Adaptive Narrative: How Autonomous Agents, Hollywood, and Multiprocessing Operating Systems Can Live Happily Ever After

Jay Douglas and Jonathan Gratch

University of Southern California
Institute for Creative Technologies
Marina del Rey, California 90292
{jdouglas, gratch}@ict.usc.edu

Abstract. Creating dramatic narratives for real-time virtual reality environments is complicated by the lack of temporal distance between the occurrence of an event and its telling in the narrative. This paper describes the application of a multiprocessing operating system architecture to the creation of adaptive narratives, narratives that use autonomous actors or agents to create real-time dramatic experiences for human interactors. We also introduce the notion of dramatic acts and dramatic functions and indicate their use in constructing this real-time drama.

1 Introduction

```
EXT - BOSNIAN VILLAGE STREET - DAY

A young lieutenant is on his way to a rendezvous with the rest
of his platoon near the village square. His RADIO crackles out
an assignment.

                           RADIO VOICE
                    We need you here at the armory
                    as soon as possible.

But the lieutenant, still a few kilometers away, is preoccupied.
We SEE a traffic accident involving one of the lieutenant's
humvees and two local CIVILIANS. One, a YOUNG BOY, is seriously
injured and hovering near death. The second, his MOTHER, is
unharmed, but in shock and hysterical. A menacing CROWD gathers.
A CAMERAMAN for an international cable channel materializes,
shooting tape for the evening news.
```

This isn't a snippet of a Hollywood movie script. It's part of an interactive story based on real-life experiences of troops assigned to peace-keeping missions in the former Yugoslavia. In this tale, a lieutenant faces several critical decisions. The

O. Balet, G. Subsol, and P. Torguet (Eds.): ICVS 2001, LNCS 2197, pp. 100–109, 2001.

platoon at the armory is reporting an increasingly hostile crowd and requests immediate aid. The boy needs medical attention, which could require establishing a landing zone prior to a helicopter evacuation. The accident area must be kept secure, but excessive force or a cultural faux pas could be construed as a cover-up with major political consequences. The lieutenant's orders prohibit the use of weapons except in the face of an immediate threat to life or property. Unlike the movies, though, this cast, with the exception of the lieutenant, consists entirely of computer-generated characters, several of which are autonomous and cognitively aware. Instead of a Balkan village, the action takes place in a virtual reality theater with a 150-degree screen, 3 digital Barco projectors and an immersive, surround sound audio system equipped with 10 speakers and two sub-woofers (a 10.2 arrangement as compared with the typical 5.1 home theater system). The agents can interact with the lieutenant through a limited natural language system, and the mother agent responds to changes in her environment in a limited way. In spite of all this technology, we still cannot guarantee our storytelling environment will deliver engaging, dramatic content. This paper presents our work-in-progress on content production for the Mission Rehearsal Exercise (MRE) [1], one of the major research efforts underway at the Institute for Creative Technologies (ICT).

What we describe here is a multiprocessing operating system-like architecture for generating story world events unknown to the interactor, and the notion of dramatic functions, a method for gathering these events into dramatic moments. These low-level tools allow a human in the storytelling loop to create dramatic content in a real-time environment.

2 Motivation

Regardless of the medium, literary theorists describe the creation of narrative content as a three-step process: selection, ordering, and generating.[1] Out of all possible occurrences in the story world, some set must be selected for telling. Next, these occurrences must be ordered. There is, after all, no requirement that a story unfold chronologically. With apologies to Aristotle, we can begin at the end, then jump back to the beginning. If we choose to organize our narrative in this way then we require a crucial condition be met: the narrator must have a temporal distance from the events, where temporal distance means the events in the narrative occurred at some time other than the time of their telling.

In traditional storytelling this is no problem, for the telling of a tale comes after the occurrence of the events comprising it. If all occurrences unfold in real time, the processes of ordering and selecting are governed more by physical, rather than narrative, concerns. Our ability to create mystery, suspense, humor, and empathy are compromised.

