Not your Grandmother's Game: AI-Based Art and Entertainment

Michael Mateas

Computer Science Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
michaelm@cs.cmu.edu

Abstract

AI-based art and entertainment opens new possibilities both for game design and for AI. For games, it points the way to intelligent entertainment that functions as High Culture. For AI, it points the way to expressive AI, a new viewpoint that can inform and direct AI research. These twin claims are discussed in light of the concrete examples provided by three AI-based art and entertainment systems: Subjective Avatars, Office Plant #1, and Terminal Time.

Introduction

Most current computer games fall into one of two camps: shoot-em-ups in which the implicit narrative is kill or be killed and adventures in which a first-person protagonist solves a series of puzzles to accomplish a goal. AI-based art and entertainment has the potential to move beyond these two forms, opening up new interactive expressive forms that can play the same role in culture as literature, cinema, or visual and conceptual art. In addition, AI-based art and entertainment can open a new research agenda in AI. The application of off-the-shelf AI techniques is not enough – novel AI research motivated by the needs of artistic expression is necessary. In this paper I will briefly describe three AI-based art and entertainment projects in which I have been involved. I will then discuss these twin claims, that AI-based art and entertainment can be the carrier of high culture and that it can serve as a new AI research agenda, in light of the concrete examples provided by the three systems.

Subjective Avatars

The goal of the Oz project (Bates, 1992) at CMU is to build dramatically interesting virtual worlds inhabited by believable agents - autonomous characters exhibiting rich personalities, emotions and social interactions. In many of these worlds, the player is herself a character in the story, experiencing the world from a first person perspective. Typically, the player's representation within the world -

her avatar - is passive. The avatar performs actions as fully specified by the player and reports events (by, for example, rendering a 3D scene or generating descriptive text) in a pseudo-objective manner (*pseudo*-objective because any description encodes the bias of the world author). An alternative is a subjective avatar (Mateas 1997): an avatar with autonomous interpretations of the world.

Why Subjective Avatars?

I want the user to step into the shoes of a character, experiencing a story from this new perspective. In this manner the user gains an empathic understanding of a character by *being* this character. In non-interactive drama (movies, theater), an audience is able to gain insights into the subjective experience of characters precisely because the experience is non-interactive; the characters in the drama make decisions different from those that audience members might make. In an interactive story, how will a user gain insight into the character she is playing when she is controlling this character's actions? If she were to immediately begin acting out of character, she will derail the story, effectively preventing any insight. With a subjective avatar, the hope is that if the user's avatar filters and interprets the world in a manner consistent with the character, the user will begin to feel like their character, gaining a deeper understanding of the message the author wants to convey. The avatar becomes an additional artistic resource for authorial expression.

I've experimented with subjective avatars within the Oz text-based world. The text-based world accepts commands from the user and presents the world to the user in a manner similar to text-based adventure games.

Subjective State

In order for the avatar to provide a subjective interpretation for the player, it responds to activity in the world by maintaining subjective state. Currently, the avatar's subjective state consists of emotional state (emotional responses to events) and story context.

To maintain emotional state, I make use of Em (Neal Reilly, 1996), the Oz model of emotion. Em is integrated with Hap (Loyall and Bates, 1991), a reactive-planning language specifically designed for writing characters. In Em, emotions are generated primarily in response to goal

processing events and attitudes. Em generates emotions as goals are created, as beliefs change about the likelihood of goals succeeding or failing, and as goals actually succeed or fail. At any given moment, an agent's emotional state will contain several emotions with non-zero values. Over time, Em decays emotions. In order for the avatar to have goal processing emotions, it must be processing some goals. Since the avatar doesn't directly take action on its own, its goals are all passive. Passive goals wait for some event to occur in the world in order to succeed or fail.

In addition to emotion processing, the avatar keeps track of where it is in the story. This is done to organize the avatar's goals and simplify the writing of behaviors. At any given moment, the avatar is pursuing some set of goals. The behaviors associated with these goals are watching for certain events or sequences of event to happen in the world. At different points in the story experience, the same event may cause different reactions in the avatar (or no reaction). Explicitly maintaining a story context pushes the context information into the tree of active goals instead of requiring this information to be included in the precondition of every behavior.

