The use of affective and attentive cues in an empathic computer-based companion

Nikolaus Bee[1], Elisabeth André[1], Thurid Vogt[1] and Patrick Gebhard[2]

[1]Institute of Computer Science, Augsburg University,Germany
[2]DFKI GmbH, GmbH, Germany

ABSTRACT

Recently, a number of research projects have been started to create virtual agents that do not just serve as assistants to which tasks may be delegated, but that may even take on the role of a companion. Such agents require a great deal of social intelligence, such as the ability to detect the user's affective state and to respond to it in an empathic manner. The objective of our work is to create an empathetic listener that is capable to react on affective and attentive input cues of the user. In particular, we discuss various forms of empathy and how they may be realized based on these cues.

INTRODUCTION

Listening is an ingredient component of any dialogue, and to make your counterpart feel understood and cared for, you need the capabilities of an empathetic listener that appears attentive and sensitive towards the emotional state of others. Consider, for example, a situation where your best friend fails an exam that was very important for him as it decides about graduating. When your friend tells you about his situation, you may show different kinds of empathetic reaction. In the easiest case, you show an emotional expression that is similar to that of your friend. Your friend will usually interpret your behavior as a sign that you share his feelings and might finally feel better, true the motto: a problem shared is a problem halved. Instead of simply mirroring your friend's emotional state, you may also show an emotional state as response to your friend's emotional state that is related to that of your friend, but not necessarily the same. For example, if your friend looks sad, you might show him by a compassionate facial expression that you are sorry about his sadness. Responding to somebody's emotional state by conveying a so-called good-will emotion, such as sorry-for or happy-for, is called reactive empathy. Another possibility for you to react is trying to understand the current situation and thinking about possibilities to make your friend feel better. For example, you might encourage your friend by telling him that you also failed some exams but there is always a second chance. We call this form of empathy deliberate empathy. Deliberate empathy requires a deeper level of affective mind reading than parallel or reactive empathy. In order to comfort your friend, you need to be able to reason about his affective appraisal processes. In cognitive science, this ability is called affective theory of mind.

Recently, a number of research projects have been started to create virtual agents that do not just serve as assistants to which tasks may be delegated, but that may even take on the role of a companion. Such agents require a great deal of social intelligence, such as the ability to detect the user's affective state and to respond to it in an empathic manner. While

a lot of work has been done on the creation of affective output for virtual characters, less work was done on the smooth integration of recognition and display processes. The objective of our work is to create an empathetic listener that is capable to react on affective and attentive input cues of the user. Such an agent has to appear sensitive by appropriately responding to the user's affective state. Furthermore, it has to appear attentive by establishing frequent eye contact with the user without being obtrusive.

RELATED WORK

Prendinger and colleagues (Prendinger, Dohi, Wang, Mayer, & Ishizuka, 2004) developed an empathic companion that accompanies a user in a virtual job interview. This system measures the users' physiological state (skin conductance and electromyography) in real-time and interprets it as emotion. The virtual agent reacts with empathic feedback on the users' current affective state. The reaction is calculated using a Bayesian net, which takes the physiological state and the user's answers to the interview into account. Gratch et al. (Gratch, Wang, Gerten, Fast, & Duffy, 2007) describe a system to establish rapport in human-machine dialogs. They detect speech and head orientation from the user to create a responsive dialogue agent. Speech detection is used for detecting backchannel opportunity points, disfluencies, questions and loudness and the head tracker detects head nods, shakes, gaze shifts and posture shifts. Their system is mainly meant to create contingent nonverbal behavior and neither takes the user's emotional state into account nor intends to react with affective behavior. McQuiggan et al. (McQuiggan, Robison, Phillips, & Lester, 2008) created a system for empathic virtual agents. Using a machine learning approach, they automatically determine appropriate parallel and reactive empathy behaviors. The system is trained by the users' in-game behavior. Thus, it does not consider the real users' actual emotional state. Ochs et al. (Ochs, Pelachaud, & Sadek, 2008) created an empathic model for virtual agents by combining a corpus-based approach with theoretical work on emotions. Their model was applied to an agent that provides users with information on their email. The system infers the user's emotional state based on an analysis of the dialogue and computes an empathetic reaction to it. Boukricha (Boukricha, 2008) proposes a computational model for empathic virtual agents that distinguishes between sharing the emotions of others (mimicry) and understanding the emotions of others cognitively (role-taking). Detection of the users' current emotional state, which will be mapped to a virtual character, is planned by using facial feature detection. There is an ongoing project named SEMAINE (Schröder, Cowie, Heylen, Pantic, Pelachaud, & Schuller, 2008) that deals with building Sensitive Artificial Listeners, which are designed to sustain the communication with a person. Their system will recognize and generate non-verbal behavior in real-time.

