Studying Situated Communication with an Embodied Agent

Ipke Wachsmuth (ipke@techfak.uni-bielefeld.de)
Artificial Intelligence Group, Faculty of Technology, University of Bielefeld
33594 Bielefeld, Germany

Keywords: embodied agents; situated communication; gesture; imitation; multimodality; emotion; virtual environments.

Motivation and Context

Embodiment has received increased attention in the study of human communication. In face-to-face interaction, interlocutors are mutually involved through different forms of embodiment (eye gaze, gesture, posture, facial expression). Meanings transmitted are multimodally encoded and strongly situated in the present context and especially the body. An important implication is that communication calls systematically on physical resources beyond those of verbal interaction. In the context of long-term collaborative research on situated communication (Rickheit & Wachsmuth, 1996), this contribution reports recent work on studying communication in a simulated environment.

Studies with Max, an Embodied Agent

With the artificial humanoid agent MAX we investigate situated communication by way of computer simulations in virtual reality. With a synthetic voice and an animated body and face, Max is able to speak and gesture, and to mimic emotions. By use of microphones and trackers, Max can also "hear" and "see" and process spoken instructions and gestures. Current research challenges are what is required for an artificial agent to exhibit multimodal behavior, imitate human gestures, and express event-related emotions.

Multimodality

Principles of human motor control are employed to animate the body of MAX, which is based on a kinematic skeleton approximating one-third of the degrees of freedom of the human body. In determining spatial and temporal relations among modalities, timing is a central issue. We successfully integrated multimodal behaviors to effect gross and subtle movements with visual acceptability and real-time responsiveness (Kopp & Wachsmuth, 2004). This allows to study, e.g., how essential features of a meaning-bearing gesture phase are produced and perceived, by way of systematically modifying Max's verbal and nonverbal behavior.

Gesture Imitation

To more fully understand the representational requirements and processes involved in gesture imitation, we investigate imitative movement by computer simulations for two cases: (1) mimicking simple gestures involving reproduction of hand positions and finger configurations; (2) imitating more complex gesture sequences related to the description of shape as elucidated in empirical study (Sowa & Wachsmuth, 2004). In a first prototype, Max was connected to a module for gesture recognition, with only the form representations of essential gesture features transferred (Kopp, Sowa, & Wachsmuth, 2004). In face-to-face interaction Max is able to imitate a great range of a human interlocutor's gestures.

Emotion

Simulation-based approaches building on appraisal theories like Ortony, Clore, & Collins' OCC model view emotions as arising from an agent's valenced reaction to events and objects in the light of goals, and attitudes, but they assume that emotion is fully cognitive and categorical. In contrast, (Becker, Kopp, & Wachsmuth, 2004) presents an approach which is based on a continuous simulation of emotion dynamics. In result, Max's facial expression, gesture, speech, and secondary behaviors as well as his cognitive functions are modulated by the emotional system that, in turn, is affected by information arising at various levels within the agent's architecture.

Conclusion

In summary, the design of human-machine interactions is of great heuristic value in the study of communication because it allows researchers to isolate, implement, and test essential properties of inter-agent communications.

References

Becker, C., Kopp, S., & Wachsmuth, I. (2004). Simulating the emotion dynamics of a multimodal conversational agent. In E. André et al. (Eds.), *Affective Dialogue Systems*. Berlin: Springer.

Kopp, S., Sowa, T., & Wachsmuth, I. (2004). Imitation games with an artificial agent: from mimicking to understanding shape-related iconic gestures. In A. Camurri & G. Volpe (Eds.), *Gesture-based Communication in Human-Computer Interaction*. Berlin: Springer.

Kopp, S., & Wachsmuth, I. (2004). Synthesizing multimodal utterances for conversational agents. *Journal of Computer Animation and Virtual Worlds, 15*, 39-52.

Rickheit, G., & Wachsmuth, I. (1996). Collaborative Research Centre "Situated Artificial Communicators" at the University of Bielefeld, Germany. *Artificial Intelligence Review, 10*, 165-170.

Sowa, T., & Wachsmuth, I. (2003). Coverbal iconic gestures for object descriptions in virtual environments: An empirical study. In M. Rector, I. Poggi & N. Trigo (Eds.), *Proceedings of the Conference "Gestures. Meaning and Use."* (pp. 365-376). Porto: Edições Fernando Pessoa.