Exploration of Affect Sensing from Speech and Metaphorical Text

Li Zhang

School of Computing
University of Teesside, UK
l.zhang@tees.ac.uk

Abstract. We report new developments on affect detection from textual metaphorical affective expression and affect sensing from speech. The textual affect detection component has been embedded in an intelligent conversational AI agent interacting with human users under loose scenarios. The detected affective states from text also play an important role in producing emotional animation for users' avatars. Evaluation of the affect detection from speech and text is provided. Our work contributes to the conference themes on engagement and emotion, virtual AI agents, narrative storytelling in education and evaluation of affective social interaction.

Keywords: Affect sensing/detection, affective speech processing, and metaphor

1 Introduction

We intend to provide anti-bullying functionalities for online interaction via affect detection in speech and text and incorporate such functionalities with an automated intelligent conversational agent, engaged in a virtual storytelling environment with human users. In order to achieve this research goal, first of all, we have developed a textual affect detection component, EMMA (emotion, metaphor and affect) on detecting simple and complex emotions, meta-emotions, value judgments etc [1]. The work presented here reports further developments on textual affect detection for one particular metaphorical expression with affect implication, affects as physical objects metaphor ("anger ran through me", "fear attacks me") and a new development on affect sensing from informal conversational speech.

The textual affect detection component has been embedded in an intelligent conversational AI agent, engaged in a drama improvisation with human users under loose scenarios (school bullying and skin cancer). In both scenarios, the AI agent plays a minor role in drama improvisation. It plays a close friend of the bullied victim (the leading role) in school bullying scenario, who tries to stop the bullying and a close friend of the sick leading character in Skin Cancer scenario who tries to give support to his friend with the decision on his life-changing operation. The animation engine adopts the detected affect implied in users' text input to produce emotional gesture animation for the users' avatars. The conversational AI agent also provides appropriate responses based on the detected affect from users' input in order to stimulate the improvisation. The newly developed component on affect sensing from speech has not been integrated with the AI agent yet. Thus we report the initial evaluation of this component separately.

We have also analyzed textual affect detection performance based on the collected transcripts from a new round of user testing by calculating agreements via Cohen's Kappa between two human judges, human judge A/the AI agent and human judge B/the AI agent respectively.

The content is arranged in the following way. We report relevant work in section 2. In section 3, we present the further developments on textual affect and intensity detection especially for the processing of affect metaphor, and the new development of emotion recognition from speech. Brief discussion on the overall system framework and how the detected affects from user's text input contribute to the emotional animation is provided in section 4. Newly produced evaluation results of affect detection from speech and text are reported in section 5. Finally we summarize our work and point out future directions for further developments in section 6.

2 Relevant Work

Automatic affect interpretation from open-ended text and speech could be a very challenging task. Affect expression in verbal communication generally differs from culture to culture, from female to male and from one age group to another, especially for metaphorical expression of affect and emotional expression in speech. In this section, we briefly review relevant well-known research work in this research area.

Textual affect sensing is a rising research branch for natural language processing. ConceptNet [2] is a toolkit to provide practical textual reasoning for affect sensing for six basic emotions, text summarization and topic extraction. Shaikh et al. [3] provided sentence-level textual affect sensing to recognize evaluations (positive and negative). They adopted a rule-based domain-independent approach, but they haven't made attempts to recognize different affective states from open-ended text input.

Although Façade [4] included shallow natural language processing for characters' open-ended utterances, the detection of major emotions, rudeness and value judgements is not mentioned. Zhe and Boucouvalas [5] demonstrated an emotion extraction module embedded in an Internet chatting environment (see also Boucouvalas [6]). It used a part-of-speech tagger and a syntactic chunker to detect the emotional words and to analyze emotion intensity for the first person (e.g. 'I' or 'we'). Unfortunately the emotion detection focused only on emotional adjectives, and did not address deep issues such as figurative expression of emotion (discussed below). Also, the concentration purely on first-person emotions is narrow. There has been relevant work on general linguistic clues that could be used in practice for affect detection (e.g. Craggs and Wood [7]).

There is also well-known research work on the development of emotional conversational agents. Egges et al. [8] have provided virtual characters with conversational emotional responsiveness. Elliott et al. [9] demonstrated tutoring systems that reason about users' emotions. They believe that motivation and emotion play very important roles in learning. Virtual tutors have been created in a way that not only having their own emotion appraisal and responsiveness, but also understanding users' emotional states according to their learning progress. Aylett et al. [10] also focused on the development of affective behaviour planning for the synthetic characters. Cavazza et al. [11] reported a conversational agent embodied in a wireless robot to provide suggestions for users on a healthy living life-style. Hierarchical Task Networks (HTN) planner and semantic interpretation have been used in this work. The cognitive planner plays an important role in assisting with dialogue management, e.g. giving suggestions to the dialogue manager on what relevant questions should be raised to the user according to the healthy living plan currently generated. The user's response has also been adopted by the cognitive planner to influence the change of the current plan. The limitation of such planning systems is that they normally work reasonably well within the pre-defined domain knowledge, but they will strike when open-ended user input going beyond the planner's knowledge has been used intensively during interaction. The system we present here intends to deal with such challenge.