EXAMPLE

Our objective is the creation of an empathic listening agent that analyzes the user's emotive state and responds to it in real-time. To realize such an agent, we have been experimenting with several metaphors, such as that of a virtual pet or that of a virtual butler. While the virtual pet is not able to verbally respond to the user's state, the virtual butler provides both verbal as well as non-verbal feedback. Imagine the user has a rather bad day and is talking to the agent with a gloomy voice. In the case of parallel empathy, the agent would simply mimic the user's emotive state and show gloom as well (see Figure 1 middle). In contrast to that, reactive empathy requires the agent to decide which emotion to display as a response to the

user's emotive state. Here, the agent's emotion does not necessarily coincide with the user's emotion. That is the agent might, for example, decide to show "pity for" (see Figure 1 right) as for example suggested by (Pelachaud & Poggi, 2002) with the head aside and the inner eyebrow raised or to sheer the user up by showing encouragement (see Figure 1 left). In addition, the agent might provide simple verbal feedback, such as "I am feeling bad as well." (in the case of parallel empathy) or "I am very sorry you are feeling so bad." (in the case of reactive empathy).

Figure 1: The virtual character Alfred is designed utilizing FACS to compose facial expressions (left: *joyful*, middle: *gloomy* and right: *pity-for*).

AFFECTIVE INTERACTION

Our architecture provides components to sense affective and attentive states from the user, a component to process affective states for parallel and reactive empathy and a component to display affect and attention with virtual characters. Currently, emotion recognition is solely based on an analysis of acoustics using a microphone. The user's focus of attention is tracked employing an eye tracker. Further, we use a realistic virtual character with highly expressive facial emotions to display appropriate facial expressions.

A lot of effort has been spent on the implementation of emotional models. Prominent examples include EMA (Gratch & Marsella, 2004) and ALMA (Gebhard, 2005). We use ALMA because it includes a language for defining emotion display appraisal rules to describe how a virtual agent appraises its own emotional displays and that of other agents. These rules are in particular useful when implementing an empathetic listener that responds to the user's emotional state. In addition, ALMA combines an appraisal mechanism with a dimensional representation of emotions which is of benefit when describing how emotions evolve over time. We combine the component for affective sensory input with an emotional model for our empathic listener that allows it to act or react to the users' feelings.

Affect and Attention Sensing

For sensing affect from users' voice, we use EmoVoice (Vogt, André, & Bee, 2008). It is a framework that provides support for the acquisition of emotional speech corpora and the training of classifiers. Furthermore, it is suitable both for offline as well as online vocal emotion recognition. The framework is intended to be used by non-experts and therefore comes with an interface to create an own personal or application specific emotion recognizer. Table 1 shows some examples for recognition accuracies achieved by EmoVoice on different corpora. The FAU Aibo Emotion Corpus (Steidl, 2009) is a spontaneous emotions corpus of children interacting by voice commands with an Aibo pet robot. The Berlin database (Burkhardt, Paeschke, Rolfes, Sendlmeier, & and Weiss, 2005) is a corpus of acted emotional speech. Furthermore, we had students record their own corpora in order to test how easy the process of classifier building and corpus creation is for people who are unfamiliar with the task.

Tabelle 1: Examples for recognition accuracies achieved by EmoVoice

Corpus	Number of emotion classes	Speaker independent?	Recognition accuracy
Aibo	4	Yes	57 %
Berlin	7	Yes	83 %
Students	4	No	24-74 %
	4	Yes	41 %
	2	Yes	60-70 %

Eye gaze may be considered one of the most important means to indicate attention in a dialogue. Kleinke shows correlations between attention and gaze (Kleinke, 1986). For instance, interviewers are evaluated more attentive when their gaze is relatively high. And further, interviewees give shorter responses when an interviewer is not looking at them. To measure the amount and orientation of the user's gaze, we connect the eye tracker with the virtual world. Ray casting allows us to map the screen coordinates obtained from the eye tracker to the objects in the virtual world. In this vein, we are able to detect whether the user looks at the virtual agent, the left eye or the right eye or something else in the virtual scene. The accuracy of eye tracking devices is limited to about 0.5-1° visual angle. 1° corresponds approximately to the size of a thumbnail at arm length (Duchowski, 2007). This restricts the interaction elements in an interface to a certain size. Further, eye gaze is recognized with a delay dependent on the frame rate. A 50 Hz system, for instance, incurs delays of 20 ms, when using webcams with 25 Hz, the delay would be 40 ms.

