AALBORG UNIVERSITY
DENMARK

Gamble

A Multiuser Game with an Embodied Conversational Agent

Rehm, Matthias; Wissner, Michael

Published in:
Entertainment Computing (ICEC)

Publication date:
2005

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Gamble — A Multiuser Game with an Embodied Conversational Agent

Matthias Rehm and Michael Wissner

Multimedia Concepts and Applications,
University of Augsburg, Germany
`rehm@informatik.uni-augsburg.de, michael.wissner@web.de`

Abstract. In this article, we present Gamble[1], a small game of dice that is played by two users and an embodied conversational agent (ECA). By its abilities to communicate and collaborate, an ECA is well suited for engaging users in entertaining social interactions. Gamble is used as a test bed for such multiuser interactions. The description of the system's components and a thorough analysis of the agent's behavior control mechanisms is followed by insights gained from a first user study.

1 Introduction

Entertainment computing has so far successfully focused on the technical problems of enhancing the user's experience (e.g., [23], [7], [20]). An equally important factor are the social interaction qualities of a game which play an especially crucial role in multiplayer games. Talking about multiplayer games often means talking about the interaction of multiple users mediated by a computer. In this article, we present a game instead that is played by multiple users together with an embodied conversational agent (ECA) that serves as one of the interaction partners.

Why do ECAs make sense for an application like a game? Apart from their conversational skills, the nonverbal behavior and the appearance of embodied interface agents become more and more realistic. Thus, ECAs serve as an ideal tool for engaging users in social interactions like tutoring, gaming, etc. According to Sidner [21], engagement "is the process by which two (or more) participants establish, maintain and end their perceived connection during interactions they jointly undertake." Engagement behaviors comprise spoken linguistic behavior, i.e. the ability to communicate by speech, collaborative behavior, i.e. the ability to do something together with others, and gestural behavior, i.e. the ability for multimodal interactions including body movements and eye gaze. All of these behaviors are also essential ingredients of ECAs. But most ECA systems constrain themselves to interactions with single users or with single users and other agents (e.g., [6], [11], [13]) because dealing with multiple users is much more

---

[1] This work is supported by the EU Network of Excellence HUMAINE (http://emotion-research.net).

difficult than dealing with a single user. Although the setting in a specific application might constrain the relevant interaction, such a multiuser interaction is much less predictable. The users might show any behavior, e.g., by collaborating against the agent, by discussing off topic matters, etc. In the dyadic situation the only interaction partner is the agent triggering the user to regard her "traffic rules" of social interaction. This might be not the case if there is a real human as an interaction partner available apart from the agent. There are few systems that risk dealing with multiple users and thus, the literature on the behavior of multiple users is sparse. An exception is the work of Vertegaal and colleagues, who studied gaze behavior in multiuser settings ([22]).

Collaborating or competing with others in a game has one important prerequisite: sharing the interaction space with others be they humans or game characters. One way to realize this is to use augmented reality techniques to integrate the virtual world into the real world. Impressive examples of mixed reality games are Invisible Train [23], Human PacMan [7], and MR Mystery game [20]. The Invisible Train installation aims at simplifying the technical prerequisites necessary for an augmented reality experience by employing standard hand-held devices as a window to the virtual world. The players control virtual trains on real wooden tracks. In the Human PacMan system, users play the roles of Pacman and the ghosts in a real world setting competing and collaborating to win the game. The Mixed Reality Mystery game at last takes place in an art gallery, where the user has to deal with a robbery. Real paintings of Hopper are augmented by virtual information, allowing e.g., to explore the back alley of the bar Nighthawks. All of these have one feature in common. There is no ECA involved in the interaction. If we change our focus towards ECA systems, sharing the interaction space is often accomplished by using large screens (e.g., Cassell's REA system [6]) and/or impressive sound systems (e.g., the MRE system [11]), by using a CAVE (e.g., the Multimodal Assembly eXpert (MAX) [13]), or by using head-mounted displays to join the other characters in the virtual world (e.g.,
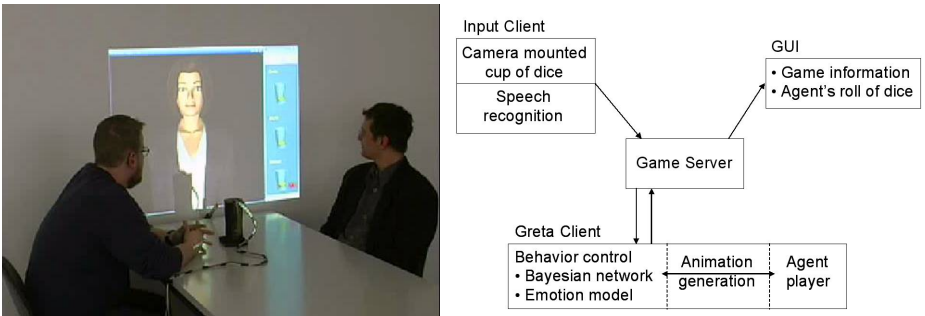

