Collaborative Use of Multimodal Applications by Lifelike Computer Characters and Humans

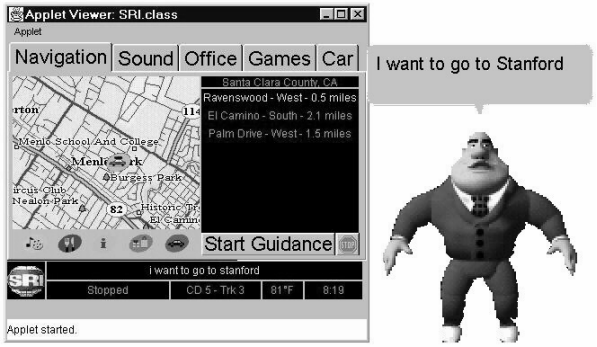
Jérôme DUBREUIL and Luc JULIA
STAR Laboratory – SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025 - USA

In this abstract we present Lifelike Computer Characters (LCC) interacting with Java programs as human users would do. The aim is to allow the LCC to issue any type of command (click, speech, drawing, gestures, etc…) to a target application and for it to interpret the application's feedback as well as its environment in order to adapt its behaviors. DemoMan, presenting a speech demonstration in collaboration with a human, shows an application of these techniques.

Looking at systems with skilled animated characters presenting multimedia data like the WWW-Persona developed at DFKI [1], we decided to go a step further by making the characters able to interact fully with complex multimodal/multimedia applications. Here are some highlights of our work:

- Giving a soul to the computer was necessary for visualization purposes as much as for psychological needs. This ability is provided by a character able to reproduce human gestures and attitudes. It can speak (speech synthesis), listen (speech recognition), interact with the visual GUI (click on buttons, move windows...), draw gestures or write on the screen.
- Most of the interest of our system is its ability to issue multimodal commands, just as if it were a human. The character can talk to the application as well as it can talk to a fellow user, human or not. Similarly, it listens to the application (in case of rejection for instance, so that it can repeat the sentence), and listens to the fellow users for interruptions or specific commands ("Now keep quiet, I want to continue alone"). If the application accepts other multimodal commands such as gesture and handwriting, the character is able to generate drawing or handwriting to the application, which will try to recognize its signals.
- The wrapping interface allowing to add those characters to any Java program, permits multiple characters to interact with the same target application. Appropriate scenarios will make them collaborate (showing different uses of a function in the application...) or act one after the other to go through the application.
- A scenario, or script, is a set commands relative either to the character itself (move, speak, be surprised, click...) or to the target application (issue command). Scripting allows to serialize (wait for an event to be completed) or parallelize (issue a synergistic command) the tasks.
- A fellow user can interrupt spontaneously another character to focus on something or make anything else, and then letting it resume its script.

Implementation and Demonstration



- The whole system is Java-based, using Nuance speech recognizer, SRI International gesture recognizer, CIC handwriting recognizer and a Microsoft Agent server.
- DemoMan is a version of Al, a Microsoft agent character animated by Animatronics.

[1] André Elisabeth, Rist Thomas and Müller Jochen: Integrating Reactive and Scripted Behaviors in a Life-Like Presentation Agent, in: Proc. of the Second International Conference on Autonomous Agents (Agents '98), pp. 261-268.