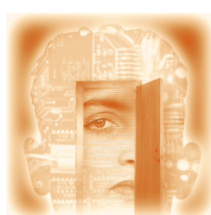
Artificial Life Meets Entertainment: Lifelike Autonomous Agents

Pattie Maes

The potential of AI in the field of entertainment is slowly but steadily being unearthed in a series of new, creative applications that reflect the even grander future of artificial life.



The relatively new field of *artificial life* attempts to study and understand biological life by synthesizing artificial life forms. To paraphrase Chris Langton, the founder of the field, the goal of artificial life is to "model life as it could be so as to understand life as we know it." Artificial life is a very broad discipline which spans such diverse topics as artificial evolution, artificial ecosystems, artificial morphogenesis, molecular evolution, and many more. Langton offers a nice overview of the different research questions studied by the discipline [6]. Artificial life shares with artificial intelligence (AI) its interest in synthesizing adaptive autonomous agents. Autonomous agents are computational systems that inhabit some complex, dynamic environment, sense and act autonomously in this environment, and by doing so realize a set of goals or tasks for which they are designed.

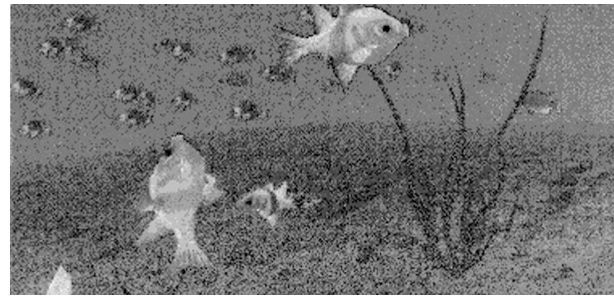
The goal of building an autonomous agent is as old as the field of AI itself. The artificial life community has initiated a radically different approach to this goal, which focuses on fast, reactive behavior, rather than on knowledge and reasoning, as well as on adaptation and learning. Its approach is largely inspired by biology, and more specifically the field of ethology, which attempts to understand the mechanisms animals use to demonstrate adaptive and successful behavior.

Autonomous agents can take many different forms, depending on the nature of the environment they inhabit. If the environment is the real physical environment, then the agent takes the form of an autonomous robot. Alternatively, one can build 2D or 3D animated agents that inhabit simulated physical environments. Finally, so-called knowbots, software agents or interface agents are disembodied entities that inhabit the digital world of computers and computer networks [8]. There are obvious applications for all these types of agents. For example, autonomous robots have been built for surveillance, exploration, and other tasks in environments that are unaccessible or dangerous for human beings. There is a long tradition of building simulated agents for training purposes. Finally, more recently, software agents have been proposed as one mechanism to help computer users deal with work and information overload [8].

One potential application area of agent research that has received surprisingly little interest so far is entertainment. This area may become much more important in the coming years, since the traditional main funding source of agent research, the defense industry, has been scaling down. Entertainment is an extremely large industry that is only expected to grow in the near future. Many forms of entertainment feature characters that act in some environment. This is the case for videogames, simulation rides, movies, animation, animatronics, theater, puppetry, certain toys, and even party lines. Each of these entertainment forms could potentially benefit from the casting of autonomous semi-intelligent agents as entertaining characters. Entertainment is a fun and very challenging application area that will push the limits of agent research.

The Challenge of Modeling Entertaining Characters

Several forms of commercial entertainment currently incorporate automated entertaining characters. Most of these characters are extremely simple; they demonstrate very predictable behavior and do not seem very convincing. This is particularly the case for characters with whom a person can interact in real-time, for example, videogame characters. When automated characters show sophisticated behavior, it is typically completely mechanical and noninteractive and the result of a painstaking and laborious process. An example of the latter is the behavior of the dinosaurs in the movie *Jurassic Park*.



