Producing Immersive Interactive Narratives in Complex 3D Open Worlds

Alexander Shoulson
University of Pennsylvania
Philadelphia, PA, USA

Introduction

Interactive virtual worlds present a unique opportunity for simulation, training, and entertainment in settings that are otherwise prohibitive to recreate or control. One way to produce a more immersive experience is to populate these 3D environments with fully articulated virtual humans capable of exhibiting functional, purposeful behavior and complex interactions with other objects in the environment. Virtual worlds with intelligent populaces create a fertile setting for dynamic storytelling, where sequences of occurrences in the environment are designed and carried out with overarching *narrative intent* rather than simply emerging from the individual motivations of the digital actors. However, fully simulating every virtual actor is computationally expensive, and wastes resources on characters that may be less important than others.

We are interested in designing a system where authors can create unfolding stories in large, continuous environments full of ambient activity, and characters from the ambient populace can be "promoted" on an as-needed basis with richer behaviors that advance a story's plot. Our goal is to allow a user to freely navigate the environment and interact with a population of virtual characters that begin as simple pedestrians, but that can be upgraded at runtime to exhibit rich behaviors relevant to the narrative and the user's interests. This facilitates immersion by affording the user more complete agency and ownership of his or her impact on the world and its characters based which who he or she chooses to interact with, and in what manner.

This form of adaptive storytelling is specifically suited for *open world* or *free-roam* environments that place an emphasis on the user's ability to explore and interact with a variety of world features. With a "primordial soup" of undifferentiated characters that can transition from background ambience to pivotal actors in a long-form narrative, we develop the ability for a user to encounter compelling narratives at any point in the virtual world without the computational cost of fully simulating every character from the start. This draws inspiration from Alibi Generation (AG) (Sunshine-Hill and Badler 2010), but where AG focuses on justifying a character's past observed behavior, we want to use a pre-authored story as a blueprint for deciding which characters to "upgrade", and how to use them in future behaviors.

There are a number of limitations that must be surmounted before a robust open world interactive narrative system can be created. First, virtual actors must be able to exhibit a rich repertoire of individual capabilities, with specialized controllers for navigation, locomotion, gaze tracking, reaching, gesture animation, and others. Second, specialized control structures are needed to facilitate both individual character autonomy and the coordination of complex multi-character interactions with a high degree of control fidelity. A user must have tools to author character coordination at varying levels of granularity, from two characters exchanging a prop to large groups of characters participating in a riot. Finally, an artificial director must be able to reason about the objects and characters in the world, and the narrative impact of different character capabilities in order to produce action sequences that accomplish predetermined story goals and to decide which characters should participate.

Character Animation

In order to develop complex virtual worlds and fill them with highly capable virtual actors, an open-world storytelling system needs to be able to control fully articulated 3D virtual humans with a rich set of control capabilities. Character animation is a well studied field in computer graphics (Pettré, Kallmann, and Lin 2008) with techniques for navigation, locomotion, gaze tracking, reaching, gesturing, and others. This functionality is essential for replicating natural human motion, but often these systems are developed in isolation using vastly different architectures, making integration into a single character prohibitively difficult. The Agent Development and Prototyping Testbed (ADAPT) (Shoulson et al. 2013b) provides a single open-source end-to-end solution integrating all of these capabilities for complete, functional virtual characters capable of maneuvering an environment and performing the animations necessary for interesting interactions with objects in that environment. ADAPT creates an interface where high-level commands like *ReachFor*, *GoTo*, and *LookAt* are automatically converted to complex joint actuations on the character's body. This is ideal for authoring behavior involving natural human motion without concerning the author with the details of those articulated movements.

Multi-Actor Interactions

Like character animation, crowd simulation is a mature field that has produced a number of techniques for producing environments with numerous autonomous virtual characters. Traditional approaches for large-scale crowd simulation incorporate decision networks (Yu and Terzopoulos 2007), predictive models (van den Berg, Lin, and Manocha 2008), and other approaches to handle the characters' navigation and decision processes. Pelechano et. al. (2008) give a more detailed survey of additional work in the field. Crowd systems generally represent virtual characters as very simple abstractions like particles with mass and velocity, and characters rarely exhibit sophisticated interactions with one another. To facilitate the creation of environments where rich interactions can occur at any point, our event-centric authoring paradigm (Shoulson and Badler 2011) introduces *events*, which are reusable, temporary centralized behavior structures that co-opt the autonomy of their participants for the duration of an interaction and control them as limbs of the same entity with high control fidelity. Events are designed using Parameterized Behavior Trees (PBTs) (Shoulson et al. 2011) that, unlike traditional behavior trees, accept as parameters multiple actors or objects and coordinate the dispatch of behavior commands to those participants.

