Proceedings of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conference

No Peanuts! Affective Cues for the Virtual Bartender

Marcin Skowron and **Hannes Pirker** and **Stefan Rank**
Austrian Research Institute for Artificial Intelligence
Vienna, Austria

Georgios Paltoglou
Wolverhampton University
Wolverhampton, UK

Junghyun Ahn and **Stephane Gobron**
Virtual Reality Lab, EPFL
Lausanne, Switzerland

Abstract

The aim of this paper is threefold: it explores methods for the detection of affective states in text, it presents the usage of such affective cues in a conversational system and it evaluates its effectiveness in a virtual reality setting. Valence and arousal values, used for generating facial expressions of users' avatars, are also incorporated into the dialog, helping to bridge the gap between textual and visual modalities. The system is evaluated in terms of its ability to: i) generate a realistic dialog, ii) create an enjoyable chatting experience, and iii) establish an *emotional connection* with participants. Results show that user ratings for the conversational agent match those obtained in a Wizard of Oz setting.

1 Introduction

An important aspect of the use of artificial agents in interactive environments, from virtual training to non-player characters in games, is to deal with the affective states of human participants. Many application domains require competences in both exchanging task-specific information with users as well as establishing, maintaining and developing an emotional connection, taking into account, e.g, a user's affective state or satisfaction level. Thus, systems need to detect changes in users' affective states, suitably reacting to them by incorporating *affective cues* into the autonomous characters' generated behaviour in all available modalities. Sentiment Analysis (also known as Opinion Mining) deals with the computational treatment of textual expressions of private states (i.e. personal states that are not open to objective observation or verification); the research area of Virtual Humans (VH) covers, among other things, the important communication channel of facial expressions; research on conversational agents focuses on the verbal behaviour of artificial entities.

This paper deals with the interaction between a virtual agent, the *Affect Bartender*, and users represented by their avatars in a 3D virtual reality (VR) bar environment (section 3). We describe the architecture of the conversational agent that manages aspects of verbal communication between this VH and its users (section 4). Next, we focus on methods for analysis of textual expressions of affective states, and how this information on affective cues is applied in dialog management (section 5). Finally, we provide an overview of the experiment results (section 6). The conversational agent is compared to a Wizard-of-Oz (WOZ) scenario, considering the agent's capacities to: achieve a realistic dialog, provide an overall enjoyable chatting experience for the experiment participants, and establish and maintain an *emotional connection* with them.

2 Relevant Research

Sentiment analysis has been a popular research topic in recent years, mostly focusing on analysing reviews, e.g. (Pang, Lee, and Vaithyanathan 2002; Blitzer, Dredze, and Pereira 2007), but also in other domains such as political debates or news (Thomas, Pang, and Lee 2006; Devitt and Ahmad 2007). The seminal book (Pang and Lee 2008) presents a thorough analysis of the field. (Pang, Lee, and Vaithyanathan 2002) were among the first to explore the sentiment analysis of reviews, focusing on machine-learning approaches. Later, the same authors presented an approach based on detecting the subjective parts of documents (Pang and Lee 2004). Most other approaches in the field have focused on extending the feature set with semantic or linguistic features, e.g. (Whitelaw, Garg, and Argamon 2005) who used fine-grained semantic distinctions to improve classification.

The development of computer systems that incorporate the modeling of emotional behavior receives a significant interest in the research community (Picard 1997; Schroeder and Cowie 2006; Petta, Pelachaud, and Cowie 2011). An important part is the recognition of affective states in multiple modalities: visual (e.g, facial expressions), auditory, as well as their combination (Sebe et al. 2007; Cowie et al. 2008; D'Mello, Picard, and Graesser 2007). This forms the background for the development of embodied conversational agents (ECAs), virtual characters and VH (Gebhard et al. 2008; Caridakis et al. 2008). The management of human-computer conversations that incorporate emotional cues is the central area of interest for Affective Dialog Systems (ADSs). This multidisciplinary field integrates work from a range of research areas, e.g., dialog processing, speech recognition, speech synthesis, computer graphics, animation, embodied conversational agents and human-computer interaction (André et al. 2004).

