PAPER Special Issue on Software Agent and its Applications

# Designing and Evaluating Animated Agents as Social Actors

**Helmut PRENDINGER**†, ***Nonmember and*** **Mitsuru ISHIZUKA**†, ***Regular Member***

**SUMMARY** Recent years have witnessed a growing interest in employing animated agents for tasks that are typically performed by humans. They serve as communicative partners in a variety of applications, such as tutoring systems, sales, or entertainment. This paper first discusses design principles for animated agents to enhance their effectiveness as tutors, sales persons, or actors, among other roles. It is argued that agents should support their perception as social actors by displaying human-like social cues such as affect and gestures. An architecture for emotion-based agents will be described and a simplified version of the model will be illustrated by two interaction scenarios that feature cartoon-style characters and can be run in a web browser. The second focus of this paper is an empirical evaluation of the effect of an affective agent on users' emotional state which is derived from physiological signals of the user. Our findings suggest that an agent with affective behavior may significantly decrease user frustration.

***key words:*** *Human-like and believable qualities of agents, emotion and personality, agent modeling, empirical methods to observe user behavior, web-based environments.*

## 1. Introduction

Recent years have witnessed a growing interest in employing animated agents for tasks that are typically performed by humans. To mention some of the more prominent applications in progress, embodied characters are used

- as virtual *tutors* in interactive learning environments [10], [11], [15], [17],
- as virtual *sales persona* on the web and at information booths [1], [3], [7], [18],
- as virtual *presenters* [2], [16],
- as virtual *communication partners* in therapy [23],
- as virtual *actors* for entertainment [29], [33],
- as virtual *personal representatives* in online communities and guidance systems [9], [34], [35], and
- as virtual *information experts* enhancing conventional web search engines [19].

The main effort in designing animated agents is to make them *life-like* or *believable* [4]. Life-likeness is intended to provide the viewer with the illusion of life, and similarly, believability should allow viewers to suspend their disbelief. Life-like characters are intended to communicate like real people and be able to engage naturally in conversation and other tasks with humans.

In this paper, we will discuss the premises under which synthetic characters can be pedagogically effective as tutors, convincing as presenters, and dramatically interesting as actors, among other roles. Rather than describing application-specific requirements for agents, we will focus on features that are common to different kinds of animated interface agents. We will argue that the key premise for successful agent applications is to design animated agents as *social actors* that display (and possibly recognize) social cues such as a certain (social) role, affective and socially appropriate verbal response, and non-verbal behavior including the 'embodied' expression of emotions and turn taking. In order to support our assumption that agents as social actors have a positive impact on users, an experiment has been carried out where an animated agent displays affective behavior (emotion and empathy) in the context of a simple quiz game.

The rest of the paper is organized as follows. In Section 2, we will motivate the design requirements of agents as social actors. Section 3 describes the main components of the SCREAM system, an architecture for emotion-based agents. In Section 4, agents' embodiment is discussed by describing its communicative functions. Section 5 illustrates our model by means of two web-based interaction scenarios that feature animated agents. Section 6 describes an empirical study that shows that an agent's affective feedback may reduce user frustration. In Section 7, we summarize and conclude the paper.

## 2. Background

A rather surprising finding about the interaction between humans and computers is that humans are already strongly biased to interpret the behavior of synthetic entities that display some human-like features as life-like and believable. This was shown by Reeves and Nass [32] who carried out a series of classical tests of human-human social interaction, but replaced one interlocutor by a computer with human-sounding voice and a particular role such as companion or opponent. The results of those experiments suggest that humans treat computers in an essentially natural way – as so-

Manuscript received August 30, 2002.
Manuscript revised January 10, 2003.

†The authors are with the Department of Information and Communication Eng., Graduate School of Information Science and Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

cial actors – with a tendency, e.g., to be nicer in 'face-to-face' interactions than in third party conversations. More support for this result is provided by Lester *et al.* [22] who investigate the impact of animated agents along the dimensions of motivation and helpfulness, and coin the term 'persona effect' "[...] which is that the presence of a lifelike character in an interactive learning environment – even one that is not expressive – can have a strong positive effect on student's perception of their learning experience." [22, p. 359].

The human interpretation process, however, is very sensitive to and easily disturbed by a character's 'inconsistent' or 'unnatural' behavior, whatever type of 'nature' (realistic or not) is applicable. The challenge here is to maintain consistency between an agent's internal emotional state and various forms of associated outward behavior such as speech and body movements (Gratch *et al.* [14]). An agent that speaks with a cheerful voice without co-occurring happy facial expression will seem awkward or even fake. Another challenge is to keep consistency of agents over time, allowing for changes in the their response tendencies as a result of the interaction history with other agents (Prendinger and Ishizuka [29], [30]).