Affect Model

For the affect simulation in real-time, we rely on ALMA Bio and extended version of the computational model ALMA (Gebhard, 2005). ALMA Bio allows processing of values directly derived from sensor signals (e.g. speech or bio signals). In this context the signals are treated as unspecific emotions that contain individual (measured) pleasure, arousal and dominance

(PAD) values. These measured emotions are used as input to change or intensify the current mood.

ALMA provides three affect types as they occur in human beings: (1) emotions reflect short-term affect that decays after a short period of time; (2) moods reflect medium-term affect, which is generally not related to a concrete event, action or object; and (3) personality reflects individual differences in mental characteristics and affective dispositions.

ALMA implements the cognitive model of emotions developed by Ortony, Clore and Collins (OCC) (Ortony, Clore, & Collins, 1988) combined with the BigFive model of personality (McCrae & John, 1992) and a simulation of mood based on the PAD model (Mehrabian, 1995). The relations between the different affect types are a central part of the affect simulation:

- A given personality defines a default mood and influences the intensities of different emotions.

- The current mood amplifies or dampens the intensities of emotions.

- Emotions as short term affective events influence the longer-term mood.

Elicited emotions influence an individual's mood. The higher the intensity of an emotion is, the higher the particular mood changes. Emotions usually do not last forever. Over a specific period the intensity of emotions decays and the influence on the current mood fades.

The current mood also influences the intensity of emotions (Neumann, Seibt, & Strack, 2001). This simulates, for example, the intensity increase of joy and the intensity decrease of distress, when an individual is in an exuberant mood. Mood is represented by a triple of the mood traits pleasure (P), arousal (A) and dominance (D). The mood's trait values define the mood class. If, for example, every trait value is positive (+P, +A, +D), the mood is exuberant.

Alfred – Affect Display

Alfred is capable of generating empathy through displaying affective facial expressions and generating attentive gaze behavior. While (Niewiadomski, Ochs, & Pelachaud, 2008) concentrate on improving empathic expressions by generating complex facial expressions like superposition and masking, we focus on generating empathic feedback by combining affective facial expressions and attentive gaze behavior that reacts on the user's eye gaze. Composing facial expressions with the FACS based approach gives Alfred a great flexibility to generate a high number of affective expressions. Three examples of affective facial expressions can be seen in Figure 1. The attentive gaze behavior can be varied to generate attention dependent on the user's state. Further, Alfred is able to speak and to nod his head. Alfred is able to talk with a text-to-speech (TTS) system or by playing prerecorded (emotional) audio files with synchronized lip movements. Although these capabilities are not yet used as empathic feedback cues, one can easily imagine how Alfred's capabilities can be enhanced to head nodding or verbal feedback such as "Mm Hm" or "Uh huh" while showing empathy.

Facial Expression

Ekman and Friesen developed the Facial Action Coding System (FACS) to classify human facial expressions (Ekman & Friesen, 1978). FACS divides the face into action units (AU) to describe the different expressions a face can display (e.g. inner brow raiser, nose wrinkler, or cheek

puffer). Although FACS was originally designed to analyze natural facial expressions, it turned out to be usable as a standard for production purposes, too. That is why FACS based coding systems are used for the generation of facial expressions displayed by virtual characters, like Kong in Peter Jackson's King Kong (Sagar, 2006). But the usage of FACS is not only limited to virtual characters in movies. The gaming industry with Half-Life 2 by Valve, also utilizes the FACS system to produce the facial expressions of their characters (Valve).

Alfred (see Figure 1), a butler-like character, uses these action units to synthesize an unlimited set of different facial expressions. The action units were designed using morph targets and thus give the designer the full power in defining the facial expression outlook. The system includes a tool to control the single action units. The tool allows storing the result in a XML file for later usage in our agent system (Bee, Falk, & André, 2009).