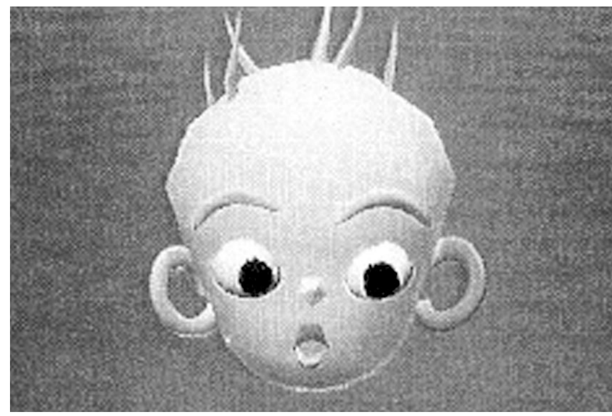


Figure 1.
Realistic fish behavior modeled by Terzopoulos et. al. to produce short animated movies

Figure 2.
Tosa's artificial "baby" reacts to sounds made by a user looking into its crib

Space-Ace says, "So why won't you kiss me, Julia?"
Julia says, "Sorry, I'm not into geeks, Space-Ace."
Space Age says, "Julia, I'm not a geek."
Julia says, "Then lose the pocket protector, Space-Ace."
Space-Ace says, "Julia, I don't have a pocket protector."
Julia says, "Congratulations."
Space-Ace says, "So kiss me, Julia."
Julia says, "I'd rather be overhauled by a Yugo mechanic, Space-Ace."
Space-Ace says, "julia, kiss me"
Julia says, "Why, did your sheep leave you?"
Space-Ace says, "julia, kiss me"
Julia says, "Take a long walk through an unlinked exit, Space-Ace." (Copyright © 1990 by Ashne)

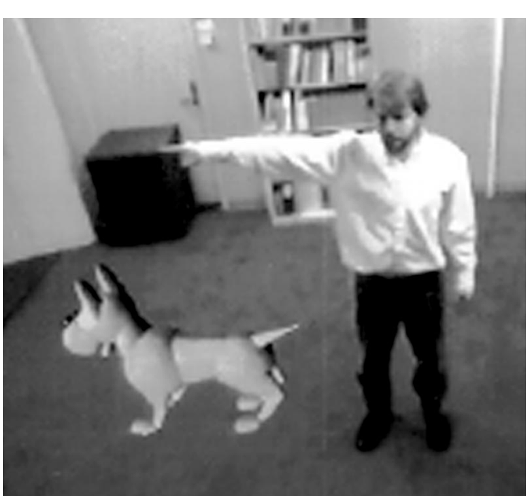
Julia is offended by Lenny's suggestion.
Julia winks at Lenny.
Julia eyes Lenny carefully...
Julia grins at Lenny.
Julia nods to Lenny.
[...]

Figure 3.
Excerpts from interactions with "Julia," Mauldin's autonomous conversing agent that inhabits a text-based multi-user environment

Figure 4.
The ALIVE system allows a user to employ natural gestures to interact with a virtual world inhabited by animated autonomous agents, such as this dog

Figure 5.
Gestures are interpreted by the agents based on the context; here, the dog walks away in the direction the user is pointing

In the last couple of years, a few exceptions have emerged. A number of researchers have applied agent technology to produce animated movies. Rather than scripting the exact movements of an animated character, the characters are modeled as agents that perform actions in response to their perceived environment. Reynolds [11] modeled flocks of birds and schools of fish by specifying the behavior of the individual animals that made up the group. The same algorithms were used to generate some of the behavior of the bats in the movie *Batman II.* Terzopoulos and colleagues [12] modeled very realistic fish behavior, including mating, feeding, learning, and predation (Figure 1). His models have been employed to make entertaining, short, animated movies.



In addition to the previous work, some researchers have used agent models to build interactive, real-time animation systems. Bates' *Woggles World* [1] allows the user to interact with a world of creatures called Woggles. In this pioneering work, a user interacts with the world and its creatures using the mouse and keyboard to directly control the movements and behavior of one of the Woggles. The different Woggles have internal needs and a wide range of emotions, which results in fairly complex interactions. Fisher's *Menagerie* system [4] allows a user to interact with animated agents in real time by using a head-mounted display and a head-tracking device. The agents in *Menagerie* are typically engaged in a single high-level behavior, such as flocking. Certain parameters of this behavior, such as direction of movement, are influenced by the user's actions. Tosa [13] used neural networks to model an artificial baby that reacts in emotional-appearing ways to the sounds made by a user looking into its crib (Figure 2). Finally, the ALIVE system (described later) allows a user to enter a virtual world and use his/her whole body to interact with animated autonomous agents (Figure 4).

