

MAKING INTERACTIVE GUIDES MORE ATTRACTIVE

Anton Nijholt

Department of Computer Science
University of Twente, Enschede, the Netherlands
anijholt@cs.utwente.nl

Abstract – We investigate the different roads that should be taken to make 2D and 3D guides on webpages and in (augmented) virtual reality environments more attractive. Currently, most of the approaches have been done from a graphics point of view. Often research stops when we have to model how these graphically well-designed agents are going to do something useful and have to interact with the user or a visitor of an environment. What does the virtual guide know about the environment, what are its goals and how does it recognize and handle the characteristics and desires of the users/visitors it has to deal with? We survey the different approaches we take and mention some shortcomings and how they can be overcome. Topics that will be mentioned are multimodal and affective interaction, social relationships, the role of humor in interactions, web-based applications, and verbal and nonverbal communication modeling.

INTRODUCTION

We investigate how to make it more attractive to interact with embodied agents that help to explore cultural heritage environments, virtual museums and virtual exhibitions. We look at several aspects of human-human interaction in environments where there is face-to-face interaction and where it is natural to make references to a visible environment. Visualization and advanced interaction techniques, together with embodied agent research make it possible to design environments to which human-to-human communication characteristics can be translated to human-computer communication characteristics.

COMPUTERS AS SOCIAL ACTORS

In translating human-human interaction to human-computer interaction we can refer to the ‘Computers Are Social Actors’ paradigm introduced by Reeves and Nass [7]. They showed that computer users attribute human characteristics to computer systems. In short, when interacting with a computer, people attribute human characteristics to this computer. They are flattered when the computer rewards their behavior, they are less honest when they are questioned by their ‘own’ computer than by a ‘neutral’ computer, etc. As designers of human-computer interfaces we can exploit this behavior. Especially, when we have some properties of a computer system embodied in a virtual and embodied agent, we may assume that users attribute all kinds of human properties to such an agent and we can try to elicit certain behavior of the human partner that will help to get a smoother communication with the system.

COMPUTERS AS EMBODIED CONVERSATIONAL AGENTS

Embodied conversational agents (ECAs) have become a well-established research area. Embodied agents are agents that are visible in the interface as animated cartoon characters or animated objects resembling human beings. Sometimes they just consist of an animated talking face, displaying facial expressions and, when using speech synthesis, having lip synchronization. These agents are used to inform and explain or even to demonstrate products or sequences of activities in educational, e-commerce or entertainment settings. Experiments have

shown that ECAs can increase the motivation of a student or a user interacting with the system. Lester et al. [4] showed that a display of involvement by an embodied conversational agent motivates a student in doing (and continuing) his or her learning task. Some examples of embodied conversational agents are shown in Figure 1. From left to right we see: Jennifer James, a car saleswoman who attempts to build relationships of affection, trust and loyalty with her customers, Karin, informing about theatre performances and selling tickets, Steve, educating a student about maintaining complex machinery, and Linda, a learning guide.

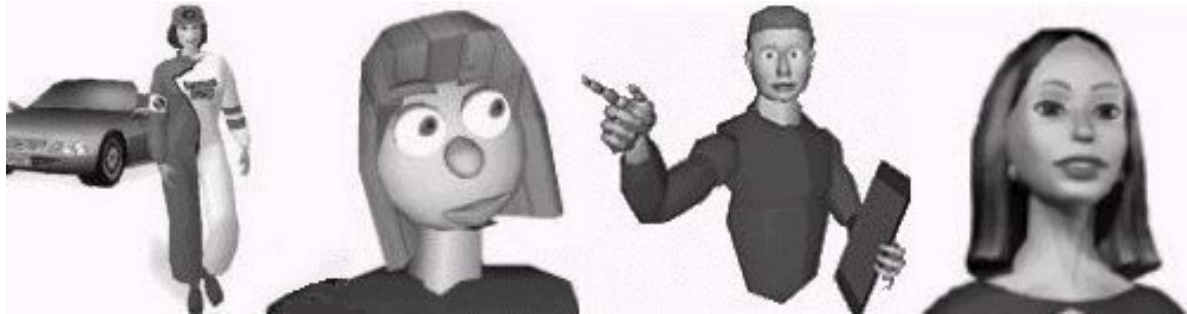


Figure 1. Examples of 2D and 3D embodied agents

Intelligence and Nonverbal Interaction

Embodiment allows more multimodality, therefore making interaction more natural and robust. Several authors have investigated nonverbal behavior among humans and the role and use of nonverbal behavior to support human-computer interaction. See e.g. [1] for a collection of chapters on properties and impact of embodied conversational agents (with an emphasis on coherent facial expressions, gestures, intonation, posture and gaze in communication) and for the role of embodiment (and small talk) on fostering self-disclosure and trust building.

Current ECA research deals with improving intelligent behavior of these agents, but also with improving their verbal and nonverbal interaction capabilities. Improving intelligent behaviour requires using techniques from artificial intelligence, in particular natural language processing. Domain knowledge and reasoning capabilities have to be modelled. Agent models have been developed that allow separation between the beliefs, desires and intentions of an agent. Together with dialogue modelling techniques rudimentary natural language interaction with such agents is becoming possible.