Fig. 1. The setting (left) and the system architecture (right). The agent is projected to the wall at the end of the table, thus sitting together with the other players around the table. Between the human players, the CamCup and the microphone are visible.

Rickel's Steve agent [19]). Due to the complexity of coordinating an autonomous agent with a freely moving user, the use of augmented reality techniques in ECA systems has just started (see [1] for an overview). A different approach uses the application's setting to augment the presence of the ECA and create a shared interaction space, e.g., in the MACK system [15]. The metaphor for the MACK system is an information booth in the entrance area of a building. The agent is situated on a screen that shows the real world area behind the agent making use of a video camera. Thus, the agent's screen is its booth in which it works its shifts. In Gamble, we follow a similar approach to augment the ECA's presence. Gamble is a small game of dice played at a table. This setting naturally constrains the possible interaction space. No sophisticated movements through space are supposed to take place, the agent is sitting together with the other players at the table by projecting it to a screen at the end of the table (see Fig. 1 (left)). In the remainder of this article, we will first describe the rules of the game (Sec. 2), followed by a description of the system itself (Sec. 3) and the insights gained by a first evaluation study (Sec. 4).

2   Gamble: The Game

In Gamble, two users play a simple game of dice (also known as Mexicali) with an embodied conversational agent. To win the game it is indispensable to lie to the other players and to catch them lying to you. The traditional (not computer-based) version of the game is played with two dice and a cup.

Let's assume it is the turn of player 1. He casts the dice and then inspects the dice without permitting the other players to have a look. The cast is interpreted in the following way: the higher digit always represents the first part of the cast. Thus, a 5 and a 2 correspond to a 52. Two equal digits (11, ..., 66) have a higher value than the other casts, the highest cast is a 21. Player 1 has to announce his cast with the constraint that he has to say a higher number than the previous player. For instance, if the dice show a 52, but the previous player already announced a 61, player 1 has to say at least 62. Now player 2 has to decide whether to believe the other player's claim or not. In this case, he has to cast next. Otherwise, the dice are shown and if player 1 has lied he has lost this round and has to start a new one.

Although the rules of the game are very simple, complex behaviors emerge from these simple rules. Blaming another player for an attempted deceit or getting away with such an attempt e.g., creates highly emotional situations that trigger rich social interactions allowing us to use Gamble as a test bed for investigating social behavior of an agent towards users and vice versa. Section 4 reveals insights gained from a first user study.

3   Gamble: The System

The architecture of our system can be seen in Figure 1 (right): The game server is the central element. It manages turn order, collects announcements and casts
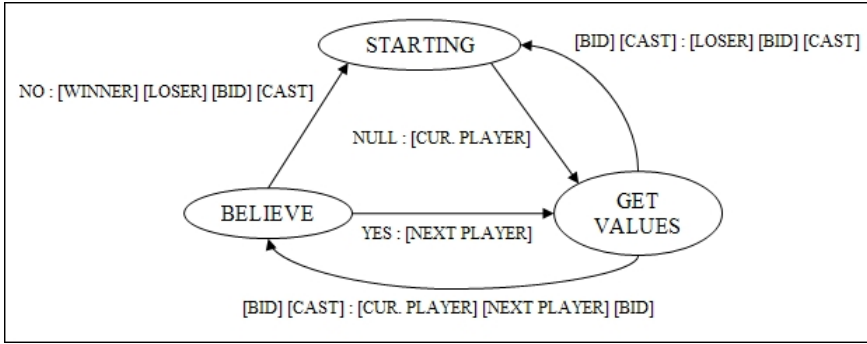
Fig. 2. The game protocol. Depending on the players input, transitions between states are initiated and corresponding output is sent by the game server to the different clients.

from both the human players and the agent, calculates the results and sends all relevant game information to the agent and the GUI. The input client gathers the user input through the CamCup (actual cast) and the speech recognition (announcements) and sends them to the game server. The Greta client contains the agent's behaviour control and animation generation. It sends the agent's announcement and cast to the game server and selects and generates the multimodal output to be displayed by the agent player.