In addition to computer animated agents, one notable original project is Julia [10], an autonomous agent that lives in a text-based Multi-User Simulation Environment (MUSE) system. Julia's behavioral repertoire includes moving around in and mapping the environment, holding discourse with players, overhearing conversations and gossiping about them later on, relaying messages between players, and helping players with navigation problems. Julia has moods, feelings, and attitudes toward players and a good memory. She remembers what people told her, and did to her when she last encountered them. Figure 3 includes some sample transcripts of interactions with Julia. This agent is an instance of what Mauldin has termed a "chatterbot." A chatterbot has several different modules for dealing with the different functions required to automate a player in a MUD. The conversation module is implemented as a prioritized layer of mini-experts, which are collections of patterns and associated potential responses. Chatterbots go beyond Weizenbaum's Eliza system in that they use more tricks and have more sophisticated memories of such incidents as past events or conversations.

Building these entertaining agents requires addressing the same basic research questions that are central to all agent research, namely perception, action selection, motor control, adaptation, and communication. The agent has to perceive its environment, which is often dynamic and unpredictable, especially when a user is able to affect it. It has to decide what to do next to make progress toward the goals it is designed to achieve. Relevant actions have to be translated into concrete motor commands. Over time, the agent has to change and improve its behavior on the basis of its past experience. Finally, the agent has to be able to communicate to other agents in its world, both human and artificial ones. The key problem is to come up with an architecture that integrates all these functionalities and results in behavior that is fast, reactive, adaptive, robust, autonomous, and, last but not least, "lifelike." Lifelike behavior is nonmechanistic, nonpredictable and spontaneous. The architectures of many of the successful entertaining agents listed earlier share a surprising number of features:

- The agents are modeled as distributed, decentralized systems consisting of small competence modules. Each competence module is an "expert" at achieving a particular small, task-oriented competence. There is no central reasoner, nor any central internal model. The modules interface with one another via extremely simple messages. Each of the competence modules is directly connected to relevant sensors and effectors. As a result, the behavior produced is robust, adaptive to changes, fast, and reactive.
- Complex behavior is the result of interaction dynamics (feedback loops) at three different levels: interactions between the agent and the environment, among the different modules inside the agent, and among multiple agents. For example, a simple Braitenberg creature that "loves" the user can be built by making it move in the direction of the user with a speed proportional to the distance from the user. As an example of complex multi-agent interaction, Reynolds' creatures demonstrate flocking behavior through the use of simple local rules followed by each of the creatures in the flock [11].
- The architecture includes a lot of redundant methods for the same competence. Multiple levels of complexity/sophistication ensure fault tolerance, graceful degradation, and nonmechanistic behavior. For example, Julia has several methods of responding to an utterance addressed to her: She can try to understand the utterance and generate a meaningful reply; if that fails, she can quote someone else on the same topic; or, if that fails, she can start a different conversation topic.

A more comprehensive discussion of the research questions in autonomous agent research and the characteristics of typical architectures for autonomous agents can be found in [7].

Software agents have been proposed as one mechanism **to help computer users deal with work and information overload**

Apart from the more standard research questions, the design of entertaining agents also requires more novel (to a large degree to the AI community and definitely to the artificial life one) questions to be dealt with—for example, how to model emotions, intentions, social behavior, and discourse [1]. Typically, these issues are even more important than making the agent very intelligent, since, to quote Bates, "The actual requirement is to achieve a persistent appearance of awareness, intention, and social interaction" [1]. Even though these topics may turn out to be of central importance to modeling and understanding intelligence, they have hardly been studied in AI to date.

Finally, building entertaining agents requires the agent researcher to think more about the user. The researcher is forced to address the psychology of the user: how the typical user will perceive the virtual characters, what behavior she or he will engage in, what misconceptions and confusing situations may arise, and so on. Other disciplines such as human-computer interaction, animation, sociology, literature, and theater are particularly helpful in answering these questions. For example, animation teaches us that users typically perceive faster-moving characters as being younger, upbeat, and more intelligent. Literature and theater teach us that it is easier for users to quickly grasp stereotyped characters, such as the "shrink" in the Eliza program.

We are investigating how autonomous videogame characters can learn *and improve their competence over time, to keep challenging a videogame player*

The ALIVE Project

A more detailed description of a particular project aimed at building entertaining agents may convey the research challenges and application opportunities of entertaining agents in a more convincing way. Artificial Life Interactive Video Environment (ALIVE) [9] is a virtual environment that allows wireless full-body interaction between a human participant and a virtual world inhabited by animated autonomous agents. One of the goals of the ALIVE project is to demonstrate that virtual environments can offer a more emotional and evocative experience by allowing the participant to interact with animated characters.

