Virtual Drama with Intelligent Agents

Josephine Anstey, Dave Pape
Department of Media Study
Stuart C. Shapiro, Vikranth Rao
Department of Computer Science and Engineering
University at Buffalo, Buffalo, N Y14260, USA
[jranstey, depape,shapiro, vbrao]@buffalo.edu

Abstract. This paper describes "The Trial The Trail" a work of interactive drama for a projection-based virtual reality system. The project marries a virtual reality authoring system and an artificial intelligence system.

1. Introduction

This paper describes *The Trial The Trail*, an interactive drama application for projection-based virtual reality systems, such as CAVEs®. We believe immersive virtual reality offers a uniquely powerful new medium for fictional experience, as it makes possible not only first person perspective but first person experience. The user is inside the virtual world with the other characters, rather than outside, viewing the world on a monitor and manipulating an avatar of herself. This facilitates the kind of psychological and cathartic fictional experiences that we are interested in producing..

The Trial The Trail is a work in progress that builds upon Anstey and Pape's previous experience creating an interactive fiction application for VR, *The Thing Growing*, [1] and upon Shapiro's experience developing intelligent agents and cognitive robots [2, 3, 4]. Our main areas of research are evolving interactive narrative strategies for virtual reality; and developing successful intelligent agents to act as characters in the drama and to guide a human user through virtual locations, moral choices, and emotional states.

In section 2, we give an overview of the use of agents in virtual drama. In section 3 we describe the dramatic concept behind *The Trial The Trail*. In section 4 we discuss the organizing principles, and in section 5 the implementation of *The Trail The Trail*. In section 6 we detail our progress so far, a first user-test, problems, and possible solutions.

2. Interactive Drama

In 1993, Laurel conceptualized IF, an Interactive Fantasy System [5], which would use agents to oversee plot development and interact with the user as characters. However, this idea has not moved far from the conceptual stage. Laurel's project *PLACEHOLDER* [6] was an example of narrative in a virtual environment (VE), but a human, not an agent, improvised the main actor. Other examples of successful immersive, interactive narratives include Disney's *Aladdin's Magic Carpet Ride* [7] and MIT's *KidsRoom* [8]. These projects have influenced the narrative strategies we use, but actor-agents were hardly used, certainly not to the extent envisaged by Laurel. Several research teams have focused on using intelligent agents for interactive, narrative experiences, including the *Virtual Theater for Children* [9], the *Alive Project* [10], and the *Oz Project* [11]. In these works the user is often addressed cerebrally - invited to create fiction with the agents, to direct them, or to marvel at simulations of autonomy and personality. The systems are impressive, the agents complex and believable. However, the kinds of experiences that they have produced do not necessarily produce nail-biting, fictional experiences for a user.

Probably the work at USC Institute for Creative Technologies, specifically the *Mission Rehearsal Exercise Project* has gone furthest in integrating a dramatic narrative, a VE, and intelligent agents. The goal of the project is pedagogic; to provide an immersive simulation of a difficult, real-world, peace keeping situation for an Army lieutenant. The project uses a variety of visualization tools for the VE and the SOAR AI system for the agents. The first version of this interactive scenario was very simple, however, the research team are planning a system called storyNet, which will include a director agent to oversee more complex plot development [12].

Mateas and Stern are currently working on a computer-based interactive drama, *Façade*, which promises to fulfill Laurel's vision. However, it is designed for a desktop. [13]

3. The Trial The Trail

The basic conceit of *The Trial The Trail* is a traditional quest or journey narrative. On the way the user encounters distractions and challenges. In our current version, the user meets a guide who helps her enter a walled zone and explains that she will reach her heart's desire in a room at the center of this zone. The first distraction the user encounters is a bed of reeds that flatten as she moves through them, encouraging play (figure 1). This is followed by a challenge where the user has the opportunity to rescue three imprisoned victims, which includes battling several morningstar-like creatures. The trope of rescuing a damsel in distress is debunked by the outcome. Instead of rescuing one of the victims the user encounters a clinging cat-like creature that does not want to let her continue her journey. If the user gets away she re-encounters the guide who leads her through an abandoned factory to an antechamber. Here, there are torch-like particle systems for the user to play with. We have storyboarded but not implemented two possible final outcomes: life or death.

