

Pedagogical Agents: A tool for the design and development of Web-based Instruction

Abstract

Pedagogical agents, when used effectively as a conversational interface, can serve as a powerful tool for instructional designers. These tools are available as web-based plug-ins (i.e. Microsoft Agent, the Haptek Player, & Pulse3D Veepers). This makes it easy for instructional designers to incorporate pedagogical agents into web pages (and even stand-alone applications). Animated Pedagogical agents by providing a variety of gestures which allow developers to create believable characters and provide “the illusion of life” (Bates, 1994). In addition to these animations, most plug-ins include text-to-speech software as an interface for voice simulation (Note Appendix A). Some even include speech recognition for voice command input. But like much of the software on the World Wide Web, most plug-ins are freeware, in other words they are free to instructional designers, developers and learners. This paper hopes to demonstrate, not only how easy it is to use and develop Pedagogical agents, but also why they are important to Instructional Designers or Developers, and finally to discuss the finer points of design, when using these powerful tools in instruction.

Overview

Perhaps the best known software agent is the fictional “character,” the HAL 9000 computer from Arthur C. Clarke’s *2001: A Space Odyssey*. In this 1968 movie, HAL was a sophisticated computer system that could reason with humans and interact with his environment (the spaceship - *Discovery*). Now imagine that HAL was programmed to teach Dave Bowman (the astronaut character on board *Discovery*). HAL would then be a pedagogical software agent.

Even though the HAL 9000 computer is still science fiction, pedagogical software agents are among us already. “Clippit®,” the familiar paperclip animation in Microsoft Word®, is an example of a common pedagogical agent. This Microsoft Office assistant communicates with the user via text. The most recent versions of Microsoft Agent incorporate text-to-speech software to allow pedagogical agents to communicate via a synthetic voice. These interface agents are designed to simplify user interaction via text-to-speech (TTS) software and gestures to create “social interfaces” (Thorisson, 1993).

For many people, the idea of synthetic speech generates visions of Stephen Hawking. Speech synthesis or text-to-speech software has undergone tremendous development since DECtalk (the software used by Stephen Hawking) was originally developed. That was 1983 (Klatt, 1987). Now in the 21st century, it may not be long before many televisions, VCRs, and cell phones have these social conversational interfaces (Müller et al., 2000). AT&T for instance has developed a “conversational natural language speech

understanding” agent called *How May I Help You?*SM, to perform telephone-based customer service (AT&T Labs, 2001). In their words it “is so good at reproducing the sounds, inflections and intonations of a human voice that it can recreate voices and even bring the voices of long-dead celebrities back to life (Guernsey, 2001).” This social conversational agent actually responds to questions and performs customer service through synthetic speech. The plug-ins mentioned above place these powerful technologies in the hands of Instructional designers and developers.

How does one develop a Pedagogical Agent?

The simplest way is to use the agent plug-in provided by Microsoft. Microsoft Agent® can be described as an agent framework. That is Microsoft Agent® is a loose set of software components that work together to produce an agent interface. This interface takes commands from the various client side and backend scripting languages (JavaScript, VBScript, ColdFusion, etc.) In addition, Microsoft allows for a number of speech engines to accompany their animated agents to produce a “voice” for the character. A common speech engine is the freeware version of Lernout & Hauspie TruVoice TTS Engine. This software component allows for a number of voices both male and female, but in addition it can reproduce the inflections and speech of ten different languages including Spanish, French, and German. In the VBscript below, the Merlin agent appears, greets the learner and then disappears. Simply download the necessary software component from Microsoft (<http://agent.microsoft.com/msagent/>) and insert the script below into a web page and “voila!” you have an animated software agent.

The only information traveling from the web server to the client browser is text (the script). Voice interaction then is scripted and encoded into the web page’s source code and then translated on the client machine by the

browser, this information is then sent to the associated text-to-speech engine. This low bandwidth multimedia format is designed specifically for the World Wide Web.

```
<OBJECT ID="AgentControl" width=0 height=0
CLASSID="CLSID:D45FD31B-5C6E-11D1-9EC1-00C04FD7081F"
CODEBASE="#VERSION=2,0,0,0">
</OBJECT>

<OBJECT ID="TruVoice" width=0 height=0
CLASSID="CLSID:B8F2846E-CE36-11D0-AC83-00C04FD97575"
CODEBASE="#VERSION=6,0,0,0">
</OBJECT>
<SCRIPT language=VBScript>
Sub Window_OnLoad()
    AgentControl.Characters.Load "Merlin", "C:\WINDOWS\MSAGENT\chars\merlin.acs"
    Set Merlin = AgentControl.Characters("Merlin")
    Merlin.LanguageID = &H409
    Merlin.Show
    Merlin.Speak "Hello and Welcome to the University of South Florida!"
    Merlin.hide
End Sub
</SCRIPT>
```

Instructional uses of Pedagogical Agents

The above description made the development of Pedagogical Agents seem simple and indeed Microsoft has made it easy. But the development process is just the beginning. The difficult part is the work of the instructional designer. It is now up to the instructional designer to take these animated characters and develop instruction.