While the persona effect suggests that animated agents add to a user's engagement and entertainment when interacting with a character-based interface, an empirical study performed by van Mulken *et al.* [25] showed that interface agents have no effect on ('objective') measures such as comprehension and recall. A commonly observed problem of human–agent interaction, however, is that users experience negative emotions such as frustration and stress. Recent studies in the 'Affective Computing' paradigm address the question whether providing appropriate affective feedback may alleviate the intensity of user frustration (Picard [28]). Klein *et al.* [20] could show that computers may undo some of a user's negative emotions by providing affective response in the form of active listening, empathy, and sympathy. Rather than using a text-based interface to interact with the user as in [20], we will describe a study that employs an affective animated agent to communicate with the user.

## 3. Emotion, Personality, Social Role

We have developed a system called SCREAM (SCRipting Emotion-based Agent Minds) that facilitates scripting a character's affect-related processing capabilities. The system allows to specify a character's mental make-up and endow it with emotion and personality which are considered as key features for the life-likeness of characters. A character's mental state can be scripted at many levels of detail, from driven purely by (personality) traits to having full awareness of the social interaction situation, including character-specific beliefs and beliefs attributed to interacting characters or even the

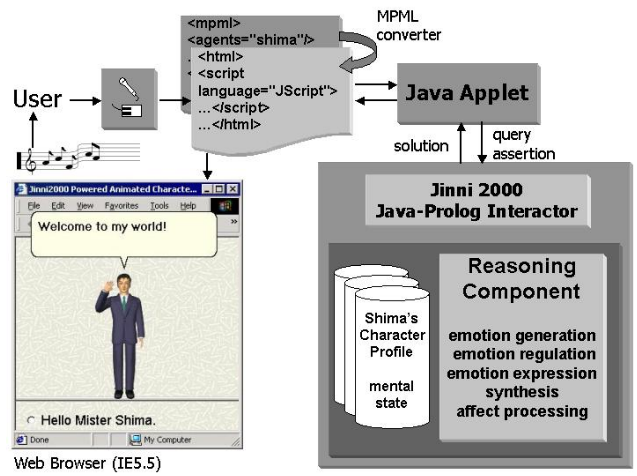


**Fig. 1** Implemented System Overview.

user. For portability and extensibility, the SCREAM system is written in Java and Jinni 2000, a Java based Prolog system [5].

The Microsoft Agent package [24] is used to embed animated characters into a web page based JavaScript interface. The package comes ready with controls for animating 2D cartoon-style characters, speech recognition and a Text-to-Speech (TTS) engine. In order to facilitate the process of scripting more complex scenarios, including, e.g., sequential and parallel embodied behavior of multiple characters, we have developed an XML-style markup language called MPML (Ishizuka *et al.* [16]). MPML provides an interface to SCREAM and thus supports both behavior autonomously generated (by SCREAM) and pre-scripted agent behavior (specified in MPML). Fig. 1 gives an overview of the implemented system. The following paragraphs offer a walk through the main components of the SCREAM system (see [29] for a detailed description): Emotion Generation, Emotion Regulation, Emotion Expression, and the Agent Model that is also responsible for updating a character's mental state.

A core activity of an emotion-based agent mind is Emotion Generation and the management of emotions, which is dealt with by three modules, the *appraisal* module, the *emotion resolution* module, and the *emotion maintenance* module. Reasoning about emotion models an agent's appraisal process, where events are evaluated as to their emotional significance for the agent (Ortony *et al.* [27]). The significance is determined by so-called 'emotion-eliciting conditions', the agent's beliefs, goals, standards, and attitudes. Following the emotion model in [27] (the OCC model), we conceive emotion types as classes of eliciting conditions, each of which is labelled with an emotion word of phrase. In total, twenty-two classes of eliciting conditions are identified: joy, distress, happy for, sorry for, resent, angry at, and so on. All emotions have associated intensities depending on the intensities of its conditions, e.g., the intensity of an agent's attitude toward its interlocutor. Since a reasonably interesting agent will have a multitude of mental states (beliefs, goals, attitudes, etc.), more than one emotion is typically triggered when the agent interacts with another agent. The emotion resolution and maintenance modules determine the most dominant (winning) emotion and handle the decay process of emotions, respectively.

The expression of emotions is governed by social and cultural norms that have a significant impact of the intensity of emotion expression. We will treat Emotion Regulation as a process that decides whether an emotion is expressed or suppressed. We categorize regulatory parameters into ones that constitute a social threat for the agent (social role, distance, power), and parameters that refer to the agent's capability of (self-)control (personality, interlocutor personality, linguistic style). Our personality model considers just two dimensions, which seem crucial for social interaction. Extroversion refers to an agent's tendency to take action (talkative), whereas agreeableness refers to an agent's disposition to be sympathetic (friendly). An overall control value is computed based on the given (possibly mutually defeating) control values. E.g., the control of an agent that is very extrovert but deals with a very unfriendly interlocutor might be neutralized to some degree.