In the research fields of VR and computer graphics, emotional expressions are a “hot topic” touching several areas such as facial expression (Pelachaud 2009), body motion (Egges, Molet, and Magnenat-Thalmann 2004), and gaze analysis (Grillon and Thalmann 2009). Su, Pham, and Wardhani (2007) used personality types and emotional states to consistently control body language of virtual characters. Recently, Tsetserukou and Neviarouskaya (2010) proposed an original sensorial based system (i.e. augmented reality) enabling users to receive affect feedback from text chatting.

3 Interaction Setting

As setting for the experiment, a 3D VR bar was created: a bar room, virtual furniture, and a virtual bartender. The virtual bartender also performs typical activities, e.g., cleaning dishes, when users take longer to reply. Technical aspects of the VR process pipelines are described in (Gobron et al. submitted). Initially, users see how their avatar enters the premises and moves towards the bar counter. Upon reaching the counter, the perspective changes and an isometric view of the scene is presented, see Fig.1(a). Two windows at the top of the screen show close-ups of the bartender’s and avatar’s faces. See Fig.1(b) for an example of an emotional facial expression (EFE). A chat interface allows the user to type utterances and displays the bartender’s responses.

Figure 1: Screenshot of the VR bar setting

4 Conversational Agent: Architecture

The *Affect Bartender*, i.e. the conversational agent, is responsible for the management of verbal communication between the virtual bartender and a user, represented in the virtual 3D bar by an avatar. The main objectives for the system in the above presented interaction scenario are:

1. achieving realistic dialogs,
2. providing an enjoyable overall chatting experience,
3. establishing and maintaining an emotional connection.

The implementation of the system is based on the concept of *Affect Listeners* (Skowron 2010), conversational agents for detection of and adaptation to affective states of users (i.e., textual expression of users’ affective states), to meaningfully respond to users’ utterances both at the content- and affect-related level.

The core tasks of the Affect Bartender in the context of the 3D virtual bar scenario include: perception and classification of affective cues in user utterances and system response candidates, the incorporation of affective cues into the dialog management, maintenance of an emotional connection with users (affective dialog management), management of task-oriented dialogs (closed-domain dialog) as well as conversations not restricted in topic (open-domain chats), and, finally, the detection of cues in the system-user interactions that enable the selection of suitable system response generation methods (balancing task oriented dialog vs. open-domain conversations). Fig.2 presents the top-level layers of the system architecture (communication, perception, control) and the interaction loop with the environment.

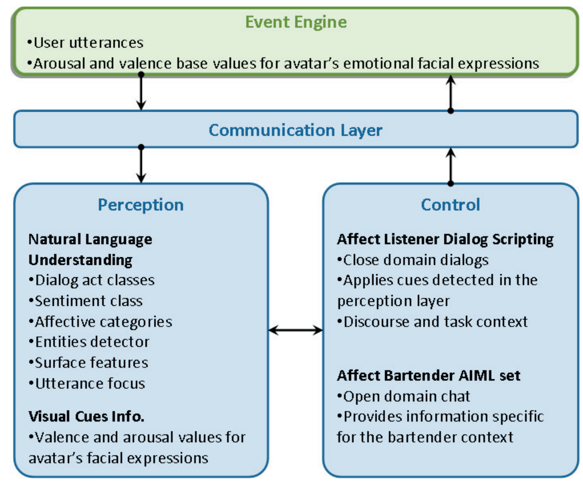

Figure 2: Interaction loop and conversational system layers

4.1 Perception Layer

The Perception Layer integrates different natural language processing tools, linguistic and affective resources to analyze user utterances and system response candidates:

- Dialog Act (DA) classifier: based on the annotation scheme used in the NPS Chat Corpus (Forsyth and Martell 2007). For the present scenario, the original taxonomy (Accept, Bye, Clarify, Continuer, Emotion, Emphasis, Greet, No Answer, Other, Reject, Statement, Wh-Question, Yes Answer, Yes/No Question) was extended with an additional class “Order” (i.e. ordering drinks) using 339 additional training instances. For this taxonomy and training set, the maximum entropy based DA classifier using a bag-of-words and bag-of-bigrams feature set achieved 10-fold cross validation accuracy of 71.2%
- Sentiment Classifier: provides information on sentiment class (SC), positive (PS) and negative sentiment values (NS): see section 5.1
- Linguistic Inquiry and Word Count - LIWC (Pennebaker, Francis, and Booth 2001) (LC): see section 5.1
- Gazetteers and regular expressions for detecting bar-context specific entities, i.e., drinks (DR) and snacks (SN)
- Surface features detector: e.g., exclamation marks (EM), emoticons (e.g., EE-sad, EE-happy)