We are trying to design interactive scenarios that will serve as indicators of the user's psychological traits that will then determine the outcome she gets. The distractions test whether the user is more intent on reaching the end, or in experiencing the journey. One of the purposes of the rescue scenario is to determine whether the user is cynical or romantic. Mateas suggests it is not the absolute number of interactive possibilities that make an interactive narrative or game compelling but the balance between the things you can do, and good reasons for doing them [14]. The very simple narrative drive of our project *The Thing Growing* severely constrained the number of things the user could do, but aimed to provide a satisfying reason for doing them. Our first challenge is to keep this balance in the more complex and open environment of *The Trial The Trail*.

4. Organizing Principles

Our basic process is to start with a simple narrative arc that we wish the user to traverse. We rely on intelligent agents to establish, populate, and maintain that narrative, draw in the user, and, by simulating emotion, stimulate her emotional and psychological responses. User testing tells us how to refine the narrative; where to add material to facilitate alternate routes through the story space; and how to build and augment the agent's behavior and responses. We will discuss our first user-test in section 6.

Figure 1: User in the reeds

In order to keep authorial control over pacing and the creation of surprise we base our VR dramas on a series of scenes with a distinct beginning, middle and end. The beginning introduces the environment and characters and an inciting incident reveals the protagonists' dilemmas or goals. In the middle the protagonists must overcome obstacles to their goals. The end contains the resolution. The narrative moves from scene to scene either as a result of the user's actions, or because a time limit has passed. At a finer granularity there are multiple ways in which the user can interact with agents or smart environments in each scene. These may have their own goals within the dramatic structure. They may act as guides to the user, or as opponents. The user has a certain amount of agency, but not absolute free will.

Our intelligent agents/actors must be believable characters in the environment. This means primarily, that they must be responsive to the user. But they can only be responsive if the application knows what the user is doing. In *The Thing Growing*, all the application knew about the user was information from the tracking system and controller. The narrative included the *Thing* character teaching the user a dance. The tracking information was interpreted to mean that the user was trying to dance, or not dancing, which in turn meant that she was obeying or not obeying the Thing. Then the character responded with praise or encouragement, criticism or anger as the scenario continued.

5. Implementation

Our VR drama projects are built using our VR authoring tool Ygdrasil [19]. This system is based on the OpenGL Performer scene-graph, and provides a framework for extension; application-specific modules (plug-ins) may be added to define behaviors for objects or characters. A text file system is used to assemble virtual scenes: all the models, objects, their locations and behaviors are described in the scene file along with messages to be passed between objects, in response to events. Ygdrasil also transparently incorporates the networking of virtual environments

The narrative structure of our virtual dramas can be created with the Ygdrasil scene file. Typically a narrative comprises sequences that explicate the story, periods of interactive possibility, and transitions from scene to scene. Each scene can be assembled under a switch; transitions from scene to scene and interactive events within each scene can be triggered by user proximity, the completion of specific events, or a timer. The text file serves as production manager for the story and can easily be edited and changed.