The Agent Model describes an agent's mental state. We distinguish static and dynamic features of an agent's mind state, such that the agent's personality and standards are considered as static whereas goals, beliefs, attitudes and social variables are considered as dynamic. Here, we are mainly concerned with change of attitude as a result of social interaction. Ortony [26] suggests the notion of *(signed) summary record* to capture the dispositional (dis)liking of another person. This record stores the sign of emotions that were induced in the agent $L$ by an interlocutor $I$ together with emotions' associated intensities. The sign is either positive (e.g., for 'joy', 'happy for') or negative (e.g., for 'distress', 'angry at'). In order to compute the current intensity of an agent's (dis)liking $\delta^\sigma$, we simply compare the (scaled) sum of intensities of elicited positive and negative emotions ($\delta^\sigma$, $\delta \in \{0, \ldots, 5\}$, $\sigma \in \{+, -\}$), starting in the initial situation $Sit_1^{L,I}$.

$$\delta^\sigma(Sit_n^{L,I}) = \frac{\sum_{i=1}^{n} \delta_w^\sigma(Sit_i^{L,I})}{n}$$

Here, we only consider the intensity of the winning emotional state $\delta_w$. If no emotion of one sign is elicited in a situation, it is set to zero. Positive values for the difference $\delta^+ - \delta^-$ indicate an agent's liking of an interlocutor and negative ones indicate disliking. The more interesting case where an interlocutor the agent likes as a consequence of consistent reinforcement (suddenly) induces a high-intensity emotion of the opposite sign, e.g., by making the agent very angry, is captured by a special update rule that weights historical and recent (affective) information to obtain the current attitude.

An agent's emotions are expressed by its bodily behaviors. Embodiment has also important functions to guide communication, which will be described in the following section.

## 4. Embodiment

When humans communicate, they employ a variety of signals in combination with verbal utterances, such as body posture, gestures, facial expressions, and gaze. In a similar way, animated agents may use their bodies to convey meaning and regulate communication. The most extensive study of non-verbal behaviors for synthetic characters, especially gestures, can be found in Cassell's work on embodied conversational agents [8].

*Emblematic* gestures are culturally specified gestures, e.g., signalling "okay" by a "thumb-and-index-finger" ring gesture. An example of a *propositional* gesture is the use of both hands to measure the size of an object in symbolic space while saying "there is a big difference" (see Fig. 2(a)).

There are four types of gestures that support the conveyance of communicative intent (so-called 'co-verbal' gestures [8]): (i) *Iconic* gestures illustrate some feature of an object or action, e.g., mimicking to hold a phone while saying that someone has been called; (ii) *Metaphoric* gestures represent a concept without physical form, e.g., a rolling hand gesture while saying "let's go on now"; (iii) *Deictic* gestures locate physical space relative to the speaker, e.g., by pointing to an object; (iv) *Beat* gestures are small baton-like movements to emphasize speech. A special form of a beat gesture is the *contrastive* gesture that depicts a 'on the one hand ... on the other hand' relationship if two items are being contrasted (see Fig. 2(b)).

An important class of gestures (including facial gestures) serves the expression of an agent's *emotional state* such as 'hanging shoulders' to signal sadness (see Fig. 2(c)). Although face may express emotions most succinctly (Ekman and Friesen [12]), we rather rely on signals involving the whole body as the size of the characters used is relatively small.

Gestures also realize *communicative functions* including conversation initiation, turn taking, back channelling ("nodding"), and breaking away from conversation [7]. The communicative behavior corresponding to the function of "giving turn" is typically realized by looking at the interlocutor with raised eyebrows, followed by silence, whereas "taking turn" is signalled by glancing away and starting to talk (see Fig. 2(d)).

Agent behaviors in our system are 2D animations controlled by MPML and have to be specified at the signal level for each given communicative situation. APML (Affective Presentation Markup Language), on the other hand, allows to specify the meaning associ-

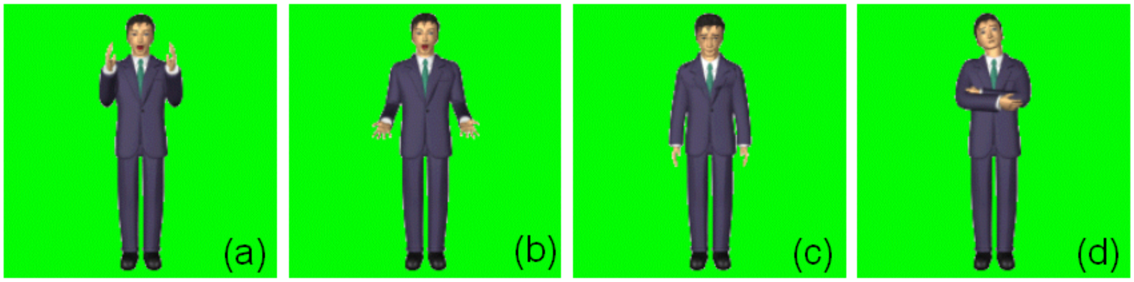


**Fig. 2** "Big difference" (a), "Contrast" (b), "Sadness" (c), and "Take turn" (d).

ated with a communicative act that may be instantiated by different signals depending on the agent's personality, gesture style, or culture (de Carolis *et al.* [6]). However, since the signals (here, animation sequences) provided for our characters are fixed, a direct scripting approach seems appropriate.