3.1 Game Server

The communication between the game server and the other components is realized via sockets, with the game server acting as a socket server while its clients (input client, Greta client, GUI) are socket clients. It is written in Java and the server's protocol is implemented through a finite state transducer (see Fig. 2).

- STARTING: At the beginning of each round the server is in the state STARTING. The server awaits no input from the clients but broadcasts whose player's turn it is now. At the beginning of a new game, this player is randomly selected, after that it is always the player who lost the last round.
- GETVALUES: In state GETVALUES the server awaits a CAST and a BID from the current player and a simple acknowledgement from everyone else. If the bid is incorrect (i.e. it was lower than the previous player's bid) the current player loses instantly, the server broadcasts the round's result (LOSER) and switches back to state STARTING for a new round. However, if the bid is correct, the server asks the next player to rate the current player's bid and switches to state BELIEVE.
- BELIEVE: If the next player answers YES, she becomes the current player and the server broadcasts that it's her turn. The state switches to GET-VALUES. With an answer of NO on the other hand, the server compares the current player's cast and bid, determines the WINNER and switches to state STARTING in order to start a new round.

Fig. 3. The CamCup (left) and different GUI states (right). The cup is augmented by a webcam which registers the casts. Basically, three GUI states can be distinguished: (i) Rating announcements (left), (ii) casting dice by user (middle), (iii) casting dice by agent (right).

3.2   Input Client

As mentioned above the input client gathers all the human players input and sends it to the game server. There are two kinds of input: announcements and casts. Announcements (bids and believe-statements) are received through speech recognition while the dice casts are registered with the CamCup.

Speech recognition. A microphone is set between the players and records their announcements. These announcements are then processed by the speaker independent ESMERALDA speech recognition system ([10]) which was trained to recognize all possible numbers and a few variations of 'yes' and 'no' like 'I believe you' or 'Never'. The result of the speech recognition process is then forwarded to the input client.

The CamCup. The dice cup used by Gamble (see Fig. 3 (left)) contains an USB-camera in order to recognize the dice cast. Once the dice recognition is started, the camera takes one picture every 200 milliseconds and compares every pair of successive pictures. If these two pictures do not differ, the dice recognition assumes that 'the dice are cast' and begins to analyze the last picture with a blob coloring procedure. Once this procedure is complete the recognised cast is returned to the input client as a two-digit number.

3.3   GUI

Gamble's GUI is designed in Macromedia Flash and displays all relevant game information to the users (see Fig. 3 (right)). Each player's score is displayed by a jigger which is more or less filled with a yellowish liquid. There are six different 'fill-states' from empty (the player has not lost until now) to full (the player has lost five times and the game is over if he loses once more) . Besides, the current player is indicated by a small symbol next to her jigger: a pair of dice if it's her
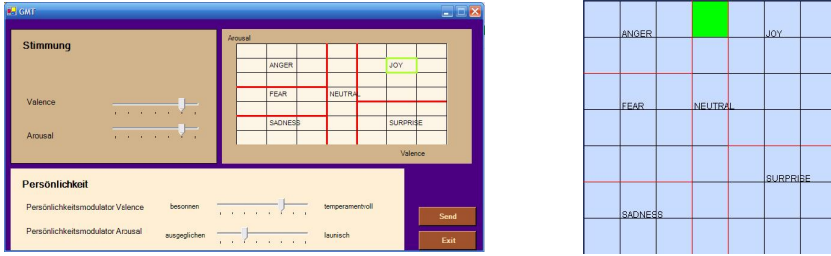
Fig. 4. Setting the agent's initial emotion (left, above) and its personality (left, below) and visualising the agent's emotional state (right)

turn to cast or a thought bubble with a question mark if it's her turn to rate another player's announcement. To indicate that the agent is currently casting the dice, a short video is shown.