How might they be used? How can a traditional instructor deliver instruction? The possibilities are virtually endless and the same could be said for pedagogical agents. These pedagogical agents could serve as virtual instructors... but is that a good use of the technology? Perhaps, to the novice, this may seem like the best thing to do with pedagogical agents, and it certainly is the easiest. But the author would have to disagree; this is not a good use of the technology.

Many have complained about shovelware - simply posting old lecture notes or syllabi on the web (Fraser, 1999). Are you going to have Microsoft's Merlin agent narrate your syllabus? That is possible of course, but it would be a poor use of the technology. Pure narration is not an optimal use of the technology. Students truly need to be engaged to learn. If the student has to pay attention and understand what is being said, they should learn more. It is these interactions that instructional designers will need to better understand to produce quality instruction. This overlap between Instructional Design and Human-Computer Interaction will become an increasingly more important part of our field.

Think of these agents as tutors not instructors. Rather than lecturing they can provide immediate feedback and instructions. Tutoring, rather than lecturing, can give one personal attention. This is a whole new medium and it requires us to think in new ways. So please don't follow in the footsteps of many web-based content providers and simply attempt to recreate a classroom experience. Be creative. These animated characters can be used like cartoons to entertain and motivate the learner.

Perhaps one of the most interesting uses of agents is to have a group of agents work together simultaneously to deliver an instructional message. Laurel (1993) was one of the first to suggest that the computer can act as theatre. In her book *Computers as Theatre* she makes several suggestions as to how Instructional Designers could incorporate these characters into instruction in an entertaining way. A more contemporary look comes from Janet Murray of M.I.T.; she wrote another useful book entitled *Hamlet on the Holodeck: the future of narrative in cyberspace* (Murray, 1997). This book describes the use of interactive drama, plots that are less linear, and stories that can act as games. Certainly Instructional designers can make use of these ideas to develop drama using multiple pedagogical agents.

Why use Pedagogical Agents?

Entertainment is just one consideration. But they are called "Pedagogical" for a reason. They are used to promote learning and transfer. In a series of studies, Richard Mayer and

his colleagues (Mayer & Anderson, 1991; Mayer & Anderson, 1992; Mayer & Moreno, 1998; Mayer and Moreno, 1999) tested Paivio's (1971) Dual Coding Hypothesis (later called Dual Coding Theory; D'Agostino, O'Neill, & Paivio, 1977) with multimedia. Mayer and his colleagues, and later other researchers, Jeung et al. (1997) and Lester et al. (1997), repeatedly found that students learning from multimedia presented with animation and narration consistently did better than those who learn from animation and text-based materials.

Mayer explains this effect from an information processing/cognitive load perspective. Information can be encoded both as text (visually) and narration (auditory). If it is encoded auditorily it reduces the cognitive load of the learner and they are better able to handle the incoming information. He has since called this the "Modality effect," or the Modality Principle (Mayer, 2001; Clark & Mayer, 2002). In Mayer's words the Modality Principle is that "Students learn better from animation and narration, than from animation and on-screen text (Mayer, 2001)."

Therefore if instructional designers develop instruction that includes both animation (usually depicting some process) along with narration (the voice of the pedagogical agent) students will learn better, that is they are better able to transfer their learning to new situations than those receiving text. Mayer & Moreno (1998, 1999) worked with Animated Pedagogical agents to test this very hypothesis, and found results similar to the above studies.

Social interfaces?

By incorporating agents into their web sites, Instructional Designers now can incorporate voice-based interactions into their instruction. As mentioned earlier these agents act as a "social interface" (Thorisson, 1993). Social interfaces may decrease cognitive load, as mentioned above, but this type of interface can complicate software design for Instructional Designers. This is because of the social nature of these interfaces.

In an attempt to better understand the social implications of Human-Computer Interactions Clifford Nass, a social psychologist from Stanford University, led a group of researchers, in a series of studies. Interestingly enough they observed that ordinary computer-literate individuals can be induced to use social rules toward computer agents and thus behave as if computers were human (Nass et al 1993; Nass & Steuer, 1994; Nass et al., 1997). Each of these studies and several that followed showed that people tend to apply social rules to computers. In a number of studies these researchers found that humans react differently depending on:

- the gender of the computer's "voices" (Nass et al., 1997);
- the ethnicity of the agent (Lee & Nass, 1998);
- the amount of praise offer (Fogg & Nass, 1996);
- and even if humor was involved in the interaction (Morkes, Kernal & Nass, 2000).

Nass describes this as a new paradigm in Human-Computer Interaction known as the “Computers are Social Actors” (CAS) paradigm (Nass et al., 1994).