## 5. Illustrative Examples

In this section, we will describe two web-based scenarios that instantiate (simplified versions of) emotion-based agents. We will use a technique called *story-morphing* to achieve a simple form of narrative. Story-morphing relies on a given fixed base plot structure ('script') that allows to generate numerous distinct stories (*story-morphs*) by varying the affective models of the involved characters (Elliott *et al.* [13]). Most importantly, the user may change the evolution of events by interacting with the characters by giving speech commands.

### 5.1 Playing Black Jack

In the casino scenario, the user and other characters can play the "Black Jack" game. Fig. 3 shows the situation where the character "Genie" in the role of an advisor practices Black Jack with the user by commenting the game of character "Al" (Genie is the character at the bottom-left of the Internet Explorer window, and Al is the male character to the right of the dealer). In an interaction session, the user is guided by the advisor Genie to play the Black Jack game whereby Genie's reactions vary according to his goals and personality profile as well as the user's decisions ("hit", "stand") and the outcome of the game.

Different story-morphs are obtained by different interaction patterns, i.e., decisions of the user. They result from the evolution of Genie's attitude toward the user, depending on whether the user follows or refuses to follow Genie's advice. In effect, Genie's attitude decides, e.g., whether he is sorry for or resents the user's lost game. Given the user either "hits" or "stands" in each round, there exist $2^n$ story-morphs ($n$ the number of rounds) for each setting of Genie's character profile.

For expository reasons, we let the user never follow Genie's advice, and we use the following character profile. Among others, Genie is assumed as rather agreeable and extrovert, he is socially close to the user and also (initially) slightly likes the user. His goals are that the user wins (with low intensity), and that the user follows his advice (with high intensity). Note that the outcome of the the game, i.e., whether the user wins or looses, is independent of her or him following Genie's advice. Let us watch the user playing the beginning three rounds of Black Jack.

**First round** (user ignores advise and looses). Genie's winning emotional state is 'distress' with high intensity, because the user did not follow his advice. He displays his distress with low intensity as his agreeable personality effects a decrease in the intensity of negative emotion expression.

**Second round** (user ignores advice and looses). Genie is sorry for the user with high intensity, since positive ('sorry for' the user's lost game) emotions decay slowly and sum up. However, Genie's personality traits let him express the emotion with even higher intensity.

**Third round** (user ignores advice and looses). Genie gloats over the user's lost game with high intensity, because at that point, the negative emotions dominate the positive ones as a consequence of the user's repeated refusal to follow Genie's advice and his attitude changes to slightly disliking the user which lets him experience joy over the user's distress. Again, Genie's friendly personality is responsible for the decreased intensity of the external emotion.

By way of example, in the third round Genie opens his arms widely while saying "Ha! Want to decide by yourself? Never listen to my advice. You got the result." He thereby expresses his joy over the user's (assumed) distress ('gloating').

Although we did not evaluate the believability of the advisor in a systematic way, informal conversations with users suggest that Genie's behavior is understood as natural, whereas simple "user wins"–"Genie is happy for user" and "user looses"–"Genie is sorry for user" reactions are conceived as odd. Those preliminary experiences are in line with a previously performed study, where the feature of 'social role awareness' was empir-