Rather than abandon these powerful literary devices, our goal is adapting these techniques to the context of a real time environment. To do this, we need

[1] The steps are taken from Bordwell [2], substituting "generating" for his term "rendering" to avoid confusion with graphics terminology.

to maintain a separation between the time of events and the time the interactor learns about them. Once an event becomes "public," we forfeit the chance to foreshadow it, and recognizing foreshadowing opportunities is complicated by the interactor's freedom of choices. One apparent solution is providing the interactor with physically partitioned spaces into which he or she can move and ask "What happened here?" Events in these spaces would be known to us and temporally distant from the interactor so we could construct our dramatic moments. Such an approach leads to narrative consistency problems. Very quickly, we can wind up with a collection of moments each inconsistent with those of other spaces. What we suggest is creating "potential" foreshadowing opportunities to serve as fodder for our narrative content.

3 An Adaptive Narrative Architecture

A viable source for such foreshadowing opportunities presented itself unexpectedly, as side effects of a series of Wizard of Oz (WOZ) experiments.[2] A number of autonomous agents were replaced by human actors, and scenarios were played out under the invisible control of a (human) wizard. An agent command interface and two-way radios closed the behavior loop, giving the wizard control over agents and actors, as well as over the timing of interactions between them. The unexpected drama we encountered encouraged us to build an infrastructure for playing out multiple scenarios in parallel, under the control of a software *narrative agent* capable of cranking through the dramatic functions and turning events into dramatic experiences.

In life, unlike in books or films, the world goes on outside the pages we read or the images accessible to us on the screen. In books and film, moreover, readers and viewers only know what the author/director wants them to know. Not so in life (or adaptive narratives). If the interactor hears a noise behind a door, he or she should have the option of discovering the source. This may mean opening the door, asking another character, or seeing "through the door" via a surveillance camera. While the reconstruction of life may tax our abilities and our patience, our WOZ experiments pointed the way to a more user-friendly computer science model: the multi-processing operating system.

In UNIX-flavored systems, the user may have one process running in the foreground, but many others operating in the background. Similar effects were recognized in our WOZ experiments. We always had a foreground narrative, one involving the lieutenant, while other narratives, background processes in effect, played out somewhat unnoticed and asynchronously "offstage." These background narratives unwound according to their own scripts, and even though their actions were not the focus of the lieutenant's attention, their unfolding generated events the wizard used to increase or decrease the lieutenant's stress level.

[2] Although these experiments were performed to collect data for dialogue systems research, the results that intrigued us were those similar to Kelso [3].

Our developing system model relies on the abilities of autonomous agents to carry on their "lives" outside of the main focus and control of a central authority. By allowing these agents to execute their own scripts somewhat out of sight, the narrative agent accumulates invisible (to the interactor) events to support dramatic effects. Our background narratives run independently of each other, eliminating timing and contention problems. In our current design, the frenzy of the crowd at the accident scene, the situation at the armory, the attempts of a TV news cameraman to interject himself into the situation, and the status of men and equipment attached to the base commander all vary at their own speed, based on parameters established by the narrative agent. Thus, the cameraman agent might accede to a soldier's order back away from a shot if the cameraman's desperation factor is low (he's got his most important shots), or hold his ground if getting the shot means earning enough for his baby son to eat that night. While the lieutenant can "swap" foreground and background narratives, in same way as the fg and bg console commands can swap UNIX foreground and background processes,[3] a background narrative can always create an "interrupt," demanding attention from the lieutenant. For example, a background character can initiate a conversation with, or send a radio message to, the lieutenant, immediately bringing itself to the fore. Perhaps most importantly for training and education, modifying agent attitudes and the speeds of background narratives means each retelling of the MRE tale opens up new interpretations, each still causally linked to the interactor's behavior, each with its own feel and intensity, and each created without additional scripting or programming.

