Anthropomorphic Systems: An Approach for Categorization

Kathryn Howe

Department of Information Systems
University of Maryland Baltimore County (UMBC)
Baltimore, MD 21250 USA
khowe@umbc.edu

Abstract. Are systems that incorporate anthropomorphic attributes better at interactivity with a user than systems that do not use such attributes? Do these systems allow a user to interact with the system in a natural way; or can the system cause more frustration then aid? It is a fact that many systems nowadays are attempting to make their interfaces more natural to use. Some systems attempt to do so by the advance of various input systems, such as touch screens, screen readers, etc. Other systems attempt to create user interfaces that a user can easily relate to. They can take on various anthropomorphic attributes such as emotion, speech, cognition and learning abilities. These systems vary dramatically in how they incorporate the attributes as well. Some systems use an interface of cartoon characters that allow a user to believe that the character can speak and learn like the user, while still keeping a separation of the virtual and real world by its physical form. Others attempt to effect human attributes so much that it could be difficult to distinguish between the two.

1 Introduction

In this paper, various levels of anthropomorphism are explored and evaluated as to its usability. Within each level, real world applications will be discussed to show both benefits and constraints. The final section will review the findings and attempt to determine whether the use of anthropomorphism is beneficial or a hindrance, and to what level its use should be limited to.

2 Background

Systems that provide anthropomorphic attributes can be classified into four levels. These levels differentiate by the sophistication and similarity to true human to human interaction. Level I discuss systems that show intelligence in learning how to operate a task. These systems demonstrate this ability through some sort of primitive anthropomorphic interface.

Level II holds the systems that have the same level of intelligence but also can reflect personality, behaviors and social etiquette while interacting with the user and other agents. The GUI used in level II agents could use the same type of physical interface as level I or can create a more human like interface.

N. Aykin (Ed.): Internationalization, Design, LNCS 5623, pp. 173–179, 2009.
© Springer-Verlag Berlin Heidelberg 2009

Level III houses the systems that show the same intelligence as level II, but also has the capabilities of showing emotion and behaviors with the use of more sophisticated physical embodiment; such as muscular distinctions in the agent's face and body posture. Level III agents are often very human like in the physical model, allowing for a very believable, human to human interface. Not only can level III agents interact with each other and display body language, but these agents are also capable of reading, reacting to and understanding the same body language given off by the users.

Level IV, the highest level in the categorization, has all of the attributes of the lower three levels, but also has the ability to self replicate and to essentially "live" both inside and outside of a virtual environment.

3 Discussion

As previously discussed, level I anthropomorphic systems are among the least realistic systems. These systems are relatively simple in graphics; often only being displayed as a line drawing or cartoon character. These systems are used often times as assistive agents to help the user by offering suggestions or to take on current tasks. These agents have the ability to think independently and to learn the processes of the user.

Examples of such systems can include email agents [6]. In this particular scenario, the agent was a simple drawing of a human face. This face had a few variations. The first variation was an inquisitive look to show the user that it was attempting to learn what the user was doing, and for what purpose. If the agent was learning about what the user was trying to accomplish, and had a suggestion, a light bulb would appear above the figures' head. Another was a look of concentration on an object. This face allowed the user to see that the agent was currently busy working on a task.

Many benefits come from the use of a level I agent. These agents lessen the workload of the user, allowing that user to become more efficient. Prendinger and Ying Zi found that level I interface can also help users concentrate on specific objects and understand the agent's goals by examining the use of an animated real estate agent named "Kosa Ku". When users looked at a virtual tour of a home that was "shown" by Kosa Ku, users often watched the agent until she pointed to the area of interest; and then, after looking at the pointed area, focused attention once again to Kosa Ku. This allowed for the type of interaction expected in human to human interaction. When using other application types, such as text screens along with voice readers, users had to spend more effort and time reading the text; less time was spent looking at the actual tour of the home [11].