Input	Perception Layer Output (excerpt)
I have a problem at work	DA-Statement SC- -1 NS- -3 PS-1 LC-Affect:Negemo:CogMech :Discrep:Present DR-0 EE-0

Figure 3: Perception Layer – annotation example

- Utterance focus and interest detector (Skowron, Irran, and Krenn 2008)

Fig.3 presents an excerpt of the Perception Layer annotation for the input "I have a problem at work"[1].

4.2 Control Layer

The Control Layer manages the dialog progression by relating the observed dialog states to the intended ones (e.g., conducting specific bartender tasks, querying and follow-up questions on the user's affective states) using cues acquired by the Perception Layer described above (e.g, linguistic or affective categories discovered in a user utterance). This layer selects the system response from a number of generated response candidates, integrating rule-based action selection—Affect Listener Dialog Scripting (ALDS)—with the command interpreter for the Affect-Bartender AIML-set[2] described below.

4.3 Communication Layer

The Communication Layer provides the conversational system with an interface to the 3D VR event engine: these components are situated on separate hosts and the Communication Layer is handling their connection via an XMLRPC protocol. It receives and decodes user utterances and arousal and valence values calculated when generating emotional facial expressions (EFE). Further, the layer formats and dispatches system responses.

5 Affective Cues

For the purpose of this work, we define *affective cues* as indicative evidence of a user's affective state that can be perceived by the agent; in particular, in the case of a conversation system, these relate to the textual expressions of users' affective states.

5.1 Perception and Classification

The capability to detect and to classify textual expressions of affective states in utterances of the users is a core prerequisite for the Affect Bartender system. In the present realization of the conversational agent, two affect detection and classification methods are used: Sentiment Classifier (Paltoglou et al. 2010) and Linguistic Inquiry and Word Count (LIWC) dictionary (Pennebaker, Mehl, and Niederhoffer 2003). The third method presented in this section, Multi-Dimensional Probabilistic Emotional Histogram (Gobron et al. 2010), models users' emotional states during interaction and forms the basis for the generation of EFE of the VH and the user's avatar.

Linguistic Inquiry and Word Count (LIWC) This dictionary enables the conversational agent to detect 64 linguistic, cognitive and emotional dimensions. This resource provides 32 word categories that are indicating psychological processes (e.g., affective such as positive and negative emotions; social such as family, friends and human; cognitive such as insight, causation, tentative), 22 linguistic processes (e.g., adverbs, negations, swear words), 7 personal concern categories (e.g., home, religion, work, leisure), 3 paralinguistic dimensions (fillers, assents, nonfluencies), for almost 4500 words and word stems. For example, the word "won" is categorized in 5 categories: affective processes, positive emotion, achievement, verb and past tense. In recent years, LIWC has been successfully applied in various psychological and psycholinguistic studies that included e.g., the investigation of speakers linguistic style, the relations between language use and speakers personality (Chung and Pennebaker 2008).

The Perception Layer relies on the LIWC dictionary to detect words in user utterances and system response candidates which are related to affective categories such as positive emotion, negative emotion, anger, anxiety and sadness. Further, the resource provides also cues about other categories useful for managing the system user interaction, taking into consideration the agent's tasks i.e., expressions from categories such as: leisure, work, swear words or health.