3.4 Greta Client

The ECA supplied in Gamble is the Greta agent by Catherine Pelachaud[2] and colleagues ([8]). Controlling the generation of animations (facial and gestural behavior) as well as speech is conversation triggered, i.e., by augmenting utterances of the agent with special tags supplied by APML (Affective Presentation Markup Language, [5]). Based on the utterance and the given tags, corresponding facial expressions, gestures, and speech are generated and afterwards played back with the agent player. To allow for a seamless interaction in Gamble, some modifications were necessary.

Behavior control. Before generating the animations, the appropriate, i.e., context and situation specific, verbal and non-verbal behaviors have to be decided on. A Bayesian network is deployed for this reasoning process. Depending on the evidence available, the network calculates probabilities for possible actions. A turn in the game can roughly be divided into two phases: rating and announcing. First, the announcement of the previous player has to be rated. This decision is based upon (i) Greta's current emotional state, (ii) the probability of the announced result, e.g., the highest cast 21 has a probability of 0.056, (iii) the number of times that the previous player was caught lying. If the agent believes the previous player or has falsely accused him of lying, it has to cast the dice next and announce a result. The announcement is based upon (i) Greta's cast, (ii) the probability of the necessary announcement, (iii) the number of times Greta was caught lying, (iv) the agent's emotional state.

The agent's emotional state is influenced by its game success and by its personality traits. Catching another player lying, getting away with a lie, or being falsely accused of a lie and thus winning the round constitute a positive emotional influence. Falsely accusing another player or being caught lying on

the other hand constitute a negative emotional influence. The emotional model is dimensional in nature (e.g., [14]) with one dimension denoting the arousal of the accompanying emotion and the other dimension denoting its valence on a positive/negative axis. The reason for the choice of this emotion model was triggered by the abilities of automatic recognition systems for emotions to detect those dimensions in physiological ([4]) or speech signals ([2]). It is planned to test such recognition algorithms with the Gamble system. The agent's initial emotional state as well as its personality traits can be set before starting the game (see Fig. 4 (left)). Instead of using a sophisticated personality model like the big five (e.g., [12]), we take the dimensional model into account directly. The user can determine modulator values for valence and arousal. These modulators allow the agent different appraisals of a given situation. A high modulator value for valence is interpreted as a highly emotion driven decision process changing fast between positive and negative evaluation of a situation whereas a low modulator value results in a more rational decision. The arousal modulator on the other hand determines how capricious the agent reacts. A high value of the arousal modulator results in a fast increase of the arousal level in a given situation whereas a low value slows the increase down making the agent more phlegmatic. The agent's emotional state can be monitored (see Fig. 4 (right)) but is visible to the user only by the agent's behavior.

The information about the actual cast and the announcement are sent to the game server. Output of the behavior module for the animation generation is the result that will be announced by Greta, the emotional state of the agent in terms of valence and arousal, and the current ability to mask a necessary lie. Because it would be frustrating to play against an agent that is able to show a perfect poker face, we modelled some facial clues of deceit that shall allow the user to catch the agent lying (see Sec. 4.1). The ability to mask a necessary lie depends in our model on the arousal and valence value of the emotional state, on the probability for the announced result of the previous player, and on the probability of Greta's own announcement.

Animation generation. Among other things, APML allows for enriching the utterance of an agent with information on the emotional state of the agent (e.g., surprised, sad), meta information about the dialogue context (e.g., greeting, information), and gestural information (e.g., iconic for small). For the use in Gamble the gestural repertoire was extended by German emblems described in detail in the Berlin lexicon of german everyday gestures ([3]). The semantic content of emblems is well defined because they generally are used instead of words, conveying a specific meaning like the OK sign in the American culture formed by thumb and index finger. Moreover, many emblems are especially suited for a game context, because there exist specific emblems, e.g., for surprise or disbelief which occur naturally in such a context. 30 emblems have been created for Gamble.