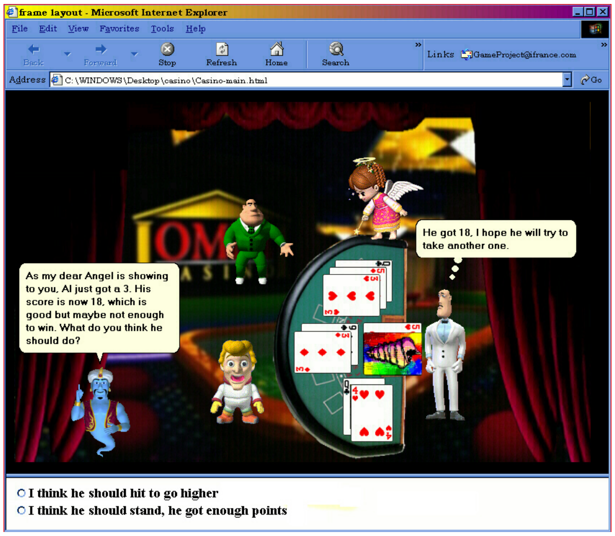


**Fig. 3** Playing Black Jack.

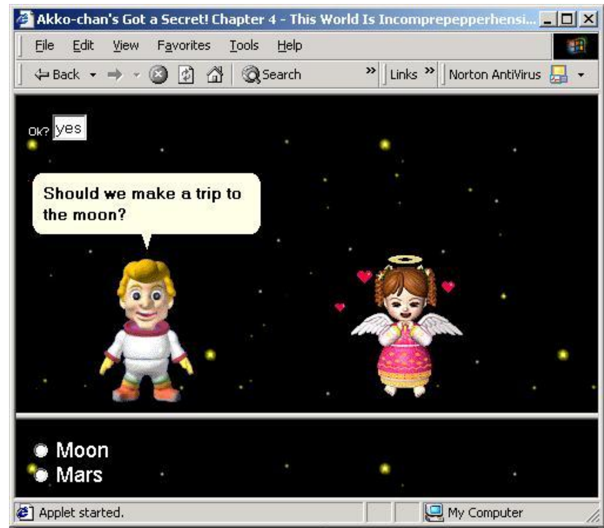


**Fig. 4** Interactive Japanese Comics.

ically tested (Prendinger and Ishizuka [31]). It could be shown that agents that express their emotions according to the social role (defined by social power and distance relations) they hold in a particular communication context are conceived as significantly more believable (or natural) than agents that are only emotion and personality driven.

## 5.2 Interactive Japanese Comics

Borrowing the idea from Fujio Akatsuka's manga series "Akko-chan's Got a Secret!", a character called "Little Akko" plays the heroine of stories for kids. Little Akko has the power to be transformed into any person upon telling her wish to a magic mirror. By this magic, she can make other people happy. Fig. 4 shows her transformed into Little Chika, a girl whom her brother Kankichi likes.

The purpose of interacting with Little Akko is to elicit her attraction emotion, which has a certain entertainment value. The interaction setting is fairly simple. The user can communicate with Little Akko by controlling an avatar, the "Space-boy" character in the role of Kankichi. By offering Little Akko items she likes, the user may increase her positive attitude and familiarity. Attitude and familiarity may change based on a small set of emotion types: joy, distress, attraction, aversion. Note that attraction and aversion are dependent on the agent's attitude and familiarity relations toward its interlocutor. Familiarity change is simply modeled by incrementing the familiarity value by a small (intensity) amount when a positive emotion (joy) is elicited.

Consider the conversation following Little Akko's statement that she likes sweet things, which is depicted in Fig. 5. When being offered a grapefruit, Little Akko shows her distress since she likes sweet things. However, as the user thereafter repeatedly selects items Little

User may select "Chocolate cake" or "Grapefruit"
Kankichi: Would you like to eat a grapefruit?
Little Akko: [*startling*] No! I do not like that. Anyway, let us now play an interesting game.

User may select "Calculate" or "Hide and Seek"
Kankichi: Do you want to play the Calculate Game?
Little Akko: [*joyful facial display*] I really like that game! I got hungry. I like rice better than bread.

User may select "Sushi" or "Sandwiches"
Kankichi: Would you like to eat some Sushi?
Little Akko: [*joyful facial display*] Yes! That is what I like! What about a trip to a beautiful place?

User may select "Moon" or "Mars"
Kankichi: Should we make a trip to the moon?
Little Akko: [*flying hearts*] I enjoy being with you!

**Fig. 5** Interacting with Little Akko.

Akko likes, there are two kinds of effects. Both Little Akko's liking value toward the Kankichi and the familiarity level increase, and hence add to the intensity of the Little Akko's 'attraction' toward Kankichi. After Kankichi offers the Little Akko a trip to the moon, her emotional state comprises two active emotions, 'joy' and 'attraction' (both medium intensity), and she expresses the combined emotion with higher intensity (see the 'flying hearts' animation in Fig. 4).

Arguably, our implementation of an episode of this comics book is simple, where the social relationship evolves in a quick and direct way. In our experience, users easily understand Little Akko's wishes by her (exaggerated) emotional reactions, and are eager to gain her sympathy. Story-morphs here are generated by the user's direction of Kankichi and the nature and intensity level of Little Akko's wishes.