The ALIVE system was demonstrated and tested in several public forums. The system is installed permanently at the MIT Media Laboratory in Cambridge, Mass. ALIVE was also demonstrated at the SIGGRAPH 93 Tomorrow's Realities show, AAAI 94's Art Show, and SIGGRAPH 95's Interactive Communities show. It will be featured in the Ars Electronica Museum, currently under construction in Linz, Austria and in the ArcTec electronic arts biennial in Tokyo, Japan.

In the style of Myron Krueger's Videoplace system [5], the ALIVE system offers an unencumbered, full-body interface with a virtual world. The ALIVE user moves around in a space approximately 16 x 16 feet. A video camera captures the user's image, which is composited into a 3D graphical world after it has been isolated from the background. The resulting image is projected onto a large screen, which faces the user and acts as a type of "magic mirror" (Figure 4) in which the user sees him- or herself surrounded by objects and agents. No goggles, gloves, or wires are needed for interaction with the virtual world. Computer vision techniques are used to extract information about the person, such as his or her 3D location and the positions of various body parts, as well as simple gestures performed. ALIVE combines active vision and domain knowledge to achieve robust and real-time performance [9].

The user's location and hand and body gestures affect the behavior of the agents in the virtual world. The user receives visual and auditory feedback about the agents' internal state and reactions. Agents have a set of internal needs and motivations, a set of sensors to perceive their environment, a repertoire of activities they can perform, and a physically based motor system that allows them to move in and act on the environment. A behavior system decides in real time which activity the agents engage in to meet their internal needs and to take advantage of opportunities presented by the current state of the environment [3].

The system allows not only for the obvious direct-manipulation style of interaction but also for a more powerful, indirect style of interaction in which gestures can have more complex meaning. The meaning of a gesture is interpreted by the agents based on the situation the agent and user find themselves in. For example, when the user points away (Figure 5) and thereby sends a character away, that character will go to a different place in the environment depending on where the user is standing (and in which direction he or she is pointing). In this manner, a relatively small set of gestures can be employed to mean many different things in many different situations.

The ALIVE system incorporates a tool called "Hamsterdam" [3] for modeling semi-intelligent autonomous agents that can interact with one another and with the user. Hamsterdam produces agents that respond with a relevant activity on every time step, given their internal needs and motivations, past history, and perceived environment, with its attendant opportunities, challenges, and changes. Moreover, the pattern and rhythm of the chosen activities is such that the agents neither dither among multiple activities nor persist too long in a single activity. They are capable of interrupting a given activity if a more pressing need or an unforeseen opportunity arises. The Hamsterdam activity model is based on animal behavior models proposed by ethologists. In particular, several concepts proposed in ethology, such as behavior hierarchies, releasers, and fatigue have proven crucial in guaranteeing the robust and flexible behavior required by autonomous interacting agents [3]. The ALIVE systems shows that animated characters based on artificial life models can look convincing (i.e., allow suspension of disbelief).

When using Hamsterdam to build an agent, the designer specifies the sensors of the agent; its motivations, or internal needs; and its activities and actions. Given that information, the Hamsterdam software automatically infers which of the activities is most relevant to the agent at a particular moment, according to the state of the agent, the situation it finds itself in, and its recent behavior history. The observed behaviors or actions of the agent are the final result of numerous activities competing for control of the agent. The activities compete on the basis of the value of a given activity to the agent at that instant, given the perceived environment, the internal needs of the agent, and the agent's recent history. The details of the behavior model and a discussion of its features are reported upon in [3].

The ALIVE system consists of different virtual worlds among which the user can switch by pressing a virtual button. Each world is inhabited by different agents: One is inhabited by a puppet, a second one by a hamster and a predator, and a third one by a dog.

The puppet follows the user around (in 3D) and tries to hold the user's hand. It also imitates some of the actions of the user (e.g., sitting down, jumping, among others.). It goes away when the user points away and comes back when the user waves. The puppet employs facial expressions to convey its internal state. For example, it pouts when the user sends it away and smiles when the user motions it to come back. It giggles when the user touches its belly.