Whereas APML allows for defining unlimited emotional facial expressions, the current lip model which relies on recorded data restricts the facial expressions to some basic emotions (anger, disgust, fear, joy, sadness, surprise, neu-

tral). Thus, the information about the emotional state of the agent has to be mapped to these basic emotions before generating the agent's utterance. According to the current emotional state of the agent and the kind of statement (belief/disbelief, announcement, comment), an utterance and an animation is chosen from a database, which is then send to the agent player. To generate realistic utterances for the agent, video recordings of human players were analysed to gain a corpus of game relevant utterances. Next step is the use of a dynamic template based speech generation approach that takes into account the emotional state of the agent and generates appropriate utterances and animations on the fly.

Agent player. In our modified version we established a client/server connection between the animation generation and the agent player via sockets. The latter is the client and waits for play-back instructions from the generation module. Having received such an instruction the player loads the corresponding animation and audio files. After finishing the play-back it switches into an idle mode and awaits the next instruction. In order to make these idle phases look more natural we implemented some idle movements (shifting, nodding, gazing around) for the agent. When entering the idle state, the player randomly chooses and plays an idle movement, depending on the agent's current emotional state. For each emotional state, at least three different variants of the idle movement exist. Moreover, the strength of the movements can be modified, e.g., if arousal is very high, the movement can be rendered more agitated and space consuming. If the player receives a play-back instruction from the generation module while playing an idle movement, the movement is blended into the new animation within 20 frames to grant a smooth transition between the two animations.

4   Gamble: The Test Bed

Gamble is a test bed for affective multiparty interactions which allows us to investigate social behavior of an agent towards users and vice versa. The CASA paradigm (Computers Are Social Actors, e.g., [16]) claims that agents are regarded and treated as social actors. Whereas in 1:1 interactions it was shown that people tend to consider their "traffic-rules" of social interaction, we know little about scenarios in which more than one user is interacting with the same agent at the same time. Gamble serves as a test bed for aspects of social interactions and the remainder of this article we present some results from a first evaluation study on deceiving agents, multiuser and affective interactions.

24 students from computer science and from communication science participated in this first user study. They were divided in 12 groups, each pair played two sessions of 12 minutes ensuring that each participant played before and after the agent. Thus, each participant had to lie to the agent some time in the game and each had to rate the agent's announcements. To ensure the participants interest in the game, winning the game would earn them 5 Euros. The general game information like actual cast, announced result, and success was logged

during the game. The sessions were recorded on video and the videos have been annotated considering for each player – human and agent – the utterances, the gaze behavior, the role in the game at that moment in time, and if the player was laughing. This last bit of information is used to train and test a system for automatic recognition of emotion from speech.

4.1   Deceiving Agents

To win the game it is indispensable to sometimes deceive the other players and to detect such attempts by the other players. Playing with an agent that is able to conceal its attempts completely, might be not a very satisfactory experience. Ekman [9] describes facial clues to deceit, that can only be controlled by the most experienced liars. Some of these clues were modelled for the Greta agent allowing it to control the level of concealment during the game. Thus, Gamble allows us to investigate in a principled way (i) if users react at all to such facial clues of deceit when an ECA shows these clues, and (ii) if this is the case, how they react to and interpret these clues, e.g., as a malfunction of the system, as an affront, or as a useful feature that makes the interaction more engaging. Although a pilot study revealed that users react to such clues shown by an agent, in the context of the game they seem to be far more engaged in the interaction to pay any attention to these clues. A thorough discussion of the results obtained by the first user study on this account can be found in [17]. Of course, also the human players try to deceive the agent and consquently we will have to deal with the question how to handle false expressions by the user, i.e. how to react to them. Such a reaction is necessary if the agent has caught the previous player lying. At the moment this situation is handled by predefined gloating behaviors. E.g., the agent displays an emblematic gestures that is recognized as gloating in Germany (holding the hand as an extension of the nose) and says "You have lost".

4.2   Multiuser Interaction

Furthermore, Gamble allows us to study emotional interaction within a multi-party scenario with several human users – an aspect which has been largely neglected so far. Our first study revealed the high engagement of the users in the interaction with the agent and with one another. And although the game is round-based and thus allows for a thorough control of turns, all kinds of social behavior could be seen, from testing the agent's domain competence over commenting on the agent's and/or the other player's moves up to a collaboration of players against the agent. What is evident in the video recordings of the first user study is the acceptance of the agent as a competent interaction participant.