Sentiment Classifier (SC) We approach the problem of sentiment analysis from the perspective of a Lexicon-based classifier. While machine-learning approaches have been the main focus of research in the field, previous work (Paltoglou et al. 2010; Neviarouskaya, Prendinger, and Ishizuka 2010) has shown that lexicon-based approaches are usually more effective for short text-messages typically found in social communication environments. The Lexicon-based classifier is based on two different emotional word-lists: The General Inquirer and LIWC. We use emotional indicators assigned to tokens in those dictionaries and enhance the produced prediction with linguistically-driven functionalities, such as negation detection, capitalization, intensifier/diminisher identification, emoticon detection. The classifier outputs a vector of two scores, {neg,pos}, where neg={-5,...,-1} and pos={+1,...,+5}. Higher absolute values indicate stronger emotional content with -1,+1 indicating lack of emotion. For example {-1,+4} would indicate a strong positive emotion, {-5,+1} a very strong negative emotion and {-3,+4} a mixed emotional response, slightly more positive. The Lexicon-based classifier initially scans the text for emotional words contained in either dictionary and adapts their original scores if modifiers are detected within the same sentence. For example the word "love" has an initial score of +4, but if detected in a sentence with exclamation marks its score is increased by 1. A negation before the word would inverse its valence and further decrease its absolute value by 1 to a final score of -3.

[1]For additional examples and detailed description of the mechanisms used for perception and application of textual affective cues in a conversational agent, refer to (Skowron and Paltoglou 2011).

[2]Artificial Intelligence Markup Language (AIML)

EmoMind and Probabilistic Emotional Histogram EmoMind is the model used for generating emotional facial expressions (EFE) in both virtual humans and avatars. It considers the affective content of utterances from the conversational system or a user. In particular, it uses four parameters obtained from the Sentiment Classifier (i.e., positive sentiment, negative sentiment, intensity, objectivity) as an input for a Poisson distribution based model. EmoMind takes this input from the Sentiment Classifier and uses a Probabilistic Emotional Histogram (PEH) for calculating values for valence and arousal (v,a), similar to Russell's circumplex model of emotions.

This approach, while only roughly mimicking any psychological reality, ensures statistical consistency. Further, it allows taking into account previous affective states, it simulates emotional ambivalence as different types of emotions (e.g. a weak joy and strong anger) can be present simultaneously, and it provides an output emotion $\{v,a\}$ suitable for both modifying the non-verbal display (i.e. facial expression) and as input to the dialog managment components for planning verbal communication. For instance, in Fig.1(a), the close-up illustrates a typical facial expression resulting from a high value of arousal (i.e. energetic emotion) and a moderate negative value of valence (i.e. a negative feeling).

5.2 Use of Cues in Dialog Management

Affective cues play an important role in the generation of response candidates, in response modification and selection. Specifically, the agent incorporates information on affective states for the following three conditions as will be explained in this section:

1. detection of high arousal in a user utterance (Sentiment Classifier),
2. rapid change in the user's affective states (based on two consecutive message exchange turns, i.e., analysis of user utterances or $\{v,a\}$ values used for the avatar's EFE),
3. occurrence of a system response that was generated by a 'confusion statement' template, (i.e., a response without direct relevance to the factual content of a user's utterance), paired with the detection of a particular affective category in an utterance of the user (LIWC).

Affect Listeners Dialog Scripting (ALDS) Introduced in more detail in (Skowron 2010), ALDS enables the creation of interaction scenarios that provide capabilities to control task-oriented parts of verbal communication spanning several dialog turns, i.e., system and user utterances, and that take advantage of the system's perception capabilities (i.e., natural language analysis, affective states analysis) that extend beyond a simple matching mechanism solely based on keywords or textual patterns. For the Affect Bartender system, the ALDS scenario relies on the affective, linguistic and cognitive categories discovered in a user utterance. In contrast to more complex communication tasks, e.g., receiving orders in a virtual bar context, the application of affective cues relies on a pre-defined link between an initiation condition (e.g., user inputs and/or system state) and a particular system response template. To prevent a discernable repetition of strategies, each category of the affect-related ALDS scenario presented here (i.e., sentiment class, EFE, particular affective or linguistic category of LIWC dictionary) is used only once during an interaction.

For the first situation mentioned above, detection of high arousal based on the Sentiment Classifier presented in section 5.1, the activation threshold was set to +5 for positive and -5 for negative sentiment. These values signal the usage of highly emotional words, which are often found in utterances that contain information important for the user or which convey strongly emotional expressions. This triggers a system response focusing on the expressed state rather than on the content of a user utterance e.g.: negative sentiment (NS= -5), example system response: "are you disappointed? ... if it is my fault, i am really sorry.".