As with any system, there are cons that come with level I anthropomorphic systems. Often times, such agents can be viewed as irritating or redundant by experienced users who do not need suggestions. Other systems that takeover tasks for the users benefit must also have a certain level of trust from the user in order for that user to be comfortable with a task takeover. One of the biggest concerns that should be reviewed, especially with animated agents, is the concern of to what level of "cuteness" should be instilled to the agent. Agents that are too cute can quickly loose credibility by users who could most benefit from them.

4 Classifications

Anthropomorphic agents classified in the second level are those agents that have all the features discussed in level I agents, but also are able to display characteristics of personality, behavior and interaction with other agents. The addition of etiquette is also often incorporated into level II agents. Behavioral and emotional additions are important when attempting to create an agent that is believable to the user. Level II agents often times use the physical shape of a human and can include not only the face, but a body as well to show simple body gestures as a supplemental means of communication. They may also take on the shapes of other animated objects such as cartoon animals. Without the use of some basic emotion though, agents lack the believability of "living", which is an essential part in making an agent believable.

Personality is very important for agents who display behavioral actions. An agent's personality will determine differences and uniqueness in physical actions among agents. An agent with a friendly personality may display behaviors such as smiling and winking; where as agents with unfriendly personalities might display behaviors such as frowning or sulking. These personalities will also determine how agents may interact with one another. An agent with an unfriendly personality will often be rude or angry towards other agents, while friendly personalities might be more amicable and patient.

Beyond the use of behavior and personalities in an agent, social role awareness, or etiquette, is also an integral part in creating a believable, anthropomorphic agent. The awareness of social roles and adherence to such roles allows an agent to not only be believable as human like, but also allows the agent to become a peer to the user or other agents. Social role awareness dictates in what manner the user, or the agent, acts.

A study by Prendinger and Ishizuka demonstrates how the use of social role awareness enhances believability of an agent. In this particular study, multiple agents are able to interact with each other in a coffee house setting. Each agent is programmed to alter their behaviors based on the role of the agent that they are interacting with. For example, a customer interacts with a waiter by asking for a beer. The waiter, who has an unfriendly personality, easily gets irritated by the customer asking for a beer in a coffee shop and speaks rudely to the customer. When that same unfriendly waiter asks the manager for some time off, the friendly manager politely says no. The waiter is irritated by the manager's response, but reacts calmly and accepts the answer because the manager's social role is higher than the waiter's; requiring respect by the waiter [10].

The use of social role awareness is also useful when an agent must interact with a user. Agents that adhere to behavioral & social standards are simpler to communicate with because a user knows what to expect while interacting with the agent. Such social standards help to dictate conversational rules and the use of (or lack of use of) language type, speech speed, volume and exclamation, all based on the environment and context of the conversation.

5 Applications

The discussions on level II agents show many benefits when these agents are used. By far the use of behavior, based on personality, and the restrictions of these behaviors

because of social roles, allows an agent to become much more lifelike then those agents in level I. The use of personalities also allow for agents to become unique among others, also bringing a level of acceptance by the users. There are, of course, downsides to the use of level II agents. Every day applications currently lack the need for agents that truly act in a human like manner. Most applications that use level II agents reside in the entertainment and gaming industry. Aside from the current limit of use, another major setback with level II agents is the higher level of user expectations. As agents become more anthropomorphic, user's expectations for the abilities of agent comprehension and response also rise. The fact is, however, that an agent is still a computer system and is bound to the parameters set by their designers. If an agent is not fast enough to respond, or able to interact naturally, frustration by the user can often become an issue.

Level III agents are very similar to level II agents in that they are able to display behaviors and personality traits that are often constrained by social roles and environmental parameters. Level III, however takes these features and limits them by their physical features. These limitations include the limitation to speech or the ability to manipulate their environment based on the limitations of the "body" that these agents inhabit.