What role do gestures play in communication and why should we include them in an agent's interaction capability? Categories of gestures have been distinguished. Well known is a distinction in consciously produced gestures (emblematic and propositional gestures) and the spontaneous, unplanned gestures (iconic, metaphoric, deictic and beat gestures). Gestures convey meanings and are primarily found in association with spoken language. Different views exist on the role of gestures in communication. Are they for the benefit of the gesturer or for the listener? In Kendon's view [3] gestures convey extra information about the internal mental processes of the speaker: ". . . *an alternative manifestation of the process by which ideas are encoded into patterns of behavior which can be apprehended by others as reportive of ideas.*" Observations show that natural gestures are related to the information structure (e.g., the topic-focus distinction) and (therefore) the prosody of the spoken utterance. In addition they are related to the discourse structure and therefore also to the regulation of interaction (the turn taking process) in a dialogue. Apart from these viewpoints on embodiment, we can also emphasize the possibility of an embodied agent to walk around, to point at objects in a visualized domain, to manipulate objects or to change a visualized (virtual) environment. In these cases the embodiment can provide a point of the focus for interaction. When we intro-

duce a guide in our virtual environments this is a main issue and more important than detailed facial expressions and much of the gestures discussed above.

Emotional Behavior, Personality, Friendship

Facial expressions and speech are the main modalities to express nonverbal emotion. Human beings do not express emotions using facial expressions and speech only. Generally they have their emotions displayed using a combination of modalities that interact with each other. We cannot consider one modality in isolation. Facial expressions are combined with speech. There are not only audio or visual stimuli, but also audio-visual stimuli when expressing emotions. A smile gesture will change voice quality, variations in speech intensity will change facial expression, etc. Attitude, mood and personality are other factors that make interpretation and generation of emotional expressions even less straightforward. In addition we can have different intensities of emotion and the blending of different emotions in an emotional expression. We should consider combinations and integration of speech, facial expressions, gestures, postures and bodily actions (see [5]). It should be understood that these are displays and that they should follow from some emotional state that has been computed from sensory inputs of a human interactant, but also from an appraisal of the events that happen or have happened simultaneously or recently. A usual standpoint is that of appraisal theory, the evaluation of situations and categorizing arising affective states. It should be understood that what exactly is said and what exactly is done in a social and emotional setting is not part of the observations above. The importance of the meaning of words, phrases and sentences, uttered and to be interpreted in a specific context, is not to be diminished. In Figure 2 we display Cyberella, an embodied agent, developed at DFKI in Saarbrücken. This agent is working as a receptionist. For example, she can provide directions to the office of a staff member. However, since she has been provided with an affective model, she also reacts emotionally to a visitor's utterances when appropriate [2].



Figure 2: Cyberella, a virtual receptionist

One of the issues we investigated was how aspects of personal attraction or friendship development [8] can be made part of the design of an embodied agent that is meant to provide an information service to a human partner. As a 'lay psychologist', we all know that people that you like (or your friends) are able to help you better, teach you better, and generally are more fun to interact with, than people that you don't like. However, 'liking' is person dependent. Not everybody likes the same person, and one person is not liked by everyone. These observations sparked our interest in the application, effects, and design of a 'virtual friend'. An agent that observes its user, and adapts its personality, appearance and behavior according to the (implicit) likes and dislikes of the user, in order to 'become friends' with the user and create an affective interpersonal relationship. This agent might have additional benefits over a 'normal' embodied conversational agent in areas such as teaching, navigation assistance and entertainment.

There is extensive knowledge about human interpersonal relationships in the field of personality and social psychology. Aspects of friendship that need to be considered in ECA design are gender (e.g., activity-based men's friendship vs. affectively-based women's friendship), age, social class and ethnic background. Effects of friendship on interaction include increase of altruistic behavior, a positive impact on task performance and an increase in self-disclosure. Interpersonal attraction is an important factor in friendship. It is governed by positive reinforcements, and similarity between subjects is a key factor. Similarity of attitudes, person-

ality, ethnicity, social class, humor, etc., reinforces the friendship relationship. Other issues are physical attractiveness (the ‘halo effect’) and reciprocity of liking (whether we think that the other person likes us). In [8] we discussed the translation of the main aspects of human-human friendship to human-ECA friendship and how we can incorporate this translation in the design process of an ECA, using a scenario-based design. One observation is that it is important to distinguish between the initial design of an ECA and the possibility to change the ECA characteristics according to an adaptation strategy based on knowledge obtained by interacting with a particular user.

Humor in Interpersonal Interaction

In previous years researchers have discussed the potential role of humor in the interface. Humans use humor to ease communication problems and in a similar way humor can be used to solve communication problems that arise with human-computer interaction. For example, humor can help to make the imperfections of natural language interfaces more acceptable for the users and when humor is sparingly and carefully used it can make natural language interfaces much friendlier. During these years the potential role of embodied conversational agents was not at all clear, and no attention was paid to their possible role in the interface. In [6] we discussed the role of humor for ECA’s in the interface.