The hamster avoids objects, follows the user around and begs for food. The hamster rolls over to have its stomach scratched if the user bends over and pats it. If the user has been patting the hamster for a while, its need for attention is fulfilled, and some other activity takes precedence (e.g., looking for food). The user is able to feed the hamster by picking up food from a virtual table and putting it on the floor. The user is also able to let the predator out of its cage and into the hamster's world. The predator tries to chase and kill the hamster. The predator views the user as a predator and attempts to avoid and flee from the user. However, as it gets more hungry, it will become more bold and dare to come closer to the user. Both the predator and the hamster are successful at reconciling their multiple internal needs (avoiding the predator, finding food, not running into obstacles).

The most sophisticated character built so far is a dog called Silas. Silas' behavioral repertoire currently includes navigating and exploring following the user and enticing the user to play, sitting (when asked by the user), going away (when sent away by the user) and performing other tricks, such as jumping, fetching a ball, lying down, and shaking. It also will chase the hamster if that agent is introduced into the dog world. Along with visual sensors and feedback, the dog world also uses simple sound input and output. In addition to the camera, an array of microphones faces the user. The resulting signal is fed into a simple pitch tracker, and high pitch (e.g., clapping, whistling, high voice) and low pitch (low voice) are interpreted respectively as positive and negative input from the user to the dog. The dog also provides auditory output, which consists of a variety of prerecorded samples.

By observing thousands of users interacting with the agents in ALIVE, several things have been learned. First of all, the gestures in which the user can engage should be intuitive with respect to the domain and should provide immediate feedback. To user interface designers the latter will seem obvious, but to artificial life and AI researchers it isn't. Examples of *natural* gestures are petting for creatures or pointing and waving for the virtual puppet. Whenever a gesture is successfully perceived by an agent, the user should receive immediate feedback, in terms of either movement or facial and body expression (the hamster rolls over when patted, the puppet smiles when tickled). This helps the user develop an understanding of the space of recognized gestures.

Second, even if the gestures are natural to the environment, it is still necessary to have a guide present to give the user hints about what he or she can do (e.g., "Try petting the hamster," "The puppet will go away if you point."). The current ALIVE system includes an artificial guide, implemented as yet another autonomous agent, which fulfills a special role: It observes the user and memorizes what kinds of interactions the user has had with the world and occasionally offers suggestions by speaking, as well as by using gestures. The guide is visualized as a parrot in the virtual world, because people expect a parrot to talk but not to be able to understand speech.

Third, users are more tolerant of the imperfections in an agent's behavior (as opposed to that of an object), such as lags and occasional incorrect or missed recognition. Having agents causes people to have appropriate expectations about the performance of the sensor system. We learned that people expect virtual inanimate objects to work reliably—the reaction of the object has to be immediate, predictable, and consistent. On the other hand, people assume that animal- or humanlike agents have perception and state; and thus, they are able to accept that the agent may not have sensed something. As a result, gestures that are hard to recognize, such as waving, can be used successfully in the context of agents (an agent might not have "seen" the user waving), but the same gesture would cause the user frustration if used in the context of some inanimate object, for example, a switch.

Fourth, it is important to visualize the motivational and emotional state of an agent in the external features of the agent. For example, a more sophisticated creature, such as Silas the dog, will lead with its eyes; that is, it turns to look at an object or a person before it actually walks over to the object or the person (e.g., to pick the object up or to invite the person to play). If a character does not lead with its eyes, its behavior looks very mechanical and as such not very lifelike. Another reason why it is necessary to show motivations and internal state is that the user may get confused or frustrated if he or she cannot perceive internal variables that determine the behavior of the agent. For example, if Silas is hungry, it may not be very obedient. It is important that users be able to observe that Silas is hungry, so that they realize this may be why it behaves very differently from the way it did a few minutes ago.

Finally, the most important lesson is that for an immersive environment to be captivating, how fancy the graphics are may be less important than how meaningful the interactions in which the user engages can be. ALIVE users reported having a lot of fun using the system and interacting with the creatures. In particular, they seemed to enjoy worlds inhabited by *emotional* agents, characters with which the user can have more of an emotional relationship. For example, users were very much intrigued by the facial expressions of the puppet and would feel bad when their actions caused the puppet to pout or happy when they caused it to smile.