Events are a way to distribute the computational cost of behavior depending on each character's importance. Smart Objects (Kallmann and Thalmann 1998) enable some behavior distribution, but alone they are not well suited for carefully coordinated interactions like riots or distraction tactics that involve a large number of actors. When an event is designed, its author specifies what kinds of objects and actors can participate, and what criteria those participants should meet, based on factors like character personality, or a history of past event participation. When an event is instantiated by the virtual director, characters are recruited from the world and the event performs the interaction. Afterwards, the characters are released with residual information that can be used for further event selection and that "progressively differentiates" members of the population. Thus, a system with a homogeneous populace can involve those characters in events for as-needed complex behaviors without incurring the computational cost of richly simulating each actor for the entire narrative. Event-centric authoring also changes the metaphor for designing character behaviors – where traditionally behavior authors are concerned with what each character can do and why, an event-centric system places the author in the position of deciding what should occur in the world and letting the system decide how to make it happen.

Figure 1 demonstrates a simulation authored using ADAPT and events. A virtual Middle-Eastern populace inhabits a marketplace and exhibits ambient behavior such as wandering and visiting market stalls. Dynamic events are created for character interactions such as conversations or haggling over the prices of goods. The populace also reacts to the presence of a soldier-robot training team, exhibiting curiosity, caution, or anger at the team's presence. We use events to author training scenarios such as disrupting an angry protest, or cordoning off an area to search for contraband (Kapadia et al. 2013).

Interactive Narrative

Most interactive narrative systems have in common a virtual director, sometimes called a drama manager or experience manager. These are centralized controllers responsible for analyzing the state of the world and steering it towards some narrative objective (Riedl and Bulitko 2013). When the current narrative trajectory is interrupted by the actions of a human user interacting with the system, it is the virtual director's responsiblity to correct for the change and ensure that the story goals are still accomplished. Several approaches exist for story adaptation, including preemption (Magerko 2005) and scene selection (Thue, Bulitko, and Spetch 2008), but we are most interested in exploring how to alter the behavior of the characters in the world in response to a change in narrative trajectory. This facilitates immersion by enabling characters to react in meaningful ways to the actions of the user.

Dynamic character behavior has been explored with respect to adaptive narrative (Mateas and Stern 2003), but many interactive narrative systems focus on a small, fixed set of characters with the same level of behavioral fidelity. The traditional virtual director operates by maintaining direct control over every character in the system, which grows computationally prohibitive in vast open worlds with large numbers of semi-autonomous characters. Our worlds have many characters, some of which the user may never interact with, and so our solution to this problem is to define the director's action space not in terms of the instructions it can give each individual character, but in terms of what events it can select to occur in the world. If a character is never needed for any event, then will always be an ambient bystander that is comparitively cheap to simulate. If, however, in a second simulation instance the user decides to interact with that character, then our director can focus more acutely on that character's simulation and involve it in more interesting behaviors using events.

As an example, if the user sets fire to an important building, the director invokes a single "put out fire" event that coopts the autonomy of some nearby firefighters and directs them to the building with the appropriate behaviors. Conversel, if the user never decides to set that fire, then the firefighters will maintain some other cheap, baseline activities with no event to tell them otherwise. An event controlling a large number of characters has the same growth in cost as an event controlling two, so this is a scalable solution for highly directable character behavior in a large virtual populace. Conceptually, an event-based system is also easier for an author to reason about than the simultaneous direction of dozens or hundreds of individual characters. Within the event structure itself, a designer can have more precise control over *how* the firefighters would perform, rather than counting on the emergence of multiple autonomous characters interacting without external control. This approach could be described as a hybrid-hierarchical system, where the virtual director's finest grain of control is at at the level of events dictating atomic actions to characters, while characters outside of events can still act autonomously outside of the director's influence until needed.

Figure 2 shows results from our ongoing work on plan-

(a) (b) (c)

Figure 1: Simulation of a Middle-Eastern marketplace with autonomous characters. (a) Ambient activity in the main thoroughfare. (b) Two characters haggling at a merchant stall. (c) Characters in the crowd reacting to the training team.

ning in the space of narrative events. We use a generalized best-first planner to produce event-centric narratives at interactive rates that adapt to both helpful and confounding intervention from a user. Using events allows us to carefully coordinate complex character interactions with precise timing and synchronization to produce more natural behavior. We are able to populate a world with characters that exhibit a rich repertoire of animation capabilities without causing combinatorial growth in the planner branching factor (Shoulson et al. 2013a).