Narrative Effects

Once the avatar is maintaining a subjective state, it must express this state in such a way as to affect the user's experience. The primary effect I've experimented with is manipulating sensory descriptions. Sensory manipulations are implemented as a set of Hap behaviors which render descriptions of events as a function of the subjective state. For example, imagine that the player-character (the character controlled by the human user) is afraid of a character named Barry. Barry, a manager in a fast food restaurant, is about to chew out the player. Without the subjective avatar, this would be rendered as follows in the Oz text-based world: "Barry is speaking to you. Barry's voice says 'wait a minute there, buster.' Barry goes to the counter area. Barry is no longer in the window area." The subjective avatar I've implemented for this world would render this exchange as follows: "With a vindictive gleam in his eye, Barry snaps 'Wait a minute there, buster.' Barry marches toward you from the drive-up window station." This description is generated by a narrative rule that matches on the current subjective state of the avatar (in this case, fear), and the current activity in the world. The important thing to note is that the same "objective" events in the world (Barry saying "wait a minute there, buster" and walking toward the player) would be rendered differently if the avatar felt differently (for example, as a result of previous events in the experience). Also, in a multi-player dramatic world in which multiple avatars are present, each player would experience a different rendering of the same event, depending on their differences in subjective state.

Subjective Avatar Conclusion

A subjective avatar is like an inverse user model. A user model watches a user's actions so as to learn a model of the user. A subjective avatar, on the other hand, has an author given model of the character. The avatar actively manipulates a user's experience so as to try and make the user feel the same way as the character. The avatar thus becomes an active expressive resource available to dramatic world authors.

Office Plant #1

Walk into a typical, high tech office environment, and, among the snaking network wires, glowing monitors, and clicking keyboards, you are likely to see a plant. In this cyborg environment, the silent presence of the plant fills an emotional niche. Unfortunately, this plant is often dying; it is not adapted to the fluorescent lighting, lack of water, and climate controlled air of the office. Office Plant #1 (Boehlen and Mateas, 1998) is an exploration of a technological object, adapted to the office ecology, which fills the same social and emotional niche as a plant. Office Plant #1 (OP#1) employs text classification techniques to monitor its owner's email activity. Its robotic body, reminiscent of a plant in form, responds in slow, rhythmic movements to express a mood generated by the monitored activity. In addition, low, quiet, ambient sound is generated; the combination of slow movement and ambient sound thus produces a sense of presence, responsive to the changing activity of the office environment. OP#1 is a new instantiation of the notion of *intimate technology*, that is, a technology which addresses human needs and desires as opposed to a technology which meets exclusively functional task specifications.

Comparable in size to a generic office plant (10x10x33 inches), OP#1 consists of a large bulb surrounded by metal fronds mounted on a base. The bulb, a hammered aluminum sphere, can open and close. Mounted on a stem, it can also rise above the fronds and remain in any intermediate position. The fronds, made of copper wire, sway slowly, moving individually or in synchrony. In addition to physical movement, OP#1 has a voice; it produces sound using a speaker housed in the bulb. These sounds provide the plant with a background presence. The force-delivering stepper motors are concealed in the lower part of the plant, discernible, though, through semitransparent plexiglas. The window in the bottom of the base would promise to reveal the inner workings of the plant, but shows, instead, a scene composed of rocks, sand and moving counterweights: the *datarium*. The datarium is the equivalent of a vivarium. In the datarium, however, the only life forms are data driven lead counterweights moving in and out of the rock and sand garden.

OP#1 is an experiment in building a companion agent, an agent that is always present, monitoring and commenting on user activity. As a constant companion, OP#1's actions must be subtle; an overactive agent would quickly becoming irritating to a user. OP#1's design attempts to maintain an air of mystery, providing a recognizable physical manifestation of a user's email

activity, but not by means of a simple one-to-one mapping. OP#1 should provide the user with an opportunity for *contemplative entertainment*, opening a window onto the pattern of a user's day.