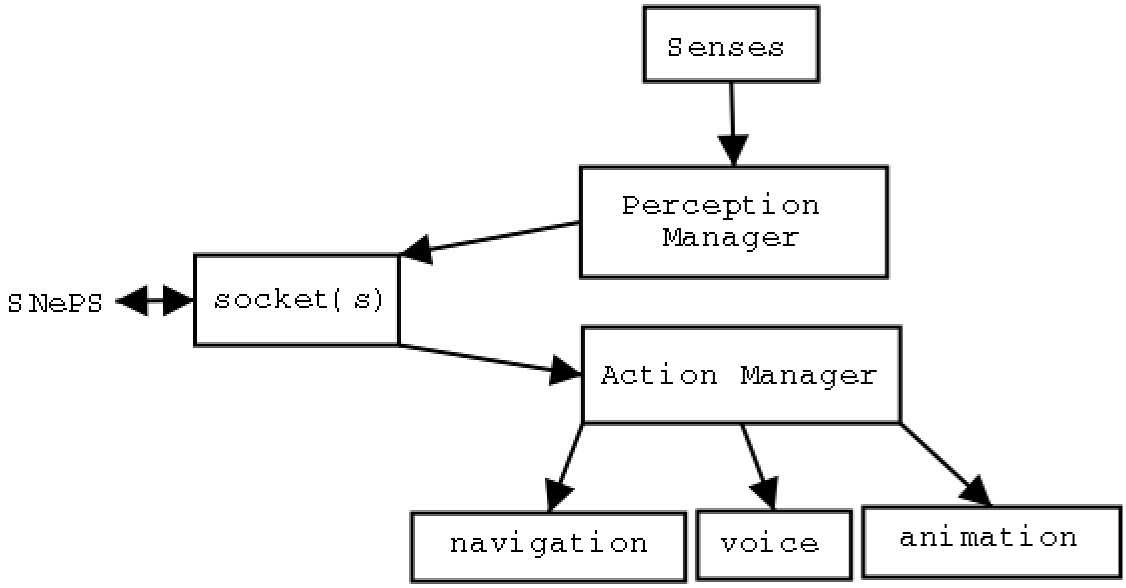

Figure 2

Specific modules were developed for the computer-controlled characters in *The Thing Growing.* The most complex character, the *Thing*, had modules controlling the behavior of its body (animation, voice and navigation) and its brain. Its physical behavior involved playing back motion-captured animations and spoken phrases. We recorded about 500 short phrases that encompassed the narrative possibilities of our story and included redundancies to avoid exact repetition. The *Thing* was animated in relation to the phrases it speaks so that its body language also reflects its mood and the meaning of the phrase. When the application is running the *Thing*'s brain picks the next appropriate animation sequence and phrase. While the agent is executing the action the brain checks the environment, the user, the narrative and its own mood in order to determine its next action. If necessary it can interrupt its current action and substitute another more appropriate one.

The agents for *The Trial The Trail* will be similar to the agents in *The Thing Growing*, although, in accord with the GLAIR architecture (described below), Ygdrasil will take care of the body's behaviors and SNePS will handle the mental functions. At this point we plan to continue using pre-recorded phrases for the voice, as current speech generating software does not do a good job of creating appropriate emotional tone. The phrases that the agent speaks will constitute the smallest narrative fragments of our story. The job of the actor/improviser agent will still be to assemble these fragments into a compelling dramatic form, with respect to the user's actions and speech, the environment and the arc of the narrative. Figure 2 shows a simplified diagram of our first plan for the agent on the Ygdrasil side.

We have conceptualized, but not implemented, a perception manager that will consolidate information coming in from the agent's senses. As yet we have not finalized what those senses should be. In our current version of the agent, messages triggered by events in the world are simply passed across the socket to the SNePs side. These triggers include information about the proximity of the user to the agent, and whether the user has taken some action that affects the world - for example opening the gate to release the imprisoned victims.

GLAIR (Grounded Layered Architecture with Integrated Reasoning) is a three-level architecture for cognitive robots and intelligent autonomous agents developed by Shapiro and his research group [16, 18]. GLAIR has been used in the design and implementation of a series of cognitive robots [3, 4, 17]. One of these has been implemented as an actual hardware robot, using an enhanced commercial Nomad robot. The others have been software-simulated robots acting in simple virtual environments. The cognitive robots all had the following capabilities: natural-language input, recognizing a fragment of English, based on a grammar written by the research group; a repertoire of primitive acts and sensory abilities; a knowledge-base of background information, domain knowledge, and acting plans; reasoning to answer natural-language questions, and to decide what to do and when; natural-language output to answer questions and to report on actions taken.

The three GLAIR levels are: the knowledge level (KL), implemented in SNePS, where "conscious" reasoning and action choice takes place; the Perceptuo-Motor Level (PML), which contains the physical-level representations of objects and the routines for well-practiced behaviors, including those that are primitive acts at the KL; the Sensori-Actuator Level (SAL), which controls the operation of sensors and actuators (either hardware or simulated). The PML is, in turn, divided into three sub-levels. The highest sub-level (PMLa) has been implemented in Common Lisp, and contains the definitions of the functions that implement the activity represented by KL primitive acts. The middle sub-level (PMLw) contains a set of Common Lisp symbols and functions defined in the World package which link to the lowest sub-level (PMLc), which is implemented in whatever language is appropriate for the hardware or software-simulated robot body. Sockets, such as illustrated in Figure 2 are used to connect the PMLw with the PMLc, which, in the case of *The Trial The Trail* is implemented in Ygdrasil, along with the SAL and the virtual world.