The ALIVE system demonstrates that entertainment can be a challenging and interesting application area for autonomous agents research. ALIVE provides a novel environment for studying architectures for intelligent autonomous agents. As a testbed for agent architectures, it avoids the problems associated with real hardware agents or robots but at the same time forces us to face nontrivial problems, such as dealing with noisy sensors and an unpredictable, fast-changing environment. It makes it possible to study agents with higher levels of cognition, without oversimplifying the world in which these agents live.

ALIVE represents only the beginning of a whole range of novel applications that could be explored with this kind of system. We are currently investigating ALIVE for interactive story-telling applications, in which the user plays one of the characters in the story and all the other characters are artificial agents, which collaborate to make the story move forward (see three short papers on ALIVE in [2]). Another obvious entertainment application of ALIVE is videogames. We have hooked up the ALIVE vision-based interface to existing videogame software, so as to let the user control a game with his or her whole body. In addition, we are investigating how autonomous videogame characters can learn and improve their competence over time, to keep challenging a videogame player. Finally, we are modeling animated characters that teach a user a physical skill in a personalized way. The agent is modeled as a personal trainer that demonstrates to the user how to perform an action and provides personalized and timely feedback to the user, on the basis of sensory information about the user's gestures and body positions.

Conclusion

Recently developed systems like as Woggles [1], Neurobaby [13], Terzopoulos' Fish [12], Julia [10] and ALIVE [8] demonstrate that entertainment can be a fun and challenging application area for autonomous agents research. However, at the same time, these early experiments demonstrate that this application area will require more of an interdisciplinary approach that combines the know-how of the human sciences with the computational models developed in artificial life and AI.

References

1. Bates, J. The role of emotion in believable characters. *Commun. ACM, 37,* 7 (July 1994).
2. Bates, J., Hayes-Roth, B., and Maes, P. Workshop notes of the AAAI Spring Symposium on Interactive Story Systems: Plot and Character, AAAI, March 1995.
3. Blumberg, B. Action selection in Hamsterdam: Lessons from ethology. In *Proceedings of the 3rd International Conference on the Simulation of Adaptive Behavior.* MIT Press, Brighton, Mass. 1994.
4. Fisher, S., et al. Menagerie. In *SIGGRAPH-93 Visual Proceedings, Tomorrow's Realities.* ACM SIGGRAPH, New York, 1993, 212–213.
5. Krueger M.W. *Artificial Reality II.* Addison-Wesley, Reading, Mass., 1990.
6. Langton, C., Ed. *J. Artif. Life 1,* 1/2 (Fall 1993/Winter 1994). MIT Press, Cambridge, Mass.
7. Maes, P. Modeling adaptive autonomous agents. *J. Artif. Life 1,* 1/2 (Fall 1993/Winter 1994). MIT Press, Cambridge, Mass.
8. Maes, P. Agents that reduce work and information overload. *Commun. ACM 37,* 7 (July 1994).
9. Maes, P., Darrell, T., Blumberg, B., and Pentland, A. The ALIVE system: Full-body interaction with autonomous agents. In *Proceedings of the Computer Animation '95 Conference* (Geneva, Switzerland, Apr.) IEEE Press, New York, 1995.
10. Mauldin, M. ChatterBots, TinyMuds, and the Turing test: Entering the Loebner prize competition. In *Proceedings of the AAAI 1994 Conference.* MIT Press, Cambridge, Mass., pp. 16–21.
11. Reynolds, C. Flocks, herds and schools: A distributed behavioral model. In *Computer Graphics: Proceedings of SIGGRAPH '87* (July) ACM Press, New York, 1987,
12. Terzopoulos, D., et al. Artificial fishes with autonomous locomotion, perception, behavior and learning, in a physical world. In *Proceedings of the Artificial Life IV Workshop,* P. Maes and R. Brooks, Eds. MIT Press, Cambridge, Mass., 1994.
13. Tosa, N. Neurobaby. In *SIGGRAPH-93 Visual Proceedings, Tomorrow's Realities,* ACM SIGGRAPH 1993, 212–213.

About the Author:

PATTIE MAES is an associate professor at the MIT Media Lab where she holds the Sony Corporation Career Development chair. Her current research interests include artificial intelligence, artificial life, and human-computer interaction. She is particularly interested in modeling and building autonomous agents. **Author's Present Address:** MIT Media Lab, Rm. E15-305, 20 Ames Street, Cambridge, MA 02139, email: pattie@media.mit.edu, http://pattie.www.media.mit.edu/people/pattie