OP#1's primary view of user activity is via their email. All incoming email is assigned labels which correspond to the social and emotional role of the message, such as FYI, intimate, chatty, request, etc. Any one email may be assigned several labels. Categorization is performed by means of Naïve Bayes and K-nearest neighbor text classification (Mitchell, 1997). Naïve Bayes classifications are made by applying Bayes law to the conditional probabilities of word occurrence given a document class and the prior probabilities of document classes. The prior terms are obtained by observing frequencies in labeled training data (an offline learning step). K-nearest neighbor classifications are found by returning the majority label among the k-nearest neighbors of the query document in the document space.

The plant's behavior is controlled by a Fuzzy Cognitive Map (FCM) (Kosko, 1997). In an FCM, nodes representing actions and variables (states of the world) are connected in a network structure (reminiscent of a neural network). At any point in time, the total state of the system is defined by the vector of node values. The action associated with the action node with the highest value is executed at each point in time. The values of nodes change over time as each node exerts positive and negative influence (depending on connection weights) on the nodes it is connected to. As email is classified, activation energy is given to appropriate nodes in the network, priming OP#1's dynamics.

OP#1 is a collaboration with roboticist and artist Marc Boehlen.

Terminal Time

Terminal Time (Domike, Mateas, and Vanouse, 1998) is a machine that constructs ideologically-biased documentary histories in response to audience feedback. Terminal Time is a cinematic experience, designed for projection on a large screen in a movie theater setting. At the beginning of the show, and at several points during the show, the audience responds to multiple choice questions reminiscent of marketing polls. Below is an example question.

Which of these phrases do you feel best represents you:

A. Life was better in the time of my grandparents.
B. Life is good and keeps getting better every day.

The audience selects answers to these questions via an applause meter – the answer generating the most applause wins. The answers to these questions allow the computer program to create historical narratives that attempt to mirror and often exaggerate the audience's biases and desires. By exaggerating the ideological position implied in the audience's answers, Terminal Time produces not the history that they want, but the history that they deserve.

Critique of Traditional Historical Narratives

Terminal Time is an exploration and critique of familiar authoritarian narratives of history. Representation is at the heart of this endeavor. The mission is to dramatize to the viewing public that the truth of history is not simple and linear. Although there are undeniable historical facts, perspective is a critical element of historical understanding. By creating fact-based histories, clearly driven by point of view, the project reveals the constructed nature of all historical representation, in particular the popular genre of the television history documentary.

Representation of Content in Terminal Time

Terminal Time represents ideological bias using a goal-tree formulation of ideology similar to Carbonell's (Carbonell, 1979). The goal tree is modified as the audience answers the polling questions. Pursuit of goals in the goal tree causes the system to search its knowledge base of historical episodes, looking for episodes which can be slanted to support the current ideological bias. In addition to historical episodes, the knowledge base also contains rhetorical devices which are used to connect episodes together to produce rhetorical flow. For example, the sentence "Yet progress doesn't always yield satisfaction" can be used to connect several episodes describing the positive effects of technological progress and several episodes describing social or environmental problems arising from technological progress. Associated with the English sentence is a formal representation constraining the meanings that episodes before and after the rhetorical device can have. Finally, Terminal Time has a media database of video clips, still images, and sounds. Each of these media elements is represented in a searchable index. Once a narrative track has been generated, Terminal Time uses the index to select media elements consistent with the narrative track.

Terminal Time is a collaboration with interactive media artist Paul Vanouse and documentary filmmaker Steffi Domike.

Intelligent Entertainment

The three projects described above provide an example of alternatives to the shoot-em-up or puzzle-solving adventure that dominate current computer gaming. All three make use of AI techniques to provide interactive experiences that do more than shallowly entertain; they also provide opportunities for introspection and exploration. AI-based interactive experiences can break free of the strictures of the current computer gaming paradigm and become a part of High Culture.