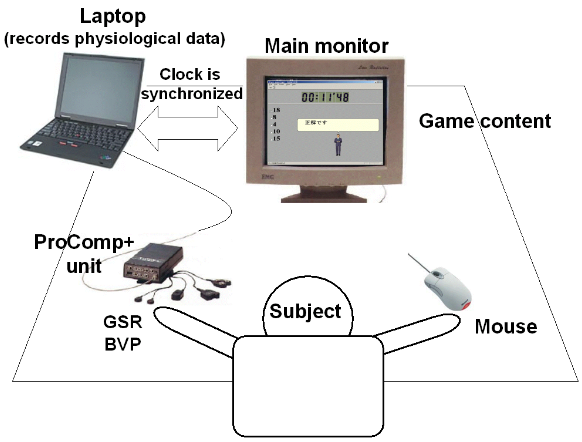


**Fig. 6** Experimental setup.

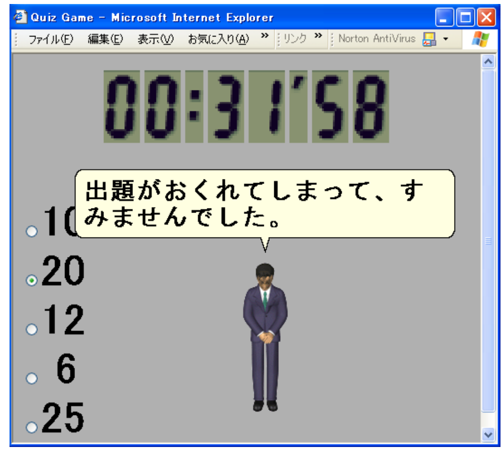


**Fig. 7** The agent apologizes for the delay.

## 6. Experimental Study

The aim of the experimental study described in this section is to show the effect of a character-based interface on users' emotional state. In order to measure the impact of an agent with affective response on users, we take their physiological signals skin conductivity and blood volume pressure. Unlike standard evaluation methods such as questionnaires, the use of physiological data allows for a more precise assessment of humans' perception of animated agents as social actors.

### 6.1 Design and Setup of the Experiment

We implemented a simple quiz game[†], where subjects are instructed to sum up five consequently displayed numbers and are then asked to subtract the $i$th number of the sequence, where $i \leq 4$. Subjects compete for the best score in terms of correct answers and time, which is awarded by a JPY 5000.- prize, besides JPY 1000.- for participation. An introductory explanation of the game given by an animated agent includes the announcement that subjects will not necessarily interact with a bug-free prototype interface. This period also serves to collect physiological data of subjects that are needed to normalize data obtained during game play. In six out of a total of thirty questions, we inserted a delay before showing the 5th number. The delay is expected to induce frustration as the subjects have to memorize the order of numbers (presumably together with its sum) for a period between 6 and 14 seconds (average 9 sec.), and their goals of giving the correct answer and of achieving a fast score are thwarted (also called 'primary frustration' in behavioral psychology).

Subjects were twenty male students, and all of them are native speakers of Japanese. Subjects were randomly assigned to one of two versions of the game (ten in each version). In the *affective version* the character expresses 'happiness' or 'sorriness' verbally and non-verbally (e.g., happy or sad facial display, hanging shoulders), depending on whether the user selects the correct (wrong) answer from a menu displayed in the game window. If a delay happens, the character also expresses empathy for the user, e.g., by saying "I apologize that there was a delay in posing the question" (English translation) after the subject gives the right/wrong answer to the quiz question. As can be seen from Fig. 7, the agent displays a gesture that Japanese people perceive as a signal of the interlocutor's apology. Here, we follow the characterization of "empathy" given in Klein *et al.* [20, p. 122]: "this is how *I* feel *about* what you're going through. Here is *my emotional response* to *your* predicament." In the *non-affective version*, on the other hand, the character does not give any affective feedback, and simply states "right" or "wrong". Furthermore, the agent does not comment on the occurrence of the delay.

Subjects are attached to two types of sensors, SC (skin conductivity) and BVP (Blood Volume Pressure) on the first three fingers of their non-dominant hand (see Fig. 6). Since SC covaries with the level of arousal, and heart rate (calculated from BVP) with negative valence of emotion, the signals can be used to infer user emotions as a location in the valence–arousal space of emotion (Lang [21], Picard [28]). Physiological signals are recorded via the ProComp+ unit and visualized using Thought Technology software.

In order to show the effect of the agent's behavior, we have been interested in two particular segments. The DELAY segment refers to the period after which the agents suddenly stops activity while the question is not completed – presumably causing user frustration – until the moment when the agent continues with the question. The RESPONSE segment refers to the agent's response to the subject's answer.

[†]The experiment was designed in close cooperation with Junichiro Mori (Univ. of Tokyo).

### 6.2 Results

The first result relates to the use of delay in order to induce frustration in subjects. All eighteen subjects showed a significant rise of SC in the DELAY segment, indicating an increased level of arousal (the data of two subjects of the non-affective version were discarded because of extremely deviant values). Due to technical problems, the BVP data of only six subjects could be taken. However, those data did not show any relevant changes in subjects' heart rate.

Our main hypothesis is that an animated agent with affective behavior may reduce user frustration. To support this hypothesis, we first calculated the mean values of SC for each subject considering only the six delay game situations. Then we computed the difference between the DELAY and RESPONSE segments on the mean values of those signals. In the following, the confidence level $\alpha$ is set to 0.05 (i.e., we accept a risk of 0.05 of making a false decision as to reject or retain the null hypothesis). In the non-affective version, the difference is even negative (mean= $-0.08$). However, in the affective version SC decreases when the character responds to the user (mean= 0.14). The $t$-test (two-tailed, assuming unequal variances) showed a significant effect of the character's affective behavior as opposed to non-affective behavior ($t(16) = -2.47$; $p = 0.025$). This result suggests that a character expressing emotions and empathy may undo some of the frustration caused by a deficiency of the interface.