3.1 Choosing Background Processes

While any background narratives might suffice, at least in theory, we want to constrain them so the events they generate lend themselves to drama in the specific narrative we are working on. One way to accomplish this is to choose background narratives congruent with the interactor's goals. In the MRE, the interactor wants to: (a) evacuate the boy, (b) maintain local security, (c) fulfill his responsibilities relative to the platoon at the armory, and (d) perform in a manner consistent with good public and press relations. Starting with these goals, our scenario writers created five background narratives: (a) a threatening situation at the helicopter landing zone caused by any number of sources, from belligerent crowds leaving the armory to engine failure; (b) a mob scene at the accident site caused by provocateurs inciting an initally curious crowd; (c) an increasingly dangerous situation at the armory, where crowds are growing in size and their demeanor is growing more threatening; (d) an aggressive TV news cameraman who insists attempts to restrain him are actions preventing him from reporting the true story of military arrogance; and, (e) a deteriorating situation at base command, where demands on men and equipment may mean a shortage of relief troops, ground vehicles, and helicopters. On their own, the

[3] Swapping foreground and background occurs when the lieutenant interacts with a character in a background narrative.

background narratives are independent of each other; however, their common focal point is the interactor. He may alter the status of one narrative based on events occurring in another. All the narratives, however, affect the interactor's ability to meet his goals. They provide the fodder for what is typically known as the drama's "second act," the part where the protagonist embarks on a certain course, passes the point of no return, and finds his way strewn with obstacles.

For the narrative agent, however, the great advantage is the interactor's relative ignorance of events occurring offstage. Unless the interactor checks, he doesn't know the state of affairs at the armory. The narrative agent does, however, so if the interactor issues a radio call to the armory, the results are liable to come back garbled. The snatches of understandable transmissions may yield the wrong impression of the scene. Or, we might find the cameraman insistent on getting a shot based on something whispered to him by the boy's mother. A high ambient noise level, stress, misunderstandings, all are at the narrative agent's disposal for presenting a narrative to the interactor of the agent's own making.

We still, however, need guidance in selecting and ordering these events. For this we introduce the notion of dramatic functions.

4 Dramatic Functions

An old writer's adage says that if you plan to shoot a character in the third act of your play the audience had better see the gun in the first act. It's a reminder that unmotivated actions appear to come from "out of nowhere" in drama. In the same vein, coincidences, obstacles, misperceptions, misunderstandings and other storytelling tools sometimes test the bounds of credulity even while creating engaging narrative experiences. We perform acts and create situations in narratives that might be judged exaggerated in real life. Gerrig [4] discusses one theory of why we, as readers, viewers, or interactors, easily accept this distortion of reality and why it does not interfere with our enjoyment and involvement. Thus, in drama we find two types of acts: acts that occur as they might under real circumstances, and *dramatic acts*, which are manipulations necessary to create emotional responses. In our work we employ the notion of *dramatic functions* to construct dramatic acts. Representing drama as a set of base dramatic functions is one of the contributions of this implementation to virtual storytelling.

4.1 A Functional Approach

When one talks about describing functional elements in narratives the work of Vladmir Propp [5] springs to mind. Working with the Russian folk tale, he identified 31 actions played out by specific character types. The actions and characters generated hundreds of tales by simply changing settings or personalities. Propp's research also discovered these actions, if they appeared in a tale, appeared in strict order. Number five always occurred after any lower-numbered functions and before any higher-numbered ones. Because of this rigid structure, Propp's functions only generate a specific narrative form. Their greatest contribution to virtual storytelling, however, is the notion that narratives can be described in functional form.

Szilas [6] carried Propp's work several steps further by developing a set of generalized functions for constructing narrative content. The general direction of his research informs our own work in constructing our narrative content. Szilas's functions are broadly applicable to content behind the narrative, such as descriptions and chronicles. Our search is for something more middle of the road, an approach somewhere between the restrictiveness of Propp and the generality of Szilas. In addition, we want a system that allows us to reason about temporal relations as well as propositions and beliefs. Towards that end we asked the question: what makes drama different from real life?