Humans employ a wide range of humor in conversations. Humor support, or the reaction to humour, is an important aspect of personal interaction and the given support shows the understanding and appreciation of humor. There are many different support strategies. Which strategy can be used in a certain situation is mainly determined by the context of the humorous event. The strategy can include smiles and laughter, the contribution of more humor, echoing the humour and offering sympathy. In order to give full humor support, humor has to be recognized, understood and appreciated. These factors determine our level of agreement on a humorous event and how we want to support the humor.

Humor plays an important role in interpersonal interactions. From the many CASA experiments we may extrapolate that humour will play a similar role in human-computer interactions. This has been confirmed with some specially designed experiments. There is not yet much research going on into embodied agents that interpret or generate humour in the interface. In [6] we discuss how useful it can be, both from the point of view of humor research and from the point of view of embodied conversational agent research, to pay attention to the role of humor in the interaction between humans and the possibility to translate it to the interactions between humans and embodied conversational agents. Graphics, animation and speech synthesis technology make it possible to have embodied agents that can display smiles, laughs and other signs of appreciation of the interaction or explicitly presented



Figure 3: “You go UP the STAIRS.”

or generated humor. There are many applications that can profit from being able to employ such embodied agents. The designer of the interface can decide when in certain scenarios of interaction agents should display such behaviour. However, much more in the line of research on autonomous (intelligent and emotional) agents we rather have an agent understand why the events that take place generate enjoyment by its conversational partner and why it should display enjoyment because of its appreciation of an humorous situation.

MULTIMODALITY: INTERPRETATION AND GENERATION

When discussing the role of embodied agents in 2D or 3D environments we have to deal with the (multimodal) integration of information coming from different sources and the generation of different actions given a multimedia presentation context. Depending on the modalities available in the system decisions on how to react on (combination of) inputs from a user can on the one hand concern issues like turntaking, reference resolution, topic shift, or asking a clarifying question, while on the other hand it may concern the choice of presentation modalities when providing information and drawing the attention of the user to certain information. When a virtual guide is embodied, apart from the social relationship properties mentioned above that we would like to include in the guide's behavior towards his or her 'customers', we also want this guide to assist the user in exploring the environment and its objects by moving around, pointing at objects, guiding the user to objects and locations, support speech communication by nonverbal gestures, etc. It is, show behavior that we can recognize in a human guide.

In conclusion, it is certainly not the case that at this moment we can come close to the modeling of a human museum or exhibition guide. However, current research allows to make comparisons and it also shows that a clever use of technology can make imperfect technology acceptable or even natural. Currently we are working on an embodied version of our navigation agent (see Figure 3, [9]). The main issue we have to deal with is the mutual interaction between pointing behavior of the agent while showing directions and the agent's utterance generation during the interactions with the visitor.

References

- [1] J. Cassell et al. (eds.). *Embodied Conversational Agents*. The MIT Press, 2000.
- [2] P. Gebhard. Enhancing Embodied intelligent agents with affective user modelling. *UM2001, 8th International Conference*, J. Vassileva and P. Gmytrasiewicz, (eds.), Berlin, Springer, 2001.
- [3] A. Kendon. Gesticulation and speech: two aspects of the process of utterance. In: *The relation of verbal and nonverbal communication*. M.R. Key (ed.), Mouton, The Hague, the Netherlands, 1980.
- [4] J.C. Lester et al. The persona effect: Affective impact of animated pedagogical agents. *CHI '97 Human Factors in Computing Systems*, ACM, 1997, 359-356.
- [5] A. Nijholt. Issues in multimodal nonverbal communication and emotion in embodied (conversational) agents. *6th World Multiconference on Systemics, Cybernetics and Informatics. Vol. II*. N. Callaos, A. Breda & Y. Fernandez (eds.). July 2002, Orlando, USA, 208-215.
- [6] A. Nijholt. Embodied Agents: A New Impetus to Humor Research. *The April Fools Day Workshop on Computational Humour*, O. Stock, C. Strapparava & A. Nijholt (eds.), In: Proc. Twente Workshop on Language Technology 20 (TWLT 20), Trento, Italy, 2002, 101-111.
- [7] B. Reeves & C. Nass. *The Media Equation: how people treat computers, televisions and new media like real people and places*. Cambridge, Cambridge University Press, 1996.
- [8] B. Stronks, A. Nijholt, P. van der Vet & D. Heylen. Designing for friendship: Becoming friends with your ECA. *Embodied conversational agents - let's specify and evaluate them!* A. Marriott et al. (eds.), Bologna, Italy, July 2002, 91-97.
- [9] M. Theune, A. Nijholt & D. Heylen. Generating embodied information presentations. Chapter in: *Intelligent Information Presentation*. O. Stock & M. Zancanaro (eds.), Kluwer Academic Publishers, 2003.