Our first experiment in creating a GLAIR-based SNePS/Ygdrasil agent has been a very simple actor-agent, the princess, who is one of the three victims that the user can try to help. We wrote an Ygdrasil module that acts as a TCP server at the body/virtual world end of the sockets and that can send and receive strings across them asynchronously. The mind end of the sockets can be different agents, or processes representing different modalities of one agent, such as speech, action, vision, etc. A simple Common Lisp program for the upper PML end of the sockets receives strings from the body/virtual world and send strings back. This was a logical extension of the Ygdrasil system's message system, which uses strings to pass messages between nodes. The messages from the virtual world come from events that had been triggered - for example by the distance of the user from the agent, or the user performing some action. We created a series of narrative zones for the agent including a zone for behavior to entice the user closer, a zone to explain her predicament, a zone to react to threatening sounds from the bad guys who have imprisoned her; a zone to explain how the user can help etc. In each of these zones were actions that the princess could perform. The initial SNePS KL was able to receive perceptual messages from the PML, reason upon them based on the narrative zone the princess is in, and initiate the appropriate rational acts in the form of messages. These messages would contain information about the narrative zone the agent was in and the action within that zone that should be performed. In some zones we wanted the agent to do only one of the possible actions, in other zones we wanted the agent to continue from one action to the next,

An immediate problem that arose was how to time the SNePS instructions. The SNePS agent does not need a full-blown model of time, but it does need to wait while the action is being performed before it sends a message about the next action. Therefore, we are currently postulating that the Ygdrasil side will send a message to the SNePS side when it is ready for a new action. When the SNePS side receives the message it will become aware, by asserting mental beliefs, that the preceding action has been completed and the next action can be initiated. We also need to migrate more knowledge to the SNePS side. At present, although it reasons based on the narrative zone, it is still blindly performing actions that it has no knowledge about. It simply has a rule that says "if I receive this message, send out an 'action' message with numbers representing a narrative zone and specific action". At present the Ygdrasil side sends the message on to the specific agent simulation.

A next version of the SNePS side is currently being developed that will contain two inherent modalities, speech and animation, which will be processed independently of each other. This version of the agent will be aware of its own repertoire of possible primitive actions, such as talking, looking, waving, jumping, cowering etc, and other more complex acts such as greeting, interacting with the user etc. The agent will represent the princess's role in a knowledge base, and based on incoming information from the Ygdrasil side about the world, the user, and the princess' current part, it can reason about the next action to take. The action manager on the Ygdrasil side will be dumber, it won't know what actions it can perform, it will only know to send information about the actions from the different modality-sockets to the animation, navigation or voice systems.

As the agent architecture evolves we must solve several problems. First, what are the senses of our actor-agent?

Clearly it needs to receive information about the user's position, orientation and actions; in order to move about it needs a sense of the structures and surfaces in the world. But in order to facilitate the drama it also needs to have a sense of where we are in the story, and when that changes. The new method of encoding awareness of completed acts will help with this. Second, we need a model of the user that will enable the agent to interpret her emotions and psychological state at any given point in the narrative, and more effectively mold the narrative to give her a unique experience. Third, it needs to be able to interact with other agents. In *The Thing Growing* none of the agents knew anything about any of the others. In order to set up scenes where they appeared to interact, we had to set up extremely inflexible scripted sequences where all the agents performed actions in a lock-stepped parallel fashion, and took turns to speak based on time. One of the main motivations for moving to a better AI system is to get round this limitation. Fourth, the system has to be fast and flexible enough to interrupt its reasoning process if things change abruptly in the virtual environment. Finally we will need to add speech recognition software to the VR display system. Although the primary mode for the agent will be to understand the speech input and reason from it, it will also have methods for dealing with speech that it does not understand. The creation of a strong personality for each character will help determine character- and context-appropriate ways for covering lack of understanding.