4.2 Dramatic Function Notation

One of the characteristics of dramatic acts, and hence dramatic functions, is their time dependency. The villain and his machete do not materialize until it appears the occupants of the haunted house need only open the front door and escape. James Bond doesn't disarm the bomb until there is no time left on the timer. Not only do we need the classical notion of events occurring before or after one another, we must reason about how long the separation between event and knowledge of the event should be and deal with events that occur over time rather than instantaneously. Thus, we require a logic that not only admits time intervals, but one robust enough to describe such commonplace relationships as during, meets, shorter, and longer. In order to reason about events in their temporal context, we represent our dramatic functions along lines outlined by Allen [7].[4]

In his temporal logic, Allen describes three primitive functions in their logic: $\mathrm{OCCUR}(e,t)$, $\mathrm{OCCURRING}(r,t)$, and $\mathrm{HOLDS}(p,t)$.[5] OCCUR is true if event e occurs strictly over interval t (that is, for any t_i less than or equal to t, $\mathrm{OCCUR}(e,t_i)$ is false). $\mathrm{OCCURRING}(r,t)$ is true if process r takes place during interval t; however, it is not necessary for r to happen at every subinterval of t. A process is distinguished from an event because we can count the number of times an event occurs. $\mathrm{HOLDS}(p,t)$ is true if property p is true over interval t (and all subintervals). We add a new primitive to this collection, $\mathrm{ASSERT}(a,b,p,t)$. For this event, during interval t, a asserts to b that proposition p is true. If ASSERT is true then we can conclude that a succeeded in asserting p to b and the act took place during interval t. Note that the function makes no claim about whether b believes a.

---

4 Allen rigorously defines arithmetic on time intervals, as well as concepts such as before, after, during, and meets in his paper. For clarity, we omit his axioms and appeal here to intuitive definitions.

5 We use lower-case letters to denote variables and uppercase letters to denote bindings.

Finally, we define a new function BELIEVE(i, p, t), which is true if a human or agent interactor i believes that proposition p holds over interval t. Modeling belief is a non-trivial undertaking, so in our work we rely on a fairly restrictive definition: BELIEVE(i, p, t) is true during interval t if p is not contradicted by any information presented in the domain during interval t and ASSERT(a,i,p,t_p)is true, where t_p precedes and meets t.

We do not deny the definition lacks a certain sophistication, especially the first clause, which presupposes a deficit of domain knowledge on i's part. The danger of such an assumption is obvious in general. We find it acceptable in the present case, because a fundamental motivation behind the MRE is that the interactor knows very little about the non-military elements of the story world. If we restrict the use of BELIEVE to propositions ranging over this domain, the definition becomes manageable, if not quite reasonable.

In the next section we give the general definitions for two dramatic functions, reversal and snare. Later in this paper we will apply these functions to the MRE in a concrete example.

4.3 Reversal Function

In a reversal (often called a reversal of fortune), the interactor sees the attainment of a goal snatched away from her—usually moments before success is at hand. Thus, for all events E, where E is a goal of the interactor, such that P is the set of preconditions, and $|t_s|$ is the length of time that event E will take, iff

$$\text{HOLDS(P,}t) \wedge |t| \geq |t_s| \rightarrow \exists t_s\text{:OCCUR(E,}t_s\text{), where } t_s \text{ is a subinterval of } t.$$

In words, once the preconditions of E are satisfied, and remain satisfied during the time it takes for E to occur, E will occur. Thus, the interactor should *expect* a successful outcome, especially if there is no perceived threat rendering HOLDS(P,t) false.

The narrative agent's role is reversing this expectation. While the interactor expects E, the narrative agent plans

$$\text{HOLDS(P,}t_1) \rightarrow \text{HOLDS(}\neg\text{P,}t_2)$$

where $|t_1| + |t_2| = |t|$, $|t_1| < |t_2|$, and t_1 precedes t_2 and t_1 meets (is adjacent to) t_2. Since P becomes false during the interval in which E requires P to be true, the goal is thwarted.