We chose the FACS-based approach for our facial animation system, because of the Facial Expression Repertoire (FER) (Repertoire), which maps over 150 emotional expressions to the action units of FACS. Not only does it explain in detail, which action unit must be activated for certain facial expressions, it further provides a rich dataset of videos which show how the action units ought to be designed.

Alfred's mesh has a resolution of about 21.000 triangles. For displaying more detailed wrinkles in the face, normal maps baked from a high-resolution mesh are used (Oat, 2007). The morph targets for the action units are modeled using the actor's templates from the FER. For rendering the character and its animations the Horde3D GameEngine (GameEngine) is used.

Gaze Behavior

Fukayama et al. propose a gaze behavior model for virtual characters based on amount and mean duration of gaze and averted gaze orientation (Fukayama, Ohno, Mukawa, Sawaki, & Hagita, 2002). They rated with two groups of attributes. One was named "friendliness" and was correlated in their study with attributes such as friendly, warm, sociable, tolerant, flexible, attentive and coordinative. The other was named "dominance" and correlated with assured, strong, successful, responsible and careful. Fukayama and colleagues found that a medium amount of gaze, a mean duration between 500 to 1000 ms conveys a "friendly" gaze behavior. The orientation of the gaze direction did not play a major role between the friendly and dominant gaze behavior, except a downward gaze was considered as less dominant.

Analogous to Fukayama et al. our gaze model has two states, looking at user and averted from user. Three parameters define how often, how long (500 to 2000 ms) and where the virtual agent looks. The gaze targets consist of a set of random points from either all over the scene, above, below or close to the user. The probabilities of changing from one state to the other or staying in the same state depend on the amount and the mean duration of the gaze parameters.

Alfred's gaze behavior was extended with further parameters as our virtual agent is capable to react on the user's current focus of attention. A confidence value defines the agent's level of self assurance and influences the probability that the agent initiates the gaze interaction. The maximal and minimal duration of mutual gaze can be set as well. Furthermore, we may indicate the maximal duration the virtual agent gazes around. Finally, we may specify how long the virtual agent waits until the user responds with mutual gaze. These parameters are

stored in a XML file and can be easily adjusted. The virtual agent is able to direct his gaze using his eyes only or his head and his eyes in combination.

Fukayama et al. evaluated their gaze behavior model with only displaying eyes to the users. Thus, we evaluated their model with a full virtual head that in addition moves his head and eyes. Although currently we focused on dominant gaze behavior, we could already confirm that setting the same parameters conveys the same dominant impression.

Speech

The system interfaces the Microsoft Speech API to synchronize the audio output with the lip movements. This allows us to use any text-to-speech system that supports SAPI 5. As the quality of common TTS systems may not be satisfactory, we integrated a module to synchronize prerecorded audio speech files with the lip movements of the virtual character. This allows us to use highly emotional sentences or affect bursts to be spoken through a virtual character. As FACS defines several action units involving mouth muscles (e.g. lip funneler, lip tightener, mouth stretch), we utilize the FACS system for lip movements. The approach is similar to displaying facial expressions. The output from the editor to modify the single action units is stored in a XML file. Reusing the FACS approach for visemes enables Alfred to display facial expressions and lip movements parallel.

EMPATHIC FEEDBACK

For the modeling of empathic feedback, we employ ALMA to simulate the agent's and the user's emotional state. Both states are independently represented. From the point of view of the agent, the user's emotional state is simulated according to the agent's own model of emotions (ALMA) but based on appraisal rules that the agent assumes the user has. Due to the fact the personality of the user is not known, ALMA is used with neutral personality values. This allows a simulation of mood tendencies, which is depicted in Figure 2.
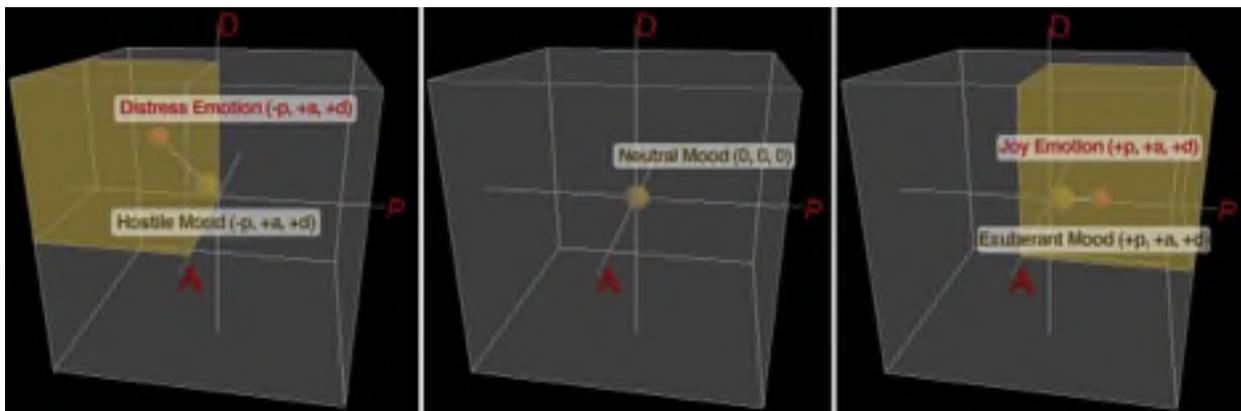