Further elements used in the ALDS scenario are dialog act classes, surface features of utterances (e.g., emoticons) and the valence and arousal values used for generation of EFEs which allows the dialog manager component to relate to these visual cues perceived by the user (condition 2). The agent generates questions which explicitly refer to the displayed EFE and their relevance for the affective states expressed verbally by the user. For example, if an utterance is classified as positive while the valence of the avatar's facial expression is negative the system might react by generating an utterance such as: "You sound optimistic, but you don't look that way... What is going on?". If a rapid change in the valence values for EFEs is detected (difference >4), and during both utterances the arousal value $\{a\}$ was larger than a set threshold, the system might generate the following response candidate: "Did something happen? You suddenly do not look that optimistic at all...".

For artificial conversational entities, the third condition mentioned above, i.e., the inability to provide a response on the basis of an analysis (i.e., semantic, discourse) of recent utterances is a relatively frequent problem, especially in open-domain applications. For the present system, this inability occurs when all of the so-far generated response candidates contain a "confusion statement". In such situations, affective cues provide the possibility to shift the focus of the system response from the semantics of the user utterance, to e.g. its affective content. The list of categories used for applying this method of response generation includes: positive or negative emotion, swear words, anger and health; as detected by the LIWC dictionary. For example, the discovery of the swear word category in a user utterance might lead to the following system response: "you look like a really decent person... please don't use this type of words excessively often ;) ... would you do it for me please?".

Affect Bartender AIML set (AB-AIML) For open domain contexts, AB-AIML provides a robust fall-back mechanism able to generate system responses for a range of inputs which do not match activation cues of the provided ALDS scenarios. The adaptation of a more generic Affect Listener AIML set (Skowron 2010) for the Affect Bartender system was twofold, aiming at enabling the system to generate response candidates that: (i) convey the Virtual Bartender's openness, interest in users' feelings, current mood, events which are of importance for them, and (ii) provide knowledge specific to the bartender tasks, and the virtual bar settings. The AB-AIML set contains 14825 patterns, 8549 response instructions, 782 'that' statements and 6999 'srai' substitution rules.

6 Experiments

The experimental setting consisted of the user, represented by an avatar (male or female according to the user's gender), interacting with a VH (male bartender) in the virtual bar as described in section 3. Each participant interacted four times, five minutes each, randomized order, in 2x2 conditions: The conversational partner was either the Affect Bartender system (AB) or a Wizard of Oz (WOZ) [3], and the generation of EFEs was either active or not. We focus exclusively on the difference between AB and WOZ here.

In both conditions, a simulation of thinking and typing speed was introduced to prevent an influence of differences in the response delivery time between the system and the human operator. 35 participants (13 female, 22 male), age between 18 and 45, completed interactions in all four experimental settings resulting in 140 interaction logs. Around 75% of the experiment participants were naive in terms of VR expertise. English, the language in which the experiments were conducted, was not their native language, but all participants had at least good communicative skills in this language.

After each of the experimental interactions, participants were asked the following questions for assessing the conversational system:

1. Did you find the dialog with the VH to be realistic?
2. How did you enjoy chatting with the VH?
3. Did you find a kind of emotional connection between you and the VH?

The participants provided their subjective ratings using a 6-point scale with '1' referring to the most negative and '6' to the most positive assessment. Fig.4 presents the aggregated results obtained for the experimental settings with the Affect Bartender and for those with a Wizard-of-Oz. In all 3 tasks, the results achieved by the conversational system match those obtained for the WOZ. In particular, the correlation coefficient for the aggregated AB and WOZ ratings varied between 0.95 (chatting enjoyment), 0.96 (emotional connection) and 0.97 (dialog realism). All these correlations differ from 0 at a significance level of .001. A repeated measures analysis of variance (ANOVA) showed no main effect of the setting (AB vs. WOZ) on the three dependent measures (all Fs (1,34) $< .50$, ps $> .49$). Pairwise comparisons with Bonferroni correction confirm the absence of significant difference between the two settings on the perception of dialog realism, chatting enjoyment, and subjective feeling of emotional connection with the system.

[3]Participants believe that they communicate with a dialog system, while responses are actually provided by a human operator. In the presented experiments, the operator was asked to conduct a realistic and coherent dialog and provided free text input to user utterances.