Level III systems can show more control over facial features which allow them to communicate to the user what the agent is doing or thinking. Facial displays in communication can offer non verbal cues to each participant of the interaction. The raising of an eyebrow, or the avoidance of eye contact; even the slight downturn of the lips can allow the user to subconsciously display confusion, understanding, or provide an emphasis on what is said or being understood. These types of features can also allow the system to spend time "thinking"; while offering a user a means to understand that it is thinking, suggesting they be more patient to the potentially perceived slower response rate.

If a computer system can take advantage of human facial displays and body language as an input as well as output, the system can determine on its own if the user is confused and stop itself and ask if the user needs clarification. This allows for a more natural means of communication between user and system. A good example of such an interface is the REA agent developed by Bickmore and Yan. REA is a virtual real estate agent that is able to interact with other agents and objects within her own environment. REA is also able to fully communicate with users, as if they were in her environment as well. REA, and her world, is displayed in life like proportions upon a blank wall. A user comes into the room where REA is located and she can detect, not only when someone has entered, but also where they are in the room. REA then focuses on the user and introduces herself. If REA is speaking, and the user motions to her as if they wanted to speak, REA is able to detect such actions and stop herself, to allow the user to speak. If the user's facial displays show signs of confusion, REA can react by asking if they need any clarification. While a user speaks, REA's facial displays give cues of attentiveness and comprehension. If REA is asked a question that requires "thought", she can give cues such as looking up and away while saying "umm", telling the user that she is thinking about how to answer the question and that they should allow her time to think. This sort of non verbal communication was once only available between humans.

Level III agents are able to avoid potential frustration by the user because they are able to react to action from the user and offer cues when "thinking" or when an agent needs to pause. As agents become more believable as anthropomorphic agents, users trust in them will undoubtedly also rise. Although this rise in trust and reliance on a system is ultimately what designers are striving to achieve, the issue of ethical limitations become more essential. Without ethical and moral regulations, such a system could easily be used to exploit the user and take advantage of them.

Level IV agents achieve, currently, the highest level of anthropomorphism. These systems embody the true meaning of artificial live, virtual realities and can also transcend into our virtual plane. Systems categorized within level IV are agents that own all qualities discussed in previous sections but are also able to evolve, adapt and replicate themselves. Level IV agents truly "live" within their own world; for lack of a better term. Other kinds of level IV agents are removed from their own virtual world and become members of the physical world in the form of robots; and are therefore distinct from their lower level anthropomorphic counterparts.

Agents who live in their own virtual reality are able to not only interact logically with their environment, but can actually sense what is there and react accordingly to the situation as it occurs. Not only can these agents react to other agents or objects within their world, but they can change their surroundings, change the behaviors of other agents and influence other's personalities. Agents can adapt to normal styles and evolve when conflicts occur without the requirement of outside influences. These agents can choose their own actions, no longer bound by the initial parameters set by their designers. These agents, and the world in which they live, is the equivalent to our own reality, including all of the positives and negatives that goes with it.

Bruce Blumburg, of MIT Media Lab, has created a system that reflects a level IV type agent. His system, called Artificial Life Interactive Virtual Environment, or ALIVE, is an entire interactive program in which users and characters within the virtual world can fully interact in a real time environment with each other. Although characters in level IV systems may or may not always embody a human form (where in contrast, level III systems are mostly in some level of human form), these characters are so lifelike in the world that they live, the boundary of human shape is not always necessary; though the use of human form would most certainly add to the believability of their true existence.

The other types of level IV agents, robots, are similar to their virtual reality counterparts. Robots are able to sense their surroundings and adapt themselves to their surroundings. Robots are able to communicate with other robots, computer systems, and human users. Applications often associated with robots are categorized as commercial, fictional and research. The more fictional robots are, the more anthropomorphic they can become. Robots are still in their infancy and require much more work and evolution to truly embody the human physical form and be able to do the things that humans are capable of doing. Even with these anthropomorphic limitations, robots are still categorized as level IV agents because they are the only agents that have been removed from a virtual environment and have transcended to the physical world.

