The Role of Emotion in Believable Agents

Joseph Bates

April 1994
CMU-CS-94-136

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

To appear in Communications of the ACM, Special Issue on Agents, July 1994.

Abstract

Artificial intelligence researchers attempting to create engaging, apparently living creatures may find important insight in the work of artists who have explored the idea of believable character. In particular, appropriately timed and clearly expressed emotion is a central requirement for believable characters. We discuss these ideas and suggest how they may apply to believable interactive characters, which we call "believable agents."

This work was supported in part by Fujitsu Laboratories and Mitsubishi Electric Research Laboratories. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of any other parties.

Keywords: artificial intelligence, emotion, believable agents, art, animation, believable characters.

1 Believability

There is a notion in the Arts of "believable character." It does not mean an honest or reliable character, but one that provides the illusion of life, and thus permits the audience's suspension of disbelief.

The idea of believability has long been studied and explored in literature, theater, film, radio drama, and other media. Traditional character animators are among those artists who have sought to create believable characters, and the Disney animators of the 1930's made great strides toward this goal. The first page of the enormous classic reference work on Disney animation (Thomas and Johnston, 1981) begins with these words:

> "Disney animation makes audiences really believe in ... characters, whose adventures and misfortunes make people laugh — and even cry. There is a special ingredient in our type of animation that produces drawings that appear to think and make decisions and act of their own volition; it is what creates the illusion of life."

Many artificial intelligence researchers have long wished to build robots, and their cousins called "agents," that seem to think, feel, and live. These are creatures with whom you'd want to share some of your life: as with a companion, or a social pet. For instance, in his 1985 American Association of Artificial Intelligence Presidential Address (Bledsoe, 1986), Woody Bledsoe told of his continuing dream to build a computer friend. He spoke of the "excitement of seeing a machine act like a human being, at least in many ways," of building a machine that could "understand, act autonomously, think, learn, enjoy, hate" and which "liked to walk and play Ping-Pong, especially with me."

Woody Bledsoe is hardly alone. Further reading on the dreams of animators and AI researchers finds both groups speaking of thinking, feeling, living creatures, of creating at least the illusion of life, of building apparently autonomous entities that people, especially the creators, would genuinely care about. Both groups also speak of achieving these dreams by finding the essence of the creatures to be simulated, and reconstructing that essence in the medium of the artist's or scientist's choice.

As AI researchers tried to find these essential qualities of humanity, they gravitated toward reasoning, problem solving, learning via concept formation, and other qualities apparently central to the capacity we call intelligence. Perhaps this happened because these qualities are characteristic of the idealized scientist, and thus are valued by the communities of which the researchers were part.

Artists, in particular the character animators, also tried to understand and express the essence of humanity in their constructions. Character animators had to be especially analytic, because they had to produce human life from nothing more than individual, hand-drawn, flat-shaded line drawings, moved frame by frame, without being able to rely on a human actor to portray the character. The practical requirement of producing hundreds of thousands of these drawings forced animators to use extremely simple, non-realistic imagery, and to seek and abstract precisely that which was crucial.

It can be argued that while scientists may have more effectively recreated scientists, it is the artists who have come closest to understanding and perhaps capturing the essence of humanity that Bledsoe, and other AI researchers, ultimately seek. If this is true, then the results of artistic inquiry, especially the insights into character animation such as those expressed in *The Illusion of Life*(Thomas and Johnston, 1981) and elsewhere (Jones, 1989; Lasseter, 1987), may provide key information for building computational models of believable *interactive* characters. We call these constructions "believable agents." Their development is one of my group's research goals in the Oz project(Bates, 1992).

2 Emotion in Believable Characters

The qualities the animators judged important are not identical to those that AI researchers emphasized, though there is some overlap. One quality the animators felt was of the greatest significance in characters was appropriately timed and clearly expressed emotion.

Quoting again from Thomas and Johnston, two of Disney's Nine Old Men (Thomas and Johnston, 1981):

> "From the earliest days, it has been the portrayal of emotions that has given the Disney characters the illusion of life."

The apparent desires of a character, and the way the character feels about what happens in the world with respect to those desires, are what make us care about that character. If the character does not react emotionally to events, if they don't care, then neither will we. The emotionless character is lifeless, as a machine.

According to Thomas and Johnston, properly portraying the emotional reactions of a character requires the animator to remember several key points:

1. *The emotional state of the character must be clearly defined.* The animator needs to know this state at each moment, so that the viewer will be able to attribute definite emotional status to the character as well.

2. *The thought process reveals the feeling.* Since Disney technique claims that action shows the thought processes, viewers must see the emotions in what the character does as its thinking is influenced by its clearly defined emotional state.