Figure 2: Screenshots of mood tendency simulation with ALMA.

Figure 2 shows the mood simulation based on neutral personality settings (center). Each elicited emotion impacts the neutral mood. The mood is altered according to an elicited emotion's relation to the PAD mood trait values (e.g. Joy is related to +P, +A, +D values, right side. Distress is related to –P, +A, +D values, left side). The simulated mood represents a

general mood tendency that is influenced by emotions. If no emotion is elicited, the mood slowly decays to the neutral mood.

Our approach for the modeling of empathic feedback consists of 3 stages:

1. Parallel Empathy (Mimicry of the user's emotion)

2. Reactive Empathy (Reacting to the user's emotion)

3. Deliberate Empathy (Altering the user's emotion)

Both, in the case of parallel empathy and reactive empathy, emotional responses are generated using the agent's current mood PAD values (representing the agent's emotional state) in order to influence the agent's facial expressions. In the case of parallel empathy, the PAD values representing the agent's emotional state correspond to the PAD values representing the user's emotional state. In the case of reactive empathy, PAD values for the agent's emotional state are assigned based on an appraisal mechanism.

While this approach is suitable for parallel empathy and less complex forms of reactive empathy, such as good-will emotions, the handling of communicative empathy that is shown with the aim to change the user's emotional state requires further work.

Parallel Empathy

In the case of parallel empathy, the agent mimics the affective state of the user. Hoffman defines mimicry as primitive, automatic and involuntary which is a process of the two steps "imitation" and "feedback". Humans automatically imitate and synchronize changes in their facial expression and voice. Through changes in the musculature, these synchronizations produce feelings that match (Hoffman, 2000).

We used ALMA to create a model of the user's emotional behavior. In the case of mimicry, the user's affective input is directly mapped onto PAD values. Currently, we just rely on vocal emotion recognition for the mapping even though the user's gaze conveys information on his or her emotional state as well. In particular, the PAD values for dominance could be derived from gaze. To avoid jumps between emotional states, we smooth the input signals during affective interpretation. Furthermore, we make use of a decay factor to account for the fact that the importance of a signal decreases over time. For more details, we refer to (Gilroy, et al., 2008). The emotional state of the user in the model is taken as a basis for the character's emotional behavior which just mirrors the user's emotional state. The user's emotional state is represented by PAD values that were derived from vocal emotion recognition. The same PAD values may be used to generate facial expressions with Alfred.

Reactive Empathy

For reactive empathy, we make use of ALMA's Emotion Display Appraisal Rules that describe how a character appraises the emotion displays of others. Those rules basically map an observed emotion display onto Appraisal Tags, which are shortcuts to the OCC emotion eliciting conditions. Given as input, ALMA computes emotions based on those tags. ALMA provides 3 Appraisal Tags classes: 12 Event Appraisal Tags (e.g. GoodEvent, BadEvent, BadEventForGoodOther, GoodLikelyFutureEvent …), 4 Action Appraisal Tags (e.g. GoodActSelf, BadActSelf, GoodActOther, BadActOther), and 2 Object Appraisal Tags (e.g. NiceThing). For

reactive empathy, the agent assumes that the user is likeable, which has an impact on how the agent appraises a user's emotional display. Figure 3 shows part of the used rules.

Disappointment, Distress, Remorse, SorryFor -> BadEventForGoodOther 1.0

Joy, Gratification, HappyFor, Satisfaction -> GoodEventForGoodOther 1.0

Figure 3. Example appraisal rules for emotional displays.