Although more results can be extracted by a more detailed evaluation of the data, the described result is the most relevant to support the general hypothesis that animated agents with affective behavior are perceived as social actors, and as such they may alleviate frustration, similar to human interlocutors.

The use of physiological data has the promise to be an accurate method to assess user emotions. In a follow-up experiment, we intend to consider signals other than skin conductance and heart rate. In order to estimate subjects' perception of a life-like character more conclusively, we also plan to run the same experiment with a text-based interface showing 'affective' text messages rather than 2D graphics and speech.

## 7. Conclusion

Recent years have seen many efforts to include life-like, believable agents as a crucial component of enhanced learning, presentation, and entertainment systems. This paper discusses design desiderata for synthetic characters that enable users to conceive them as veritable communicative partners or simply, as social actors. As two core features of animated agents as social actors we describe agents with an emotion-based artificial mind and embodied conversational behavior.

The believability and life-likeness of agents is shown from two perspectives. First, animated agents are endowed with believability enhancing features such as emotion and personality, as well as social role dependent and adaptive behavior, which is illustrated by two interaction scenarios from the entertainment genre. However, those scenarios do not lend themselves easily to a direct evaluation. Therefore, we discuss agents from a second perspective, by performing an experiment where users are frustrated on purpose and get feedback from an agent displaying affective behavior in the form of expressing his sorriness. Subjects' physiological data strongly suggest that an agent as social actor may decrease the (induced) frustration. In the future, we hope to implement and evaluate a variety of affective response strategies for animated agents.

## Acknowledgement

We would like to thank the anonymous referees for their very helpful and valuable comments. Special thanks go to Junichiro Mori (Univ. of Tokyo) who carried out the experiment and performed the data analysis. This research is supported by the JSPS Research Grant (1999-2003) for the Future Program ("Mirai Kaitaku").

### References

[1] E. André, T. Rist, S. van Mulken, M. Klesen, and S. Baldes. The automated design of believable dialogue for animated presentation teams. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors, *Embodied Conversational Agents*, pages 220–255. The MIT Press, Cambridge, MA, 2000.

[2] N.I. Badler, J. Allbeck, R. Bindiganavale, W. Schuler, L. Zhao, and M. Palmer. Parameterized action representation for virtual human agents. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors, *Embodied Conversational Agents*, pages 256–284. The MIT Press, Cambridge, MA, 2000.

[3] S. Baldes, P. Gebhard, M. Kipp, M. Klesen, P. Rist, T. Rist, and M. Schmitt. The interactive CrossTalk installation: Meta-theater with animated presentation agents. In H. Prendinger, editor, *Proceedings PRICAI-02 International Workshop on Lifelike Animated Agents. Tools, Affective Functions, and Applications*, pages 9–15, 2002.

[4] J. Bates. The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125, 1994.

[5] BinNet Corp. *Jinni 2000: A high performance Java based Prolog for agent scripting, client-server and internet programming*, 2000. URL: `www.binnetcorp.com`.

[6] B. De Carolis, V. Carofiglio, M. Bilvi, and C. Pelachaud. APML, a mark-up language for believable behavior generation. In *Proceedings AAMAS-02 Workshop on Embodied conversational agents—let's specify and evaluate them!*, 2002.

[7] J. Cassell. More than just another pretty face: Embodied conversational interface agents. *Communications of the ACM*, 43(4):70–78, 2000.

[8] J. Cassell. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, editors, *Embodied Conversational Agents*, pages 1–27. The

MIT Press, Cambridge, MA, 2000.

[9] J. Cassell and H. Vilhjálmsson. Fully embodied conversational avatars: Making communicative behaviors autonomous. *Autonomous Agents and Multi-Agent Systems*, 2:45–64, 1999.

[10] C. Conati. Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence*, 16:555–575, 2002.

[11] F. de Rosis, B. de Carolis, and S. Pizzulito. Software documentation with animated agents. In *Proceedings 5th ERCIM Workshop on User Interfaces For All*, 1999.

[12] P. Ekman and W.V. Friesen. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1:49–98, 1969.

[13] C. Elliott, J. Brzezinski, S. Sheth, and R. Salvatoriello. *Story-morphing* in the Affective Reasoning paradigm: Generating stories semi-automatically for use with *emotionally intelligent* multimedia agents. In *Proceedings 2nd International Conference on Autonomous Agents (Agents-98)*, pages 181–188, New York, 1998. ACM Press.

[14] J. Gratch, J. Rickel, E. André, J. Cassell, E. Petajan, and N. Badler. Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems*, pages 54–63, 2002.

[15] B. Hayes-Roth. Adaptive learning guides. In *Proceedings of the IASTED Conference on Computers and Advanced Technology in Education*, 2001.