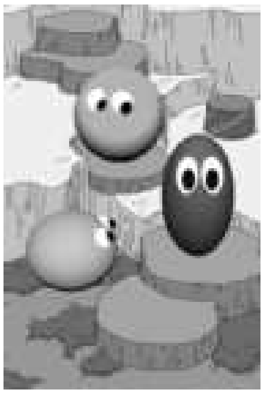

Figure 1: The Woggles in *Edge of Intention*

3. *Accentuate the emotion. Use time wisely to establish the emotion, to convey it to viewers, and to let them savor the situation.* People often cannot grasp emotional state (and other situations) immediately. Thus, believable characters depend on appropriate use of a variety of mechanisms to convey emotion, such as fore-shadowing the emotion, exaggerating it, and toning down other simultaneous action to get the main point across.

Of course, there is far more to be said about emotion in believable characters, but let us see where these comments might lead us in thinking about computational models for believable agents. In particular, we'll look briefly at how the Oz group attempted to build on animation criteria, such as the significance of emotion, and at some of the results of trying to construct believable agents in this way.

3 Using Emotion to Construct Believable Agents

One of the efforts of the Oz project in 1992 was to build a small simulated world containing several realtime, interactive, self-animating creatures based on the principles of traditional character animation. We constructed an art work, called "Edge of Intention", that was first shown at the AAAI-92 AI-based Arts Exhibition and then subsequently at SIGGRAPH-93, at Ars Electronica '93, and in the Boston Computer Museum. Figure 1 shows an image of three of these self-animating creatures, called "Woggles," posing for the camera. The three particular creatures are named Shrimp, Bear, and Wolf.

The Disney suggestions on portrayal of emotions do not tell us much about the representations of emotion inside an agent. There may be many computational models consistent with the appearance of a definite emotional state, and with the other suggestions. In particular, there is no requirement that the emotional state be represented in any obvious way within the agent, or within the agent/world system.

In building the Woggles, we chose to develop a goal-directed, behavior-based architecture for action (Loyall and Bates, 1993; Brooks, 1986; Maes, 1989), and to couple it to a distinct component for generating, representing, and expressing emotion based on the work of Ortony, Collins, and Clore (OCC) (Ortony et al., 1988; Bates et al., 1992). The action system does no planning and almost no world modeling, but it does use a minimalist conception of goals to manipulate a dynamically changing set of behaviors. These goals, and the agent's appraisals of events with respect to the goals, are key to producing a clearly defined emotional state in the creature. Our choices to make emotions explicit and to allow composition of behaviors through goals were intended to make it easy to build characters with specific personalities, while maintaining the robustness of robots that use minimal representation.

Building on the OCC theory of emotion gives us creatures with definite emotional reactions to events. A simple example is a Woggle creating an analog of anger when it both experiences an important goal failure and judges that the failure was caused by another Woggle. It often occurs, however, that many emotions of various intensities exist simultaneously. Part of our work has been in finding ways to combine these to get one or two primary overall emotions of adequate clarity to satisfy maxim 1. We are not sure if our solutions are adequate (Reilly and Bates, 1992), and we are continuing to explore this problem.

According to the second maxim, a creature's emotional state must change its thought processes so that the emotional state is visible in its actions. In the Woggles, each emotion was mapped in a personality specific way to a behavioral feature. For instance, the "fear" emotion mapped to the "alarmed" feature in Shrimp, but to the "aggressive" feature in Wolf. These features had both local and systemic effects on the action generating components of the architecture. We attempted to build behaviors so that any combination of features would produce action that conveyed those features.

Behavioral features were intended as an abstraction mechanism, to let us build a library of modular behaviors that later could be used in any character. However, as we constructed individual Woggles from these modules, we found that we wanted emotion to affect action in very ad hoc ways that seemed to depend on breaking down the abstraction barriers frequently. There is debate within the Oz group about whether this is an essential quality of characters, or whether it will be possible to find durable abstractions in the mapping from emotion to action.

An argument against abstraction comes from our experience with Shrimp. Due to a programming error, Shrimp occasionally peppers his ongoing behavior with what seems to be a nervous tick causing him repeatedly to hit his head against the ground. This attracts people's attention immediately, but to our surprise they build psychological theories, always incorrect, about Shrimp's mental state and seem to find him much more interesting and alive for having this behavior. Their reaction appears consistent with what Chuck Jones, famous for his success animating Bugs Bunny and other Warner Brothers characters, says was his key insight into animation. As an eight year old he had a cat named Johnson who liked to eat grapefruit and then wear the eaten rind on his head. Jones says he discovered that it is the oddity, the quirk, that gives personality to a character, and it is personality that gives life. If so, our architectures must support quirks, and this may mean they need to allow regularities, such as are expressed in abstraction barriers, to be broken.

Finally, let us consider the third maxim. We took care to design an architecture that provided Woggles with strong internal emotional states. However, we did not use the animation techniques of exaggeration, anticipation, or staging nearly to the extent necessary to make those emotions quickly recognizable.

An example of insufficient exaggeration was the moping behavior. A moping Woggle directly expresses sadness by decreasing its muscle tone, decreasing the length of its jumps, and slowing the timing of its actions. The creator of the behavior reported that he "started with what I thought would be a reasonable looking behavior, but it wasn't clear enough so I tripled and then doubled again [the relevant parameters] and still only got a mope that people recognize when it is pointed out to them."

