilsson 1972) representation of a story domain, narrative planners search for a sequence of events that satisfies the author's goal. Later researchers used this as a groundwork to implement a version of intentionality, wherein characters are limited to taking actions that causally contribute to the achievement of their goals (Riedl and Young 2010; Ware et al. 2014).

Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

There have been several approaches to modeling agent beliefs, with tradeoffs in realism and efficiency (Bates, Loyall, and Reilly 1992; Ryan et al. 2015; Porteous, Cavazza, and Charles 2010; Eger and Martens 2017). In my previous work we presented a model for representing belief that is realistic, but inefficient (Shirvani, Ware, and Farrell 2017). I have since proposed a new implementation of that model which may improve its efficiency, and moreover could make it more useful for experience management because it explicitly tracks history. In the future I plan to first thoroughly compare the two spaces defined by the different implementations; and second, investigate several potential experience management techniques that this representation enables. In particular I am looking to incorporate the notions of *desire* and *commitment*.

Previous Work

We have proposed a narrative planning framework that supports infinite layers of nested belief (Shirvani, Ware, and Farrell 2017). That is, we can model not only what character x believes, but also what x believes y believes, and what x believes y believes x believes, and so on.

In this model, we use *epistemic edges* to represent the beliefs of characters: Let $\beta(c, s)$ denote the state that character c believes the world to be in when the current state is s. We use an epistemic edge $s_1 \xrightarrow{c} s_2$ to represent $\beta(c, s_1) = s_2$. We define a valid solution as one which satisfies the author's goal and for which every step is *explained.* A step is explained iff the step is part of a valid plan for every acting character in that step. We consider plans to be valid for a character c in state s iff there exists a path of temporal edges starting from $\beta(c, s)$ and ending at a state where c's goal is met, and every step on that path is explained. In short, characters can only take actions which they believe will achieve their goals, and can only anticipate actions of other agents when it is reasonable to do so.

Unfortunately, a naive implementation of this model is inefficient even in small domains due to the large number of nodes it often requires just to represent a single state. In attempt to make it more efficient I have proposed a new representation which models all the same information as the naive model, but is structured very differently:

- Nodes in this space represent actions (or "events") instead of states.
- Rather than explicitly modeling the values of all the fluents in each state according to all (possibly infinitely nested) agent minds, I instead explicitly track the event history of each state, and use a query function to determine on demand the value of any fluent according to any mind.
- I encode the logic of belief updates and observability into this query function. The function determines whether or not this event can modify the given fluent, and if so, whether or not the given mind would observe it. If so, it returns the associated value; if not, it queries the parent node. If no action in the node's history could answer the query, it queries the initial state using the procedure defined in the original model.

An important distinction between the two representations is when they consider two nodes to be duplicates. The original model considers two nodes equivalent iff they have the same values for all fluents according to all minds. The new representation I am proposing considers two nodes equivalent iff their histories are the same sequence of actions. Because of this we see an interesting pattern when evaluating the size of the spaces of the two planners.

For a domain and problem designed to test belief tracking, I used a prototype of each of the two implementations to generate the full state space up to a fixed depth (maximum number of steps). For small depths of up to 4, my implementation was more efficient than the original model; it used fewer nodes to represent the same set of stories. However, for depths greater than 4 the original implementation was more efficient because it detects "duplicate" nodes and avoids re-creating them.

Future Work

This sets the stage for my future work in multiple ways. First, I want to conduct a thorough comparison of the two implementations. I designed the test domain to be easily scalable so that I can test a large number of problems with different values for several parameters, including the number of agents, number of locations, number of items, number of goals, number of 1-layer wrong beliefs, number of 2-layer wrong beliefs, and so on. In my preliminary evaluation, the only parameter that seemed to give my new implementation an edge over the original was the added *layers* of wrong beliefs. That is, increasing the number of layers of a single wrong belief caused my implementation to become more efficient than the naive model for one extra depth (n=5) before being surpassed at n=6. This suggests that the new model may actually perform better in domains that require a lot of epistemic reasoning.

Second, I want to investigate several potential experience management techniques that are facilitated by tracking history. Namely, I want to separate the notion of *desire* from that of *intention;* the difference being that the latter implies a degree of commitment to taking some action. I believe that when a character takes an action that can only be explained by a particular plan, that character is expressing to the audience some amount of commitment towards that plan. If the character then deviates from that plan without a clear reason for doing so, it violates the audience's expectations and seems unrealistic.

A simple approach to representing desire might be to add the notion of rank to character goals, allowing the author to specify not only what a character wants, but how important that goal is compared to their other goals. This could be used to determine what constitutes a "clear reason" for abandoning a plan or adopting a new one; for example, because something occurs that newly enables a plan to achieve a higher ranked goal.

Finally, I want to further explore the believability of characters governed by systems with a complete belief, desire, and intention model. In a recent (forthcoming) evaluation of the believability of character plans that reflected realistic beliefs, intentions, both, and neither, we concluded (as we expected) that humans prefer character plans when both belief and intention are accounted for. However, as we did not expect, we were unable to show that either belief or intention by itself is significantly preferred over neither. This suggests that the combination of the three features may be more important than any one in particular, underlining the need for a balanced approach.

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.

Bates, J. 1992. Virtual reality, art, and entertainment. *Presence: Teleoperators & Virtual Environments* 1(1):133–138.

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

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.

Farrell, R., and Ware, S. G. 2016. Predicting user choices in interactive narratives using indexter's pairwise event salience hypothesis. In *Proceedings of the 9th International Conference on Interactive Digital Storytelling*, 147–155.

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.

Nelson, M. J.; Mateas, M.; Roberts, D. L.; and Isbell, C. L. 2006. Declarative optimization-based drama management in interactive fiction. *IEEE Computer Graphics and Applications* 26(3):32–41.

Porteous, J.; Cavazza, M.; and Charles, F. 2010. Applying planning to interactive storytelling: Narrative control using state constraints. *ACM Transactions on Intelligent Systems and Technology* 1(2):1–21.

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

Roberts, D. L.; Bhat, S.; Clair, K. S.; and Isbell Jr, C. L. 2007. Authorial idioms for target distributions in TTD-MDPs. In *Proceedings of the 22nd Conference of the Association for the Advancement of Artificial Intelligence*, 852–857.

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.

Thue, D., and Bulitko, V. 2012. Procedural game adaptation: framing experience management as changing an MDP. In *Proceedings of the 8th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment Conference*, 44–50.

Ware, S. G.; Young, R. M.; Stith, C.; and Wright, P. 2014. The Best Laid Plans.

Weyhrauch, P. W. 1997. *Guiding interactive drama*. Ph.D. Dissertation.

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.