Subjective Avatars is an example of work in Interactive Drama. Here, the goal is to create a story-like experience in which the focus is on interactions with characters, not on solving puzzles. The Subjective Avatar offers an opportunity for a user to experience the world from a new viewpoint. The hope is to combine the empathic

understanding of a character achieved by books and movies with the intensity of a first-person interaction.

Office Plant #1 is a cross between a companion agent and an art object that someone would keep in their office. It shares a focus on long-term engagement with virtual pets such as Dogz and Catz (Stern, Frank, and Resner, 1998). However, virtual pets are intended for circumscribed, high-intensity interaction. The user interacts with the pets for specific periods of time during which this interaction is the primary activity. In contrast, OP#1 is always on, providing a background ambient commentary on the day's activity.

Terminal Time provides a mass audience an opportunity to reflect on historical construction. The audiences to which we've shown Terminal Time (in prototype form), have been engaged by the question periods, applauding to the various questions and laughing and commenting on their own emergent crowd behavior. The questions and comments during the discussion period following the performance indicate that the audience is asking the questions about ideology and historical construction that we hoped to raise. Terminal Time succeeds in being entertaining while functioning as a critical art work.

A New Direction for AI

AI has traditionally been used to study the possibilities and limitations inherent in the physical realization of intelligence (Agre, 1997). The focus has been on understanding AI systems as independent entities, studying the patterns of computation and interactions with the world that the system exhibits in response to being given specific problems to solve or tasks to perform. In AI-based art and entertainment, however, the focus turns to *authorship*. The AI system becomes an artifact built by authors in order to communicate a constellation of ideas and experiences to an audience. This focus on authorship opens a new direction for AI.

A new conception of AI that makes authorship central can be understood in relationship to a schematic map that divides AI into "traditional"(sometimes called GOFAI, or Good Old Fashioned AI) and "behavioral" (sometimes called interactionist) AI. Though crude, this map is useful as a tool for comparison.

Traditional AI is characterized by its concern with symbolic manipulation and problem solving (Brooks, 1991). A firm distinction is drawn between mental processes happening "inside" the mind and activities in the world happening "outside" the mind (Agre, 1997). Traditional AI's research program is concerned with developing the theories and engineering practices necessary to build minds that exhibit intelligence. Such systems are commonly built by expressing domain knowledge in symbolic structures and specifying rules and processes that manipulate these structures. Intelligence is considered to be a property that inheres in the symbolic manipulation happening "inside" the mind. This intelligence is exhibited by demonstrating the program's ability to solve problems.

Where traditional AI concerns itself with mental functions such as planning and problem solving, behavioral AI is concerned with embodied agents interacting in a world (physical or virtual) (Brooks, 1991 and Agre, 1997). Rather than solving complex symbolic problems, such agents are engaged in a moment-by-moment dynamic pattern of interaction with the world. Often there is no explicit representation of the "knowledge" needed to engage in these interactions. Rather, the interactions emerge from the dynamic regularities of the world and the reactive processes of the agent. As opposed to traditional AI, which focuses on internal mental processing, behavioral AI assumes that having a body which is embedded in a concrete situation is essential for intelligence. It is the body that defines many of the interaction patterns between the agent and its environment.

Traditional AI can be characterized as building brains in vats - disembodied minds solving complex symbolic problems. Behavioral AI can be characterized as building emergent insects - embodied agents engaged in relatively simple patterns of sensory-motor interaction with their environments. Historically, behavioral AI appeared as a reaction to recurring problems appearing in traditional AI, particularly in the design of robots (Brooks, 1990, 1991). And certainly behavioral AI has been successful in opening up new design spaces, particularly in inverting the hierarchical opposition between the center of mind and periphery of environment (Agre 1997). However, both traditional and behavioral AI reify the notion of intelligence. That is, intelligence is viewed as an independently existing entity with certain essential properties. Traditional AI assumes that intelligence is a property of symbolic manipulation systems. Behavioral AI assumes that intelligence is a property of embodied interaction with a world. Both are concerned with building something that **is** intelligent; that unambiguously exhibits the essential properties of intelligence.