6. First User Test

This spring we constructed a first version of the virtual environment. In this version the world is a set of increasingly small concentric circles. The outermost circle is the wall around the zone; inside this is a hilly, tree-covered area, next a circle of reeds, then a river bottom, finally an island with a factory building. On their first entry into this building the users encounter a castle-like interior and the damsel in distress scenario, on their second entry they find an abandoned factory. Our deadline for this version was early June 2003, as we had scheduled a two-week show/user test in Hallwalls Contemporary Arts Center in Buffalo.

We used the Linux-based VR system from the Department of Media Study, UB, which consists of a 5x7 foot screen, passive stereo, a tracking system with sensors on the head and two hands, and a wand control device. Users came for an hour. First they experienced *The Thing Growing* project to get them used to the VR setup, then they experienced *The Trail The Trail.* Afterwards we interviewed them informally. Ideally we wanted one person at a time to experience the VR dramas, as we have observed that people may get performance anxiety if they are interacting with our projects in front of an audience. Occasionally two or three people who knew each other came together. A total of forty-three people attended the show/test over thirty hours. Two only had time to experience *The Thing Growing.* When people were in groups, one person was the tracked, main interactor, while the others participated in the experience, making suggestions about how to interact to that "driver."

The goals of this user test were two-fold; to test out a first version of our narrative arc and the psychological scenarios within it; and to do a Wizard of Oz experiment to flesh our the characteristics of the main guide character and help us conceptualize how to build the actor-agents. The main character, was improvised around a basic script by Anstey, who was networked into the environment from a separate desktop computer. Users saw an animated avatar of Anstey in the shared virtual environment, and communicated with her over microphones. During the test period Anstey changed the basic script several times, experimenting with different suggestions and directions to the user, to try improve her understanding of the narrative.. Other characters – the victims, bad guys, fighting morningstars, and clinging cat-like creature – were computer-controlled, but rather simply and within Ygdrasil. Due to the short deadline, and software license difficulties, we did not include any agents that were controlled by SNePS for this test.

Our test revealed that however free the guide was to improvise, she was not able to provoke dramatic tension or provide a clear narrative arc without the effective support of the virtual environment. We learned that we need to make changes to the visuals, the narrative, the characters, and the underlying structure of the psychological scenarios to provide this support.

Our graphics must be streamlined so that they do not contain visual clues indicating possible user actions that are not actually implemented. An example of this was the project's use of claw-like hands to represent the user's hands in the virtual environment - chosen because we already had a model. The claw worked very well as the users pushed the reeds aside. However when they came to the damsel in distress scenario some of the doors of the victims' prison were blocked with sheets of corrugated iron. Looking at their claws and the iron, some users concluded they could pull the iron aside, but they could not. User tests help us identify things people think of doing, so that we can then implement them or contrari-wise remove material which suggests they are possible.

Overall the narrative did not work in a way that seemed satisfyingly taut and dramatic. In order for the psychological scenarios to succeed, we need to make the interactive potential and significant choices of each section clearer. For example, people tended not to spontaneously play with the distractions such as the reeds, although they said they like them. Our challenge is to produce a story where people feel they have permission to play and wander, while the story still imparts some urgency. The damsel in distress scenario was not clear to people; some remarked that they did not know who was good and who was bad. This led some people to question whether they should release the victims, who may have been locked up for a good reason. Because people can drive where they want in the virtual environment, some people did not see that when they opened the gate of the prison, the victims did not escape but were captured by the bad guys and dragged further into the building. We have to rethink how to communicate the storyline to an audience that is far from captive. It is crucial that we strengthen understanding of the story because it provides motivations for the user's actions.

In this first version of the interactive drama, we had not evolved the behavior and personalities of the virtual characters very far. The test strengthened our belief that the addition of a character that we have storyboarded but not implemented will help us clarify the story-line. This character, the fellow traveller, takes the role of another person trying to reach their heart's desire. We can use conversations between this character and the guide to communicate information; we can use this character's behavior to suggest different behavior than simply following the guide's injunctions. We are considering working with actors on improvisations of these two main characters. The work of refining the virtual characters' roles and the narrative will go on in parallel with the development and general implementation of the SNePS side of the agents. We look forward to a second test with an improved narrative structure, and implemented SnePS agents in all minor roles.