	1	2	3	4	5	6	mean	sd
Dialog realism								
AB	3	6	18	11	19	13	4.09	1.42
WOZ	3	7	14	17	16	13	4.07	1.41
Chatting enjoyment								
AB	2	8	15	18	19	8	3.97	1.23
WOZ	3	6	19	18	20	4	3.83	1.23
Emotional connection								
AB	4	14	14	22	15	1	3.47	1.24
WOZ	5	11	18	18	14	4	3.53	1.33

Figure 4: Experiment results: number of participants using a specific rating on the assessment questions for the AB and WOZ conditions on a 6-point scale.

7 Conclusions

In this work, we focused on the presentation of methods for the perception of textual affective cues from system-user communication and their application for the management of virtual agent-user interactions.

In many communication tasks, the ability of the artificial systems to correctly identify the existence and polarity of emotions expressed by users, based on the analysis of short informal messages is a prerequisite for the affective analysis of the ongoing communication and the basis for managing affective aspects of the interactions with users. Such tools also allow the annotation of textual communication between the users in open channels, e.g., blog discussions or exchanges and status updates on social networking sites, which provides insights on the role of emotions in human-human communication. These are useful for modelling, and in the future, for application of the acquired insights to the next generation of affective conversational agents.

The presented methods provide a basis for creating more realistic interaction scenarios by enriching the character traits and expressiveness of VHs, leading to more immersive and satisfying interactions for the users. The conducted experiments demonstrate that in the used interaction settings, the participants' evaluation of the conversational agent that incorporated affective cues into the dialog management was on par with the results recorded in the WOZ settings. This included the participants' assessments of their enjoyment of the interaction, the perceived level of realisticness of the generated dialogs and the participants' degree of establishing an emotional connection with the conversational agent.

Acknowledgments

The work reported in this paper is partially supported by the European Commission under grant agreements CyberEmotions (FP7-ICT-231323) and IRIS (FP7-ICT-231824).

References

André, E.; Dybkjaer, L.; Minker, W.; and Heisterkamp, P., eds. 2004. *Affective Dialogue Systems, Tutorial and Research Workshop (ADS 2004)*, LNAI 3068. Springer.

Blitzer, J.; Dredze, M.; and Pereira, F. 2007. Biographies, Bollywood, Boom-boxes and Blenders: Domain adaptation for sentiment classification. In *Proc.of ACL '07*, 440–447.

Caridakis, G.; Raouzaiou, A.; Bevacqua, E.; Mancini, M.; Karpouzis, K.; L., M.; and Pelachaud, C. 2008. Virtual agent multimodal mimicry of humans. *Language Resources and Evaluation, special issue on Multimodal Corpora For Modelling Human Multimodal Behavior* 31(3-4):367–388.

Chung, C. K., and Pennebaker, J. W. 2008. Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of Research in Personality* 42:96–132.

Cowie, R.; Douglas-Cowie, E.; Karpouzis, K.; Caridakis, G.; Wallace, M.; and Kollias, S. 2008. Recognition of emotional states in natural human-computer interaction. In Tzovaras, D., ed., *Multimodal User Interfaces*, Signals and Communication Technology. Springer. 119–153.

Devitt, A., and Ahmad, K. 2007. Sentiment polarity identification in financial news: A cohesion-based approach. In *Proc.of ACL'07*, 984–991.

D'Mello, S.; Picard, R.; and Graesser, A. 2007. Toward an affect-sensitive autotutor. *IEEE Intelligent Systems* 22(4):53–61.

Egges, A.; Molet, T.; and Magnenat-Thalmann, N. 2004. Personalised real-time idle motion synthesis. In *PG'04: Proc.of the Computer Graphics and Applications, 12th Pacific Conf.*, 121–130.

Forsyth, E., and Martell, C. 2007. Lexical and discourse analysis of online chat dialog. In *Proc. of the First IEEE Int. Conf. on Semantic Computing*, 19–26.

Gebhard, P.; Schroeder, M.; Charfuelan, M.; Endres, C.; Kipp, M.; Pammi, S.; Rumpler, M.; and Tuerk, O. 2008. Ideas4games: Building expressive virtual characters for computer games. In *Proc.of the 8th Int. Conf. on Intelligent Virtual Agents*, LNAI 5208, 426–440. Springer.

