Autonomous Communicative Behaviors in Avatars

Hannes Högni Vilhjálmsson
The Media Laboratory
Massachusetts Institute of Technology
20 Ames Street, E15-320R
Cambridge, MA 02139
hannes@media.mit.edu

Most networked virtual communities, such as MUDs (Multi-User Domains), where people meet in a fictitious place to socialize and build worlds, have until recently been text-based. However, fueled partly by the VRML standardization of 3D graphics interchange on the internet, such environments are going graphical, displaying models of colorful locales and the people that inhabit them. When users connect to such a system, they choose a character that will become their graphical representation, termed an avatar, in the world. Once inside, the users can explore the environment, often from a 1st person perspective, by moving their avatar around. The avatars of all other users, currently logged onto the system, can be seen and approached to initiate a conversation.

Although these systems have now become graphically rich, communication is still mostly based on text messages or digitized speech streams sent between users. That is, the graphics are there simply to provide fancy scenery and indicate the presence of a user at a particular location, while the act of communication is still carried out through a single text based channel. Face-to-face conversation in reality, however, does make extensive use of the visual channel for interaction management where many subtle and even involuntary cues are read from stance, gaze and gesture. I believe the modeling and animation of such fundamental behavior is crucial for the credibility of the interaction and should be in place before higher level functions such as emotional expression can be effectively applied.

The problem is that while a user is engaged in composing the content of the conversation, i.e. typing or speaking, it would distract her too much if she had to simultaneously control the animated behavior of her avatar through keyboard commands or mouse movements. And the fact that the user resides in a world that is different from the one that the avatar is in, means that directly mapping the user's body motion onto the avatar is not appropriate in most cases. For instance, if the user glances to the side, she would be staring at the wall in her office, not the persons she is having a conversation with in the virtual world.

Therefore I am exploring how low level visual communicative behaviors can be automated in an avatar, based on the user's (a) chosen personality; (b) communicative intentions (as also manifested in the accompanying text or speech), and (c) the dynamics of the social situation.

I have implemented a program to demonstrate the concept of automated avatars. This program models the conversation between two avatars and allows the user to play with different variables that control the automation of their conversational behavior. In particular, I looked at those behaviors that give cues to an approaching 3rd person, about whether that person is welcome to join the conversation or not. The behaviors include eye gaze and body orientation, based on parameters such as awareness of other avatars and openness to a 3rd party joining the conversation. An example of a backchannel feedback was also implemented in the form of head nods.

I will use this first testbed as a launchpad for a project that focuses on how to represent user's communicative intentions non-verbally and exploit the function of embodiment in the construction of animated avatars.

Topics: Dialogue management, multi-modal interfaces, modeling of personality and emotion.