Gaze behavior is generally interpreted as an indicator of the user's engagement in an interaction. Annotating the users gaze behavior in the video recordings revealed a high interest in the agent (see [18] for a thorough discussion). Overall, users looked 1.5 times longer towards the agent than to their human interaction partner. Moreover, users accepted the agent as a social interaction

partner although there is also a another real partner present. Users showed the same gaze patterns as in natural interactions ([22]). As a speaker, they first looked away and then towards the agents which was also described for human face to face interactions. As a listener, they looked towards the agent during its announcements which again is a natural behavior. At the moment, the agent shows no active gaze behavior but only randomly gazes towards the users. Thus, the results from the first user study will be used to develop a gaze model for such a multiuser setting.

4.3   Affective Interaction

Additionally, using Gamble, we can create highly emotional situations, for example when the agent blames the user for deceit or when the user detects such an attempt and the agent has to react to it. Until now, there is some anecdotal evidence from the video recordings for this effect. Reacting to an attempted lie, e.g., with the utterance "Du hast einen totalen Knall"[3] accompanied by a gesture where the agent is waving a splayed hand to and fro in front of its face, often elicits an appropriate reaction by the users, e.g., comments like "Ich wollt's mal versuchen"[4] or "Hätt ja klappen können"[5]. In another example, the agent starts swearing after the user has repeatedly caught it lying. This results always in a big laugh from the users because it is an appropriate social behavior in the context that was not expected from a technical artefact like an ECA and renders the agent even more interesting. Consequently, Gamble enables us to investigate how emotional signals employed by the agent are perceived by the human user. In particular, we want to study how emotions need to be conveyed in order to increase the user's trust in an agent. Measuring the user's affective states by means of physiological sensors and automatic speech recognition system, we will investigate how different expressive behaviors of the agent exert an influence on these states.

5   Summary

In this article we presented an ECA system that engages two users in an entertaining interaction. Because of their ability to interact in a socially appropriate manner, ECAs are well suited as a competent and natural interaction partner. The agent in Gamble provides a number of engagement behaviors that are necessary for drawing the users into the interaction. Its conversational behavior, i.e., ist ability to communicate by speech and gestures, was tailored to the domain by analyzing video recordings of human players to create game relevant utterances as well as by augmenting the agent's gestural repertoire to allow for appropriate accompanying gestures (German emblems). The agent's collaborative behavior,

---

[3] You are stupid.
[4] I just wanted to try.
[5] Could have worked.

i.e., its ability to join the human players in the game of dice, renders it a competent game partner. Enhancing the behavior control by an emotional model and some simple personality features, the agent's behavior is no longer predictable by the basic facts of the game like the probability of a cast, but becomes much more interesting by enhancing the decision process with uncertainty factors. The setting of the game (round based, played at a table) allows us to integrate the agent in a natural way. Projecting it to a screen at the end of the table creates the impression of sitting at the table together with the other players. What is lacking in terms of engagement behaviors at the moment is an active gaze behavior of the agent. The observations of the first user study show some similarities with natural multiuser communication and also some peculiarities like increased attention towards the agent. These observations will inform our gaze model for the agent allowing for an even more engaging interaction.