Conclusions

In addition to acting as tools for Instructional Designers and Developers, then these social interfaces offer educational researchers a wealth of material, for they offer complex social interactions. Perhaps future research should consider these “parasocial” interactions of pedagogical agents, and the effects that they can have on learning.

There is ample evidence then that the use of pedagogical agents offers instructional designers an important tool for the development of distance learning. Pedagogical agents can be summed up in a single word, they are a tool. They are a means to an end, and that end, of course, is to promote learning.

References

- AT&T Labs. (2001). *AT&T Labs Research News - How May I Help You?* Retrieved August 3, 2001, from <http://www.research.att.com/news/2001/hmihy.html>
- Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, 37, 122-125
- Clark, R., & Mayer, R. (2002). *E-learning & the science of instruction*. San Francisco: Jossey-Bass Wiley.
- D'Agostino, P., O'Neill, B. & Paivio, A. (1977). Memory for pictures and words as a function of level of processing: Depth or dual coding? *Memory and Cognition*, 5(2), 252-256.
- Fraser, A. (1999). Colleges Should Tap the Pedagogical Potential of the World-Wide Web. *Chronicle of Higher Education*, Vol. 48
- Fogg, B. & Nass, C. (1996). *Silicon sycophants: The effects of computers that flatter*. Paper presented to the annual International Communication Association meeting, Chicago, IL.
- Guernsey, L. (2001). *Software Called Capable of Copying Any Human Voice* Retrieved August 3, 2001, from <http://www.nytimes.com/2001/07/31/technology/31VOIC.html>
- Jeung, H., Chandler, P., & Sweller, J. (1997). The role of visual indicators in dual sensory mode instruction. *Educational Psychology*, 17(3). 329-344.
- Klatt, D. (1987). Review of text-to-speech conversion in English. *Journal of the Acoustical Society of America*. 82 (3), p. 737-793.

- Laurel, B. (1993). *Computers as Theatre*. Addison-Wesley Publishing Company, Inc: Menlo Park, CA
- Lester, J., Converse, S., Kahler, S., Barlow, T., Stone, B., and Bhogal, R. (1997). The persona effect: Affective impact of animated pedagogical agents. In the *Proceedings of CHI '97*, pp. 359-366. retrieved March 22, 2002 from <http://www.acm.org/sigchi/chi97/proceedings/paper/jl.htm>
- Mayer, R. (2001). *Multimedia Learning*. Cambridge: Cambridge University Press
- Mayer, R. & Anderson, R. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83, 484-490.
- Mayer, R. & Anderson, R. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R. & Moreno, R. (1998). *A cognitive theory of multimedia learning: implications for design principles*. Paper presented at the annual meeting of the ACM SIGCHI Conference on Human Factors in Computing Systems. Los Angeles, CA.
- Morkes, J., Kernal, H. & Nass, C. (2000). Effects of Humor in Task-Oriented Human-Computer Interaction and Computer-Mediated Communication: A Direct Test of Social Responses to Communication Technology Theory. *Human-Computer Interaction*. 14(4) 395-435.
- Moreno, R., & Mayer R. (1999). Cognitive Principles of Multimedia Learning: the role of modality and Cogniguity. *Journal of Educational Psychology*. 91 (2) 358-368.
- Murray, J. (1997). *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. MIT Press: Cambridge, Mass.
- Nass, C., Steuer, J., & Tauber, E. (1994). *Computers are social actors*. Paper presented to CHI'94 conference of the ACM/SIGCHI, Boston, MA.
- Nass, C., & Steuer, J. (1994). Voices, boxes, and sources of messages: Computers and social actors. *Human Communication Research*, 19(4), 504-527
- Nass, C., Steuer, J., Tauber E., and Reeder, H. (1993). Anthropomorphism, Agency, & Ethopoeia: Computers as Social Actors. *Proceedings of the International CHI Conference*, Amsterdam, the Netherlands
- Nass, C., Moon, Y., & Green, N. (1997). Are computers gender-neutral? Gender stereotypic responses to computers. *Journal of Applied Social Psychology*, 27(10), 864-876.

Lee, E., & Nass, C. (1998) Does the Ethnicity of a Computer Agent Matter? An Experimental Comparison of Human-Computer Interaction and Computer-Mediated Communication. *Proceedings of the WECC Conference*, Lake Tahoe, CA.

Müller W., Spierling U., M. Alexa, and Rieger, T. (2000). *Face-to-face with your Assistant Realization Issues of Animated User Interface Agents for Home Appliances* Paper presented at Intelligent Interactive Assistance & Mobile Multimedia Computing 2000. Retrieved November 13, 2001, from <http://www.rostock.igd.fhg.de/veranstaltungen/workshops/imc2000/files/proceedings/09Alexa.pdf>

Paivio, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston.

Thorisson, K. (1993). Dialogue Control in Social Interface Agents. *InterCHI Adjunct Proceedings (Amsterdam, Holland)*. Reading, MA: ACM Press.

Appendix A