4.4 Snare Function

A snare is a misrepresentation, a deception, usually one that deepens a mystery.[6] In our application, a snare represents an attempt by the narrative generator to lead the interactor astray (and thwart one of his goals) by deliberately presenting the world as it is not.

[6] See Barthes [8] for a discussion of snare and other function types.

Let P be the set of preconditions for E, where E is a goal of the interactor, I. By the reasoning above, the interactor expects E to occur because

$$\forall i\text{:BELIEVE(I,}P_i\text{,}t_s\text{)}$$

Let us also define P' such that

$$\forall i \neq j\text{:BELIEVE(I,}P_j\text{,}t_s\text{)} \wedge (\neg P_j \wedge \text{BELIEVE(I,}P_j\text{,}t_s\text{)})$$

In the snare, the narrative agent's role is to construct a P' based on P, such that the interactor believes Pj and ultimately expects E, whereas the truth is P' (and therefore $\neg E$).

5 Concrete Examples

To see how dramatic functions and background processes work together let's consider two possible sequences in the MRE. In the first, the interactor orders a helicopter evacuation of the boy. Let P be the conjunction of the four conditions: landing zone (LZ) is empty; LZ is marked with green smoke; LZ surrounded by one squad of troops; and, helicopter is over the LZ. When all these conditions are met, the helicopter can land.

The narrative agent must reverse E, the event "helicopter lands" at the last possible moment. The agent has, for this narrative, the following domain information: a crowd of 50 people are marching from the armory and are only a few blocks from the LZ; helicopters are complex machines and can develop problems requiring them to return to base and the base is not near the LZ. The narrative agent must search for a plan that results in $\neg P$. Which one to choose is a matter of style and dramatic familiarity. Mataes and Stern [9] suggest that when creating conflict one should choose situations referencing past occurrences. The story agent might check its history to see if, for example, the interactor was warned about overextending his troops (recommendation from platoon sergeant), or if the interactor was warned by the armory platoon leader that a crowd was heading towards the accident scene (background narrative), or if the base reported it was having trouble keeping its helicopters mechanically sound (background narrative). Since E is associated with an interval over which it occurs, the narrative agent can reason about not only how to create $\neg P$, but when to create it as well.

A common Hollywood use of the snare is the false ally, in which an antagonist presents herself to the protagonist as a friend while secretly working to foil the protagonist's plans. As an example of an MRE snare, consider what happens when a crowd of villagers forms at the accident site. Since neither side speaks the other's language, the crowd can only judge the soldiers' intentions towards the young victim by observing their actions; and, the soldiers can only judge the crowd's intentions by interpreting body language and tone of voice. Here is a situation ripe for misunderstandings. A restless crowd, unfamiliar with military medical techniques, on one side, nervous soldiers easily capable of misinterpreting emphatic, but harmless, gestures on the other. Certainly, it is in the interactor's best interests to keep the crowd calm.

Let E be the goal "trusted person tells crowd boy is getting good care," which we will denote by BELIEVE(C, a, B, t), where C is the crowd, a is an agent (possibly the interactor), B is the proposition "boy is getting good care," and t is some time interval over which the belief holds.

A priest, speaking broken English, materializes from the crowd and offers his services (background narrative). He will inspect the boy and report back to the crowd on the aid being administered. What the interactor does not know is the priest is a provocateur and will incite the crowd no matter how attentive the soldiers are to the boy's needs. The narrative agent expects the interactor will trust the priest (domain knowledge) and therefore will expect E is achieved over a long interval, t, once the priest talks to the crowd. However, the priest's words will be interpreted as inflammatory by the agents controlling the crowd's behavior. Sometime in t the crowd's fury will boil over (as determined by the narrative agent), hopefully surprising and distressing the interactor.