References

1. Elisabeth André, Klaus Dorfmüller-Ulhaas, and Thomas Rist. Embodied Conversational Characters: Wandering between the Digital and the Physical World. *it-Information Technology*, 2004.
2. A. Batliner, V. Zeißler, C. Frank, J. Adelhardt, R. P. Shi, and E. Nöth. We are not amused - but how do you know? User states in a multi-modal dialogue system. In *Proc. EUROSPEECH 2003*, pages 733–736, Geneva, 2003.
3. BLAG. Berliner lexikon der alltagsgesten. *http://www.ims.uni-stuttgart.de/ projekte/nite/BLAG/*, last visited: 22.05.2005.
4. Wauter Bosma and Elisabeth André. Exploiting Emotions to Disambiguate Dialogue Acts. In *Proceedings of the 9th International Conference on Intelligent User Interfaces*, pages 85–92. ACM Press, 2004.
5. Berardina De Carolis, Valeria Carofiglio, Massimo Bilvi, and Catherine Pelachaud. APML, a Mark-up Language for Believable Behavior Generation. In Z. Ruttkay and C. Pelachaud, editors, *Workshop AAMAS02: Embodied conversational agents — let's specify and evaluate them!*, 2002.
6. Justine Cassell, Timothy Bickmore, Lee Campbell, Hannes Vilhjalmsson, and Hao Yan. Designing embodied conversational agents. In Justine Cassell, Joseph Sullivan, Scott Prevost, and Elisabeth Churchill, editors, *Embodied conversational agents*, pages 29–63. MIT Press, Cambridge, MA, 2000.
7. Adrian David Cheok, Siew Wan Fong, Kok Hwee Goh, Xubo Yang, Wei Liu, Farzam Farzbiz, and Yu Li. Human Pacman: A Mobile Entertainment System with Ubiquitous Computing and Tangible Interaction over a Wide Outdoor Area. In L. Chittaro, editor, *Mobile HCI 2003*, pages 209–224. Springer, Berlin, Heidelberg, 2003.
8. Fiorella de Rosis, Catherine Pelachaud, Isabella Poggi, Valeria Carofiglio, and Berardina De Carolis. From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human-Computer Studies*, 59:81–118, 2003.
9. Paul Ekman. *Telling Lies — Clues to Deceit in the Marketplace, Politics, and Marriage*. Norton and Co. Ltd., New York, 3rd edition, 1992.
10. G. A. Fink. Developing HMM-based recognizers with ESMERALDA. In V. Matoušek, P. Mautner, J. Ocelíková, and P. Sojka, editors, *Lecture notes in artificial intelligence*, pages 229–234. Springer, Berlin, Heidelberg, 1999.

11. Randall W. Hill, Jonathan Gratch, Stacy Marsella, Jeff Rickel, William Swartout, and David Traum. Virtual humans in the mission rehearsal exercise system. *KI – Künstliche Intelligenz*, (4):5–10, 2003.
12. O. P. John. The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin, editor, *Handbook of personality: Theory and research*, pages 66–100. Guilford, New York, 1990.
13. Stefan Kopp, Bernhard Jung, Nadine Lessmann, and Ipke Wachsmuth. Max — A Multimodal Assistant in Virtual Reality Construction. *KI – Künstliche Intelligenz*, (4):11–17, 2003.
14. Peter J. Lang. The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*, 50(5):372–385, 2002.
15. Yukiko I. Nakano, Gabe Reinstein, Tom Stocky, and Justine Cassell. Towards a Model of Face-to-face Grounding. In *Proceedings of the Association for Computational Linguistics*, Sapporo, Japan, July 1–12 2003.
16. Byron Reeves and Clifford Nass. *The Media Equation — How People Treat Computers, Television, and New Media Like Real People and Places.* Cambridge University Press, Cambridge, 1996.
17. Matthias Rehm and Elisabeth André. Catch me if you can — Exploring lying agents in social settings. In *Proceedings of AAMAS'05*, 2005.
18. Matthias Rehm and Elisabeth André. Where do they look? Gaze Behaviors of Multiple Users Interacting with an Embodied Conversational Agent. In *Proceedings of Intelligent Virtual Agents (IVA)*, 2005.
19. J. Rickel and W. J. Johnson. Task-oriented collaboration with embodied agents in virtual worlds. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors, *Embodied conversational agents*, pages 95–122. MIT Press, Cambridge, MA, 2000.
20. Jorge Santiago, Luis Romero, and Nuno Correia. A mixed reality mystery game. In *Proceedings of the second international conference on Entertainment computing*, pages 1–8, Pittsburgh, Pennsylvania, 2003.
21. Candace L. Sidner, Cory D. Kidd, Christopher Lee, and Neal Lesh. Where to look: a study of human-robot engagement. In *Proc. of the 9th Int. Conf. on Intelligent User Interface*, pages 78–84, 2004.
22. Roel Vertegaal, Robert Slagter, Gerrit van der Veer, and Anton Nijholt. Eye Gaze Patterns in Conversations: There is More to Conversational Agents Than Meets the Eyes. In *Proceedings of SIGCHI 2001*, Seattle, WA, 2001.
23. Daniel Wagner, Thomas Pintaric, Florian Ledermann, and Dieter Schmalstieg. Towards Massively Multi-User Augmented Reality on Handheld Devices. In *3rd International Conference on Pervasive Computing (Pervasive 2005)*, Munich, 2005.