Future Work

Planning and heuristic search are a promising direction to explore in terms of developing reactive narratives that enable user agency while still maintaining cohesive structure. While these methods would produce most of the results desirable for creating narratives, there are a number of limitations imposed by the scale required for rich open worlds.

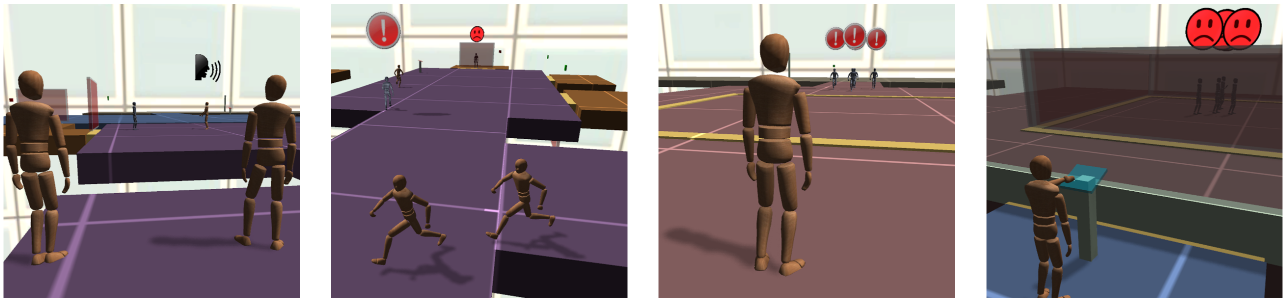
Computational Efficiency. One of the strengths of events is that they are authored, reusable structures that take their participants as parameters and can be performed by any set of actors that meet its requirements. This allows a virtual director to dispatch events to any appropriate group of characters in any location in the world. However, this flexibility has the side effect of increasing the branching factor of any heuristic search that considers the possibility of instantiating sequences of events. This is due to a combinatorial growth caused by the process of satisfying all possible collections of actors and props from the world that can participate. There are a number of ways to mitigate this growth, including (1) assigning narrative "roles" (like firefighter or school teacher) to characters in the world and requiring that each participant in an event have a matching assigned role, (2) ordering the selection of event parameters in such a way that the director first selects a location for the event, and then only selects participants within a proximity of that location, or (3) performing a random sampling of possible participants rather than an exhaustive search. All three of these possibilities sacrifice either flexibility in what events can occur where, or guarantees that an event can occur in the first place. Since these are not desirable qualities, there is room for further investigation.

Authoring Burden. In a traditional planning environment using the space of events, authoring an event requires not only designing the behavior to execute when the event is initialized, but also the pre- and postconditions for that event's execution. Not only does this increase the effort required for creating a large library of interesting events, but with more complex sets of behaviors, this action space metadata can be difficult to manually discern. Some events may represent more esoteric actions for which explicit effects can be difficult to estimate. To counteract these limitations, it is possible to model the virtual world at the level of Smart Object affordances. Then the system can exhaustively explore the space of affordance activations, and use that understanding to simulate events using a generated affordance transition function. This would enable an offline process to produce a model of the pre- and postconditions of each authored event without any manual annotation by an author. Figure 3 demonstrates preliminary work in a system to explore the space of Smart Object affordances.

Nondeterminism. Currently, events are designed with a set outcome in mind, though they are capable of success or failure. If system error or user input causes an event to produce an unexpected outcome, the entire story sequence must be invalidated and a new narrative must be generated. This limits the types of events that can occur in the world, prohibiting the creation of complicated events with unpredictable results. Furthermore, 3D character animation systems are inherently chaotic, particularly when large groups of actors are involved – systems like navigation and procedural reaching can fail unexpectedly or produce undesirable results. Since the goal of the system is to produce interesting, sophisticated narratives in well-populated worlds that react to user input, constantly generating new plans is not a desirable property of the system. A potential solution to accommodate more robust story generation is to integrate probabilistic planning methods into the heuristic search process when producing sequences of events to execute.