References

[1] J. Anstey *et al*., The Thing Growing: Autonomous Characters In Virtual Reality Interactive Fiction In the Proceedings of IEEE Virtual Reality 2000, New Brunswick, NJ, March 18 - 22, 2000.

[2] S. C. Shapiro,, Embodied Cassie, In Cognitive Robotics, Papers from the 1998 AAAI Fall Symposium, Technical Report FS-98-02, AAAI Press, Menlo Park, CA, 1998, 136-143.

[3] S. C. Shapiro *et al*., Our Dinner with Cassie In Working Notes for the AAAI 2000 Spring Symposium on Natural Dialogues with Practical Robotic Devices, AAAI, Menlo Park, CA, 2000, 57-61.

[4] J. F. Santore and S. C. Shapiro, Crystal Cassie: Use of a 3-D Gaming Environment for a Cognitive Agent, In Papers of the IJCAI 2003 Workshop

[5] B. Laurel, Computers as Theater. Addison-Wesley, New York, 1993.

[6] B. Laurel et al, PLACEHOLDER: Landscape and Narrative in Virtual Environments In Clark Dodsworth Jr (ed.), Digital Illusion, ACM Press, New York, NY, 1998, pp. 181 – 208

[7] R. Pausch *et al*., Disney's Aladdin:First Steps Toward Storytelling in Virtual Reality In Proceedings of SIGGRAPH '96, ACM SIGGRAPH, New Orleans, LA, Aug 4-9, 1996, pp. 193-203

[8] A Bobick *et al*., The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment, In Presence Vol 8, Number 4, August 1999. pp. 367-39

[9] B. Hayes Roth *et a.*, Directed Improvisation by Computer Characters, Stanford Knowledge Systems Laboratory Report KSL-95-04, 1995.

[10] B. Blumberg, and T. Galyean, Multi-Level Direction of Autonomous Creatures for Real-Time Virtual Environments In Proceedings of SIGGRAPH 95, ACM SIGGRAPH, Los Angeles, CA, August 1995. pp. 47-54.

[11] M. Mateas, An Oz-Centric Review of Interactive Drama and Believable Agents, Technical Report CMU-CS-97-156, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA. June 1997

[12] W. Swartout, *et al*., Toward the Holodeck: Integrating Graphics, Sound, Character and Story In Proceedings of 5th International Conference on Autonomous Agents, Montreal, Canada, June 2001

[13] M. Mateas and A. Stern, Towards Building a Fully-Realized Interactive Drama, The Fourth International Digital Arts and Culture Conference, Providence, Rhode Island, April 26-28, 2001 http://www.stg.brown.edu/conferences/DAC/subs_in/Mateas.html

[14] M. Mateas. A preliminary Poetics for Interactive Drama and Games. In SIGGRAPH 2001 Electronic Art and Animation Catalog, ACM Press, Los Angeles, CA (1999) 51-58

[15] S. C. Shapiro and the SNePS Implementation Group, "SNePS 2.6 User's Manual", Department of Computer Science and Engineering, University at Buffalo, 2002.

[16] S. C. Shapiro and W. J. Rapaport, The SNePS Family, Computers & Mathematics with Applications 23, 2-5 (January-March, 1992), 243-275, Reprinted in F. Lermann, Ed, Semantic Networks in Artificial Intelligence, Pergamon Press, Oxford, 1992, 243-275

[17] S. C. Shapiro and H. O. Ismail, Symbol Anchoring in a Grounded Layered Architecture with Integrated Reasoning, *Robotics and Autonomous Systems* Special Issue on Anchoring, 2003, in press

[18] H. Hexmoor *et al*.,, Embodiment in GLAIR: a grounded layered architecture with integrated reasoning for autonomous agents, In D. D. Dankel II and J. Stewman, eds. Proceedings of the Florida AI Research Symposium (FLAIRS '93),) The Florida AI Research Society, Pensacola, FL, 1993 2325-329.

[19] D.Pape, *et al*., Ygdrasil - a Framework for Composing Shared Virtual Worlds, In Future Generation Computing Systems, Elsevier Press, 2003, in press