Anticipation denotes broadcasting in advance that a certain action is about to occur, to let people prepare to see it. This is necessary because realistic timing causes events to happen quickly, and people will miss them unless they are experienced watchers. An analogy is that people watching an unfamiliar sport often seem unable to see what is happening until they learn to predict likely near future events. A classic example of anticipation in animation occurs when a character rears back in preparation for a rapid forward motion. Strictly speaking, this is unrealistic, but it is essential for people to grasp the meaning in motion.

Staging is used in theater and animation to refer to controlling the action of several characters in time and space to make sure the main point of a situation is conveyed. At each moment in "Edge of Intention," several Woggles are often moving, jumping, and gesturing socially. This makes it very difficult for people interacting with the creatures to stay aware of what is happening. While this simultaneous action may be realistic, unless the goal is specifically to produce confusion in the user, it is not helpful in producing the subjective feeling of realism.

4 Conclusion

We've talked about the idea of believable agents, by which we mean an interactive analog of the believable characters discussed in the Arts. We've tried to argue that these artists hold some of the same central goals as AI researchers, and that their insights could be important as we try to build engaging, apparently living creatures. Such creatures, besides sparking the interest of AI researchers and artists, may serve as a component of new user interfaces for the broad human population.

Chuck Jones, describing animation at Warner Brothers (Jones, 1989), said:

> "Believability. That is what we were striving for ... belief in the life of the characters. That, after all, is the dictionary definition and meaning of the word 'animation': to invoke life."

Emotion is one of the primary means to achieve this believability, this illusion of life, because it helps us know that characters really care about what happens in the world, that they truly have desires.

In contrast, the fighting characters of current video games show almost no reaction to the tremendous violence that engulfs them. Such showing of reaction is a key to making characters believable and engaging, which is perhaps why these characters engender no concern — they are merely abstract symbols of action.

As we've mentioned and implied several times, the "illusion of life" refers to the conveyance of a strong subjective sense of realism. One might wish to distinguish this from a notion of genuine life, which would refer to entities that "really are" alive, as opposed to merely seeming so. Just what to make of such a distinction is, of course, the subject of deep debate. Turing's test, Dennett's "intensional stance" (Dennett, 1987), and many other ideas of substance try to get at this issue.

We would note here that whatever one makes of the fundamental philosophical concern, there remains a practical concern of constructing agents of increasing degrees of believability in the short term. Indeed, sufficient success in this regard may alter or eliminate the philosophical debate, as has happened before where practice meets theory. In any case, it is our view that believability will not arise from copying reality. Artists use reality in the service of realism, for example by carefully studying nature, but they do not elevate it above their fundamental goal. Mimicking reality is a method, to be used when but *only* when appropriate, to convey a strong subjective sense of realism. It is art which must lead the charge here.

Believability places a variety of demands on an interactive character. These include the appearance of reactivity, goals, emotions, and situated social competence, among others. The emphasis in "alternative AI" on reactivity could be seen as choosing one of the believability requirements and elevating it to a position of importance, while downgrading other qualities, such as those related to our idealized view of intelligence. We offer the thought that researchers could make this same methodological move with other requirements, such as emotion or personality, that artists tell us are needed to present the convincing, persistent, illusion of life.

5 Acknowledgments

We gratefully acknowledge research support for the Oz project provided by Fujitsu Laboratories and Mitsubishi Electric Research Laboratories.

References

Bates, J. (1992). Virtual reality, art, and entertainment. *PRESENCE: Teleoperators and Virtual Environments*, 1(1):133–138.

Bates, J., Loyall, A. B., and Reilly, W. S. (1992). An architecture for action, emotion, and social behavior. In *Proceedings of the Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World*, S.Martino al Cimino, Italy.

Bledsoe, W. (1986). I had a dream: AAAI presidential address. *AI Magazine*, 7(1):57–61.

Brooks, R. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2:14–23.

Dennett, D. (1987). *The Intentional Stance.* The MIT Press, Cambridge, MA.

Jones, C. (1989). *Chuck Amuck: the Life and Times of an Animated Cartoonist.* Farrar, Straus & Giroux, New York.

Lasseter, J. (1987). Principles of traditional animation applied to 3D computer animation. In *Proceedings of SIGGRAPH '87*, pages 35–44. Association for Computing Machinery.

Loyall, A. B. and Bates, J. (1993). Real-time control of animated broad agents. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, Boulder, CO.

Maes, P. (1989). How to do the right thing. *Connection Science*, 1(3):291–323.

Ortony, A., Clore, G., and Collins, A. (1988). *The Cognitive Structure of Emotions.* Cambridge University Press.

Reilly, W. S. and Bates, J. (1992). Building emotional agents. Technical Report CMU-CS-92-143, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Thomas, F. and Johnston, O. (1981). *Disney Animation: The Illusion of Life.* Abbeville Press, New York.