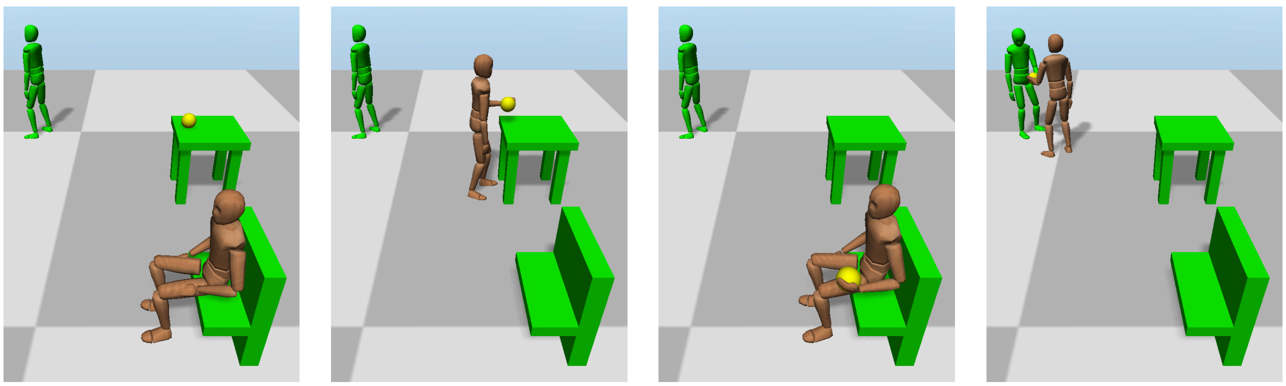
References

Kallmann, M., and Thalmann, D. 1998. Modeling objects for interaction tasks. In *Eurographics Workshop on Animation and Simulation*.



(a) (b) (c) (d)

Figure 2: Narrative planning with events to produce cooperative behaviors. (a) Two characters hide while a third distracts a guard, then run to open the door while the guard is lured away (b); (c) A character lures four guards towards him while a second character (d) presses a button to activate a trap.



(a) (b) (c) (d)

Figure 3: Exploring the space of Smart Object affordances. (a) A character activates an affordance to sit in a chair; (b) The character goes and picks up an object, then explores the effects of sitting back down in the chair with it (c) or handing that object off to another character (d).

Kapadia, M.; Shoulson, A.; Cockburn, J.; Solomon, Y.; Novich, C.; Xhu, Z.; Masand, H.; and Badler, N. I. 2013. Authoring an interactive, functional, purposeful, virtual populace. In *MIG*, In Review.

Magerko, B. 2005. Evaluating Preemptive Story Direction in the Interactive Drama Architecture. *Journal of Game Development* 3(2):25–52.

Mateas, M., and Stern, A. 2003. Façade: An Experiment in Building a Fully-Realized Interactive Drama. In *GDC*.

Pelechano, N.; Allbeck, J. M.; and Badler, N. I. 2008. *Virtual Crowds: Methods, Simulation, and Control*. Synthesis Lectures on Computer Graphics and Animation. Morgan & Claypool Publishers.

Pettré, J.; Kallmann, M.; and Lin, M. C. 2008. Motion planning and autonomy for virtual humans. SIGGRAPH 2008 Classes, 42:1–42:31. ACM.

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

Shoulson, A., and Badler, N. I. 2011. Event-centric control for background agents. In *ICIDS*, 193–198.

Shoulson, A.; Garcia, F.; Jones, M.; Mead, R.; and Badler, N. 2011. Parameterizing behavior trees. In *MIG*.

Shoulson, A.; Gilbert, M. L.; Kapadia, M.; and Badler, N. I. 2013a. An event-centric planning approach for dynamic real-time narrative. In *MIG*, In Review.

Shoulson, A.; Marshak, N.; Kapadia, M.; and Badler, N. I. 2013b. ADAPT: The Agent Development and Prototyping Testbed. In *I3D*. ACM Press.

Sunshine-Hill, B., and Badler, N. I. 2010. Perceptually realistic behavior through alibi generation. In *AIIDE*. AAAI Press.

Thue, D.; Bulitko, V.; and Spetch, M. 2008. PaSSAGE: A Demonstration of Player Modeling in Interactive Storytelling. In *AIIDE '08*. AAAI Press.

van den Berg, J.; Lin, M. C.; and Manocha, D. 2008. Reciprocal velocity obstacles for real-time multi-agent navigation. In *ICRA*, 1928–1935. IEEE.

Yu, Q., and Terzopoulos, D. 2007. A decision network framework for the behavioral animation of virtual humans. In *SCA '07*. Eurographics Association.