The three systems described above are informed by a different conception of AI: expressive AI. If traditional AI builds brains in vats, and behavioral AI builds embodied insects, then expressive AI builds *cultural artifacts*. The concern is not with building something that *is* intelligent independent of any observer and their cultural context. Rather, the concern is with building an artifact that *seems* intelligent, that participates in a specific cultural context in a manner that is perceived as intelligent. Expressive AI views a system as a performance. Within a performative space the system expresses the author's ideas. The system is both a messenger for and a message from the author. Expressive AI thus changes the focus from the system as a thing in itself (presumably demonstrating some essential feature of intelligence), to the communication between author and audience. At the technical level of building the artifact, the technical practice becomes one of exploring which architectures and techniques best serve as an inscription device within which the authors can express their message. For example, Sengers developed a new agent architecture for believable agents by thinking

explicitly about the relationship between author and audience (Sengers, 1998).

Expressive AI is not a technical research program calling for the overthrow of traditional or behavioral AI. Nor does it single out a particular technical tradition as being peculiarly suited for expression. For example, subjective avatars draw from behavioral AI, Office Plant #1 draws from statistical AI, and Terminal Time draws primarily from traditional AI. Rather, expressive AI is a stance or viewpoint from which AI techniques can be rethought and transformed. New avenues for exploration are opened up; research values are changed.

Conclusion

AI-based art and entertainment opens new possibilities both for game design and for AI. For games, it points the way to intelligent entertainment that functions as High Culture. For AI, it points the way to expressive AI, a new viewpoint that can inform and direct AI research.

References

Agre, P. 1997. *Computation and Human Experience.* Cambridge, UK: Cambridge University Press.

Bates, J. 1992. Virtual Reality, Art, and Entertainment. *Presence: The Journal of Teleoperators and Virtual Environments* 1(1): 133-138.

Boehlen, M., and Mateas, M. 1998. Office Plant #1: Intimate space and contemplative entertainment. *Leonardo*, Volume 31 Number 5: 345-348.

Brooks, R. 1991. Intelligence Without Reason, A.I. Memo 1293. Artificial Intelligence Lab. MIT.

Brooks, R. 1990. Elephants Don't Play Chess. *Robotics and Autonomous Systems* 6: 3-15.

Carbonell, J. 1979. Subjective understanding: Computer models of belief systems. Ph.D. diss., Computer Science Department, Yale University.

Domike, S.; Mateas, M.; and Vanouse, P. 1998. *The recombinant history apparatus presents: Terminal Time*. Forthcoming.

Kosko, B. 1997. *Fuzzy Engineering.* New York: Simon & Schuster, pp. 499–525.

Loyall, A. B.; and Bates, J. 1991. Hap: A Reactive, Adaptive Architecture for Agents. Technical Report CMU-CS-91-147. Department of Computer Science. Carnegie Mellon University.

Mateas, M. 1997. Computational Subjectivity in Virtual World Avatars. *Working notes of the Socially Intelligent Agents Symposium, AAAI Fall Symposium Series.* Menlo Park: Calif.: AAAI Press.

Mitchell, T. 1997. *Machine Learning*. New York: McGraw-Hill, p. 180.

Neal Reilly, W. S. 1996. Believable Social and Emotional Agents. Ph.D. diss., School of Computer Science, Carnegie Mellon University.

Sengers, P. 1998. Anti-Boxology: Agent Design in Cultural Context. Ph.D. diss., School of Computer Science, Carnegie Mellon University.

Stern, A.; Frank, A.; and Resner, B. 1998. Virtual Petz: A hybrid approach to creating autonomous, lifelike Dogz and Catz. In Proceedings of the Second International Conference on Autonomous Agents, 334-335. Menlo Park, Calif.: AAAI Press.