In the case the agent observes that the user displays a negative emotion (Watson and Clark, 1984) (e.g. SorryFor or Distress) the agent appraises the situation as BadEventForGoodOther. As a first approach, the intensity value 1.0 for appraisal is used. Due to the fact that the Appraisal Tag BadEventForBadOther mapped to the OCC-variable-value pairs pos. liking, and pos. desirability, and agency other, ALMA can derive (according to the OCC model of emotions) the emotion SorryFor. This will impact the facial expression of Alfred, as shown in Figure 1 (right side).

For events that trigger an emotional response on a user, our agent is capable to have the following reactive empathy emotions and show appropriate reactions: HappyFor, SorryFor, Resentment, and Gloating. This depends basically on the fact if the agent likes the user. For actions, our agent is capable to show appropriate reactions for the emotions Shame, Reproach and for the compound emotion Anger due to the fact those are observable. For objects, Alfred is able to show an appropriate reaction to the emotion Dislike due to the fact it can be easily observed (e.g. a face showing expressing disgust).

Deliberate Empathy

At a higher level of reactive empathy, the agent does not display emotions it feels, but tries to express deliberately emotions that might help to put the user into a better emotional state. Reactive empathy requires the agent to put himself into the shoes of the user in order to decide on an appropriate emotional display. For example, if the user is afraid of failing in an exam, the agent might tell the user that he will manage due to his good preparation. That is the agent praises the user which is perceived by the user a positive event. In addition, the agent utters a statement which decreases for the user the likelihood of the negative event (namely failing in the exam). Furthermore, in the deliberate mode, it is crucial for the virtual agents to be able to express themselves verbally - specifically if their emotional display does not coincide with the user's emotion. For example, if the user is sad and the agent is smiling, the user might not perceive this as encouragement, but as gloating (see Figure 1 middle). ALMA does can potentially serve as affect model for higher level of reactive empathy. In detail, it could be used to realize an affective theory of mind by running a simulation on a planned utterance. The provided appraisal methods directly supports this approach (e.g. OCC appraisal variable-value pairs, Appraisal Tags, Emotion Display Tags, Dialog Act Tags, and direct PAD input). Suppose the agent intends to generate an utterance, such as "Don't worry about the exam. You are so well prepared." The good preparation of the user decreases the likelihood of the negative event (Fail Exam). According to the OCC model, the intensity of the emotion Fear would be decreased.

CONCLUSION

Our approach of creating an affective and attentive virtual listener that reacts in an empathic manner to the user's input focuses on the interpretation and generation of affective and attentive cues. We have shown how an emotional model can be used to simulate different levels of empathy. Parallel empathy can be handled by mapping input signals onto PAD values which are then transformed into appropriate output signals. In contrast, reactive emotions, such as sorry-for or happy-for, require an appraisal mechanism that determines how the agent responds to the display of other agents. Even if the agent does not understand what the user is saying, he might still be able to communicate via affective cues that he is caring for the user simply by emotionally responding to the user's emotions. A great challenge represents deliberate empathy that aims at changing the counterpart's emotional state. In this case, the agent usually needs to understand why the user is feeling a certain emotion in order to give appropriate feedback.

Another non-trivial issue is to decide which form of empathy a virtual agent should show. One criterion to consider might be the personality the agent is supposed to portray. For example, an agreeable agent tends to show parallel empathy more frequently than a disagreeable agent. Our system provides the possibility to test different behavior variants in order to figure out under which conditions which form of empathy is more appropriate.

BIBLIOGRAPHY

Bee, N., Falk, B., & André, E. (2009). Simplified Facial Animation Control Utilizing Novel Input Devices: A Comparative Study. *International Conference on Intelligent User Interfaces (IUI '09)*, (pp. 197-206).

Boukricha, H. (2008). A First Approach for Simulating Affective Theory of Mind Through Mimicry and Role-Taking. *The Third International Conference on Cognitive Science, Symposium: Emotional Computer Systems and Interfaces*.

Bringsjord, S., Shilliday, A., Clark, M., Werner, D., Taylor, J., Bringsjord, A., et al. (2008). Toward Logic-Based Cognitively Robust Synthetic Characters in Digital Environments. *Proceedings of the First Conference on Artificial General Intelligence (AGI-08)*.

Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., & and Weiss, B. (2005). A database of German emotional speech. *Interspeech 2005*.

Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and Practice*. Springer.

Ekman, P., & Friesen, W. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press.

Fukayama, A., Ohno, T., Mukawa, N., Sawaki, M., & Hagita, N. (2002). Messages embedded in gaze of interface agents --- impression management with agent's gaze. *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, (pp. 41-48).

GameEngine, H. (n.d.). *University of Augsburg*. Retrieved from http://mm-werkstatt.informatik.uni-augsburg.de/projects/GameEngine/

Gebhard, P. (2005). ALMA - A Layered Model of Affect. *Proc. of the 4th Int. Joint Conference on Autonomous Agents and Multiagent Systems* (pp. 29-36). ACM.

Gilroy, S. W., Cavazza, M., Mäkelä, R. C., Niranen, M., André, E., Vogt, T., et al. (2008). An affective model of user experience for interactive art. *ACE '08: Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*, (pp. 107-110).

Gratch, J., & Marsella, S. (2004). A domain-independent framework for modeling emotion. *Cognitive Systems Research , 5* (4), 269-306.

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007). Creating Rapport with Virtual Agents. *Intelligent Virtual Agents (IVA 2007)*, (pp. 125-138).

Hoffman, M. L. (2000). *Empathy and Moral Developmen*. Cambridge University Press.

Kleinke, C. L. (1986). Gaze and Eye Contact: A Research Review. *Psychological Bulletin* , 78-100.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality , 60*, 175-215.

McQuiggan, S. W., Robison, J. L., Phillips, R., & Lester, J. C. (2008). Modelling Parallel and Reactive Empathy in Virtual Agents: An Inductive Approach. *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, (pp. 167-174).

Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. *Genetic, social, and general psychology monographs , 121* (3), 339-361.

Neumann, R., Seibt, B., & Strack, F. (2001). The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing. *Journal of Cognition & Emotion , 15* (6), 725-747.

Niewiadomski, R., Ochs, M., & Pelachaud, C. (2008). Expressions of Empathy in ECAs. *Intelligent Virtual Agents (IVA '08)*, (pp. 37-44).

Oat, C. (2007). Animated wrinkle maps. *SIGGRAPH '07: ACM SIGGRAPH 2007 courses* (pp. 33-37). New York, NY, USA: ACM.

Ochs, M., Pelachaud, C., & Sadek, D. (2008). An Empathic Virtual Dialog Agent to Improve Human-Machine Interaction. *7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, (pp. 89-96).

Ortony, A., Clore, G. L., & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge, MA: Cambridge University Press.

Pelachaud, C., & Poggi, I. (2002). Subtleties of facial expressions in embodied agents. *The Journal of Visualization and Computer Animation* , 301-312.

Peters, C., Pelachaud, C., Bevacqua, E., Ochs, M., Chafai, N. E., & Mancini, M. (2006). Social capabilities for autonomous virtual characters. *International Digital Games Conference*, (pp. 37-48).

Prendinger, H., Dohi, H., Wang, H., Mayer, S., & Ishizuka, M. (2004). Empathic embodied interfaces: Addressing users' affective state. *In Proceedings Tutorial and Research Workshop on Affective Dialogue Systems, LNAI 3068*, (pp. 53-64).

Repertoire, F. E. *Filmakademie Baden-Württemberg*.

Sagar, M. (2006). Facial performance capture and expressive translation for King Kong. *SIGGRAPH '06: ACM SIGGRAPH 2006 Sketches* (p. 26). New York, NY, USA: ACM.

Schröder, M., Cowie, R., Heylen, D., Pantic, M., Pelachaud, C., & Schuller, B. (2008). Towards responsive Sensitive Artificial Listeners. *Fourth International Workshop on Human-Computer Conversation*. Bellagio, Italy.

Steidl, S. (2009). *Automatic Classification of Emotion-Related User States in Spontaneous Children's Speech*. Berlin: Logos Verlag.

Valve. (n.d.). *Facial Expressions Primer from Half-Life 2*. Retrieved from http://developer.valvesoftware.com/wiki/Facial_Expressions_Primer

Vogt, T., André, E., & Bee, N. (2008). EmoVoice - A framework for online recognition of emotions from voice. *Proceedings of Workshop on Perception and Interactive Technologies*.

D. Watson and L. A. Clark. Negative affectivity: The disposition to experience aversive emotional states. In *Psychological Bulletin*, number 96, pages 465–490, 1984.