[16] M. Ishizuka, T. Tsutsui, S. Saeyor, H. Dohi, Y. Zong, and H. Prendinger. MPML: A multimodal presentation markup language with character control functions. In *Proceedings Agents'2000 Workshop on Achieving Human-like Behavior in Interactive Animated Agents*, pages 50–54, 2000.

[17] W.L. Johnson, J.W. Rickel, and J.C. Lester. Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11:47–78, 2000.

[18] Y. Kitamura, T. Sakamoto, and S. Tatsumi. A competitive information recommendation system and its behavior. In *Proceedings Sixth International Workshop on Cooperative Information Agents VI (CIA-02)*, pages 138–151, Berlin, 2002. Springer Verlag.

[19] Y. Kitamura, H. Tsujimoto, T. Yamada, and T. Yamamoto. Multiple character-agents interface: An information integration platform where multiple agents and human user collaborate. In *Proceedings First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02)*, pages 790–791, New York, 2002. ACM Press.

[20] J. Klein, Y. Moon, and R.W. Picard. This computer responds to user frustration: Theory, design, and results. *Interacting with Computers*, 14:119–140, 2002.

[21] P.J. Lang. The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5):372–385, 1995.

[22] J.C. Lester, S.A. Converse, S.E. Kahler, S.T. Barlow, B.A. Stone, and R.S. Bhogal. The Persona effect: Affective impact of animated pedagogical agents. In *Proceedings of CHI-97*, pages 359–366, 1997.

[23] S.C. Marsella, W.L. Johnson, and C. LaBore. Interactive pedagogical drama. In *Proceedings 4th International Conference on Autonomous Agents (Agents-2000)*, pages 301–308, New York, 2000. ACM Press.

[24] Microsoft. *Developing for Microsoft Agent*. Microsoft Press, 1998.

[25] S. Van Mulken, E. André, and J. Müller. The Persona Effect: How substantial is it? In *Proceedings Human Computer Interaction (HCI-98)*, pages 53–66, 1998.

[26] A. Ortony. Value and emotion. In W. Kessen, A. Ortony, and F. Craik, editors, *Memories, thoughts, and emotions: Essays in the honor of George Mandler*, pages 337–353. Hillsdale, NJ: Erlbaum, 1991.

[27] A. Ortony, G.L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge, 1988.

[28] R.W. Picard. *Affective Computing*. The MIT Press, 1997.

[29] H. Prendinger, S. Descamps, and M. Ishizuka. Scripting affective communication with life-like characters in web-based interaction systems. *Applied Artificial Intelligence*, 16(7–8):519–553, 2002.

[30] H. Prendinger and M. Ishizuka. Evolving social relationships with animate characters. In *Proceedings of the AISB-02 Symposium on Animating Expressive Characters for Social Interactions*, pages 73–78, 2002.

[31] H. Prendinger and M. Ishizuka. Let's talk! Socially intelligent agents for language conversation training. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 31(5):465–471, 2001.

[32] B. Reeves and C. Nass. *The Media Equation. How People Treat Computers, Television and New Media Like Real People and Places*. CSLI Publications. Cambridge University Press, 1998.

[33] D. Rousseau and B. Hayes-Roth. A social-psychological model for synthetic actors. In *Proceedings 2nd International Conference on Autonomous Agents (Agents-98)*, pages 165–172, New York, 1998. ACM Press.

[34] Y. Sumi and K. Mase. AgentSalon: Facilitating face-to-face knowledge exchange through conversations among personal agents. In *Proceedings 5th International Conference on Autonomous Agents (Agents-01)*, pages 393–400, New York, 2001. ACM Press.

[35] T. Takahashi, H. Takeda, and Y. Katagiri. Script language for embodied agents as personal conversational media in online communities. In *Proceedings AAMAS-02 Workshop on Embodied Conversational Agents: Let's Specify and Compare Them!*, 2002.

**Helmut Prendinger** is a research associate at the Dept. of Information and Communication Eng., Graduate School of Information Science and Technology, Univ. of Tokyo. Earlier he worked as a Junior Specialist at the Univ. of California, Irvine. He received his M.A. and Ph.D. degrees from the Univ. of Salzburg, Dept. of Logic and Philosophy of Science and Dept. of Computer Science. His research interests include artificial intelligence and human-computer interaction, in which areas he has published more than 50 papers in international journals and conferences.

**Mitsuru Ishizuka** is a professor at the Dept. of Information and Communication Eng., Graduate School of Information Science and Technology, Univ. of Tokyo. Previously, he worked at NTT Yokosuka Laboratory and Institute of Industrial Science, Univ. of Tokyo. During 1980–81, he was a Visiting Assoc. Professor at Purdue University. He received his B.S., M.S. and Ph.D. degrees in electronic engineering from the Univ. of Tokyo. His research interests are in the areas of artificial intelligence, multimodal agent interface/contents and intelligent WWW information space. He is a member of IEEE, AAAI, IEICE Japan, IPS Japan and Japanese Society for AI.