One of the biggest constraints to level IV agents is the theory that an agent can become too realistic and that they can become counter productive. This theory, called the "uncanny valley" was created by Mashiro Mori and it specifies that, as robots increase their anthropomorphic attributes, user's comfort level with them begin to diminish.

Mori provides an example: “if you shake an artificial hand (that you perceive to be real) you may not be able to help jumping with a scream, having received a horrible, cold, and spongy grasp”. The point of this theory is that, at least in terms of computer agents that transcend to the physical world, there may be a cap to which these agents can evolve. Not in terms of whether anthropomorphic advances can happen, but whether or not humans will accept them.

6 Findings and Conclusions

This paper attempts to categorize various anthropomorphic systems into four levels of evolution. Each level adds more human like attributes to the systems housed in them. Level I being the lowest and most primitive in terms of physical and mental anthropomorphism. Level II adding the use of emotion, behavior and etiquette to their systems. Level III incorporates a more refined physical structure allowing for physical constraints to the agents, which help with more natural, non verbal communication; as well as the ability to read such cues from the users. Level IV is the creation of a truly independent, self sufficient virtual world in which agents “live”. Systems that have entered the physical world in the form of robots also are considered level IV agents. Each level has real life applications that show both benefits and downfalls attributed to them.

When it comes to an anthropomorphic system’s usability, the requirements needed for a system should determine to what level of anthropomorphism is needed, if indeed, any is needed at all. Without a doubt, the use of these systems allow for ease of use for a novice user, or those users who might be uncomfortable with the idea of learning new and complex systems. The area of systems training could benefit greatly from the use of more human like interfaces and are already being adapted into such systems today. In the future, more systems will likely adopt the use of anthropomorphism in their designs. As the use of these systems rise, it is likely that the level of anthropomorphism within those systems will also rise.

References

1. Bates, J.: The role of emotion in believable agents. Communications of the ACM 37(7), 122–125 (1994)
2. Boling, E., Sousa, G.A.: Interface design issues in the future of business training. Business Horizons 36, 50–54 (1993)
3. Cassell, J., Bickmore, T., Vilhjalmsson, H., Yan, H.: More Than Just a Pretty Face: Affordances of Embodiment, pp. 52–59. ACM, New York (2000)
4. DiSalvo, C., Gemperle, F., Forlizzi, J., Kiesler, S.: All Robots are Not Created Equal: The Design and Perception of Humanoid Robot Heads. In: ACM DIS 2002, pp. 321–326 (2002)
5. Kirlik, A.: Modeling strategic behavior in human-automation interaction: Why an "aid" can (and should) go unused. Human Factors 35(2), 221–242 (1993)
6. Maes, P.: Agents that reduce work and information overload. Communications of the ACM 37(7), 30–40 (1994)
7. Marcus, A.: The Cult of Cute: The Challenge of User Experience Design. Interactions, 29–34 (November/December 2002)

8. Millier, C.: Human-Computer Etiquette: Managing Expectations with Intentional Agents. Communications of the ACM 47(4), 31–34 (2004)
9. Norman, D.A.: How might people interact with agents. Communications of the ACM 37(7), 68–71 (1994)
10. Prendinger, H., Ishizuka, M.: Social Role Awareness in Animated Agents, pp. 270–277. ACM, New York (2001)
11. Prendinger, H., Ma, C., Yingzi, J.: Understanding the Effect of Life-Like Interface Agents Through User's Eye Movements, pp. 108–115. ACM, New York (2005)
12. Takeuchi, A., Nagao, K.: Communicative Facial Displays as a New Conversational Modality, pp. 187–193. ACM, New York (1993)
13. Terzopoulos, D.: Artificial life for computer graphics. Communications of the ACM 42(8), 32–35 (1999)