Gobron, S.; Ahn, J.; Paltoglou, G.; Thelwall, T.; and Thalmann, D. 2010. From sentence to emotion: a real-time three-dimensional graphics metaphor of emotions extracted from text. *The Visual Computer* 26:505–519.

Gobron, S.; Ahn, J.; Quentin, S.; Thalmann, D.; Skowron, M.; Rank, S.; Paltoglou, G.; Thelwall, M.; and Kappas, A. submitted. 3D-emochatting: an interdisciplinary communication model for VR chatting. *in review process*.

Grillon, H., and Thalmann, D. 2009. Simulating gaze attention behaviors for crowds. *Comput. Animat. Virtual Worlds* 20(2–3):111–119.

Neviarouskaya, A.; Prendinger, H.; and Ishizuka, M. 2010. User study on AffectIM, an avatar-based Instant Messaging system employing rule-based affect sensing from text. *Int. Journal of Human-Computer Studies* 68(7):432–450.

Paltoglou, G.; Gobron, S.; Skowron, M.; Thelwall, M.; and Thalmann, D. 2010. Sentiment analysis of informal textual communication in cyberspace. In *Proc. Engage 2010, Springer LNCS State-of-the-Art Survey*, 13–25.

Pang, B., and Lee, L. 2004. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proc.of ACL'04*, 271–278.

Pang, B., and Lee, L. 2008. *Opinion Mining and Sentiment Analysis*. Now Publishers Inc.

Pang, B.; Lee, L.; and Vaithyanathan, S. 2002. Thumbs up? sentiment classification using machine learning techniques. In *Proc.of EMNLP 2002*, 79–86.

Pelachaud, C. 2009. Modelling multimodal expression of emotion in a virtual agent. *Philosophical Trans. of the Royal Society B: Biological Sciences* 364(1535):3539–3548.

Pennebaker, J. W.; Francis, M. E.; and Booth, R. K. 2001. *Linguistic Inquiry and Word Count: LIWC 2001*. Erlbaum Publishers.

Pennebaker, J. W.; Mehl, M. R.; and Niederhoffer, K. 2003. Psychological aspects of natural language use: our words, our selves. *Annual Review of Psychology* 54:547–577.

Petta, P.; Pelachaud, C.; and Cowie, R., eds. 2011. *Emotion-Oriented Systems: The Humaine Handbook*. Cognitive Technologies Series. Springer.

Picard, R. 1997. *Affective Computing*. MIT Press, Cambridge, MA.

Schroeder, M., and Cowie, R. 2006. Developing a consistent view on emotion-oriented computing. In Renals, S., and Bengio, S., eds., *Machine Learning for Multimodal Interaction*, 194–205. Springer.

Sebe, N.; Lew, M.; Sun, Y.; Cohen, I.; Gevers, T.; and Huang, T. 2007. Authentic facial expression analysis. *Image and Vision Computing* 25(12):1856–1863.

Skowron, M., and Paltoglou, G. 2011. Affective cues and their application in a conversational agent. In *Workshop of Affective Computational Inteligence, IEEE SSCI 2011*.

Skowron, M.; Irran, J.; and Krenn, B. 2008. Computational framework for and the realization of cognitive agents providing intelligent assistance capabilities. In *ECAI. The 6th Int. Cognitive Robotics Workshop*, 88–96.

Skowron, M. 2010. Affect listeners. acquisition of affective states by means of conversational systems. In *Development of Multimodal Interfaces - Active Listening and Synchrony*, LNCS, 169–181. Springer.

Su, W.-P.; Pham, B.; and Wardhani, A. 2007. Personality and emotion-based high-level control of affective story characters. *IEEE Transactions on Visualization and Computer Graphics* 13:281–293.

Thomas, M.; Pang, B.; and Lee, L. 2006. Get out the vote: determining support or opposition from congressional floor-debate transcripts. In *Proc.of EMNLP'06*, 327–335.

Tsetserukou, D., and Neviarouskaya, A. 2010. iFeel IM: Augmenting emotions during online communication. *IEEE Computer Graphics and Applications* 30:72–80.

Whitelaw, C.; Garg, N.; and Argamon, S. 2005. Using appraisal groups for sentiment analysis. In *Proc.of CIKM '05*, 625–631.