The narrative agent functions only if it recognizes the interactor's current goals, and this recognition represents another open issue. In a general solution, a plan model would provide feedback to the narrative agent about the interactor's intentions. We have no such mechanism, but we do have the structured domain of a military operation. In the MRE, the interactor's goals are typically expressed as orders, as when the interactor orders a helicopter evacuation. Recognition of these orders is necessary for other parts of the MRE, and the narrative agent can piggyback on these software modules, grabbing the orders as necessary and turning them into goals to be manipulated.

6 Future Work

What we've outlined here is only part of the story. The three-step narrating process as a model for a multi-processor-like storytelling environment, the notion of potential foreshadowing, and the encapsulation of primitive elements of drama into dramatic functions are promising tools; however, we believe a fully autonomous narrative agent is within reach of the state of the art. Notwithstanding our optimism, the future finds us with major obstacles to overcome. So far, we have not considered how the narrative agent combines the use of dramatic functions into a cohesive narrative. Mataes and Stern [9] provide a clue, but for complete generality, an agent will need to make far more subtle decisions, such as which dramatic function to choose for a particular effect, when to inject drama and when to allow the narrative room to "breathe," how far ahead to look when planning dramatic content, and how to recover when the interactor upsets the narrative agent's plan. Right now, a human still needs to interpret the interactor's goals in order to thwart them. The general recognition problem is still an open issue, and will most likely entail a model for recognizing narratives being constructed in the interactor's mind [4], [10], [2] combined with a mechanism for sharing these narratives, along the lines described by Young [11].

Despite these challenges, our research inches us closer to narrratives with more dramatic content than currently available, and to narratives that vary considerably in the "retelling," without the need for reprogramming or re-scripting. While there remains much work to be done, the combination of knowledge from both computer science and Hollywood offers exciting possibilities for the future of virtual storytelling.

Acknowledgements. The authors wish to thank David Traum and Michael van Lent for their assistance in clarifying and formalizing many of the concepts in this paper, and for their patience in reviewing and commenting on the manuscript.

References

1. Swartout, W., et al.: Toward the holodeck: Integrating graphics, sound, character and story. In: Proceedings of the Fifth International Conference on Autonomous Agents, ACM Press (2001)
2. Bordwell, D.: Narration in the Fiction Film. U of Wisconsin Press, Madison, Wisconsin (1985)
3. Kelso, M.T., et al.: Dramatic presence. Technical Report CMU-CS-92-195, Carnegie Mellon University (1992)
4. Gerrig, R.: Experiencing Narrative Worlds: On the Psychological Activities of Reading. Yale UP, New Haven (1993)
5. Propp, V.: The Morphology of the Folktale. 2nd edn. U of Texas Press (1988) Trans. Laurence Scott.
6. Szilas, N.: Interactive drama on computer: Beyond linear narrative. In Mateas, M., Sengers, P., eds.: Narrative Intelligence. AAAI Fall Symposium, Menlo Park, California, AAAI (1999)
7. Allen, J.F.: Towards a General Theory of Action and Time. In: Readings in Planning. Morgan Kaufmann (1990)
8. Barthes, R.: S/Z. Hill and Wang, New York (1974) Trans. Richard Miller.
9. Mateas, M., Stern, A.: Towards integrating plot and character for interactive drama. In Dautenhahn, K., ed.: Socially Intelligent Agents: The Human in the Loop. AAAI Fall Symposium, Menlo Park, California, AAAI (2000)
10. Branigan, E.: Narrative Comprehension and Film. Routledge, London and New York (1992)
11. Young, R.M.: Using plan reasoning in the generation of plan descriptions. In: AAAI/IAAI, Vol. 2. (1996)
12. Chatman, S.: Story and Discourse: Narrative Structure in Fiction and Film. Cornell UP, Ithaca and London (1980)
13. Genette, G.: Narrative Discourse: An Essay in Method. Cornell UP, Ithaca (1980) Trans. Jane E. Lewin.