Moreover, there is also much work in the area of emotion recognition in speech. Murray & Arnott [12] have suggested five vocal effects associated with several basic emotions such as 'pitch average', 'speech rate' and 'intensity' etc from their study. Nogueiras et al. [13] have also used Hidden Markov Models to recognise emotion from speech. Their study proved that the structure of HMM was useful to capture the temporal behaviour of speech. In their work, low level features such as pitch, energy, articulation and spectral shape were employed to recognize emotion. Grimm et al. [14] have used articulatory features from speech signal and mapped them to an emotion state in a multi-dimensional, continuous-valued emotion space to recognize driver's emotional state while driving. Amir and Cohen [15] have also attempted to characterise emotion in the soundtrack of an animated film. Cichosz and Ślot [16] reported a symbol-based learning approach to classify emotion in speech. They used a binary decision tree based classifier. Emotions have been used as the nodes of the tree which were assessed by feature triplets. They have evaluated their approach using two databases of emotional speech on German and Polish. Oudeyer [17] made attempts to detect emotion from speech using genetic algorithm with a set of optimal features. Although he also made attempts in using several different machine learning approaches (such as neural networks, support vector machines etc) to perform the task, naïve bayes classifier hasn't been mentioned at all.

Our work is distinctive in the following aspects: (1) affect detection in textual metaphorical expression; (2) real-time affect sensing for basic and complex affects, meta-emotions, value judgments etc (including 25 affective states) in improvisational role-play situations from open-ended textual user input; (3) affect recognition from speech using naïve Bayes classifier; (4) and real-time simple facial and gesture emotional animation activated by the detected affective states from users' text input.

3 Affect Sensing from Text and Speech

3.1 Affect Recognition from Textual Metaphorical Expression

Affect terms have been used intensively during online interaction. Besides they have been used literally to convey users' emotional states (e.g. "I am angry", "I get bored"), affect terms have been mentioned in affective metaphorical language. One category of such metaphorical expression is 'Ideas/Emotions as Physical Objects" [18, 19], e.g. "joy ran through me", "my anger returns in a rush", "fear is killing me" etc. In these examples, emotions and feelings have been regarded as external entities. The external entities are often, or usually, physical objects or events. Therefore, affects could be treated as physical objects outside the agent in such examples, which could be active in other ways [18]. Implementation has been carried out to provide the affect detection component the ability to deal with such affect metaphor.

WordNet-affect domain (part of WordNet-domain 3.2) [20] has been used in our application. It provides an additional hierarchy of 'affective domain labels', with which the synsets representing affective concepts are further annotated. Rasp has been used to detect statements with a structure of 'a singular common noun subject + present-tense lexical verb phrase' or 'a singular common noun subject + present-tense copular form + -ing form of lexical verb phrase'. Various user inputs could possess such syntactic forms, e.g. "the girl is crying", "the big bully runs through the grass" etc. We use WordNet-affect to refine the user inputs in order to obtain metaphorical affective expression. The singular common noun subject is sent to WordNet-affect in order to obtain the hierarchical affect information. If the subject is an affective term such as 'panic', then the hierarchical affect information obtained from WordNet-affect is 'negative-fear -> negative-emotion -> emotion -> affective-state -> mental-state'. The system realizes that a mental state has been used as a subject which carries out an activity indicated by the verb phrase(s). Thus the system regards such expression as affective metaphor belonging to the category of 'affects as entities'. A further processing based on the hierarchical affect result leads to the exact affective state conveyed in user's input – fear (negative emotion). If such input has a first-person object, 'me' (such as "panic is dragging me down"), then it indicates the user currently experiences fear. Otherwise if such input has a third-person object, 'him/her' (such as "panic is sweeping over and over him"), it implies that it's not the user who currently experiences 'fear', but another character. The step-by-step analysis is listed in the following for the user input "panic is dragging me down":

1. Rasp recognizes the input with a structure of 'a singular common noun subject (panic) + present-tense copular form (is) + -ing form of lexical verb phrase (dragging) + object (me)';
2. The subject noun term, 'panic', has been sent to WordNet-affect;
3. The obtained hierarchical affect information from WordNet-affect indicates the input is interpreted as a semantic syntactic structure of 'a mental state + an activity + object (me)';
4. The user input is regarded as affect metaphor belonging to the category of 'affects as entities';
5. The detected affective state ('fear') is recovered from the hierarchical affect information;

6. Since the object is ‘me’, then the system concludes that the user is experiencing ‘fear’ implied in his/her input.

If the subject of the user input is not an affect term (e.g. “the girl is crying”, “the boy sweeps the floor”), other suitable processing methods (e.g. checking syntactic information and affect indicators etc) are adopted to extract affect. On the whole, such processing is indeed at a very initial stage. However, it provides a useful way to recognize both affect from textual user input and affect metaphor in which emotions are used as external entities.

We have also implemented procedures to detect affect from food metaphor (“X is walking meat”), animal and size metaphor (“X is a fat big pig”, “shut ur big fat mouth”). Size metaphor also plays an important role in indicating affect intensities.

3.2 Affect Recognition from Speech

Because of the online chat nature of our application, first of all, we have constructed our own specialized speech database. In our application, we mainly intend to recognize 6 basic emotions from speech: neutral, happiness, sadness, fear, anger and surprise. At the initial stage, we have adopted 10 neutral and 10 emotional informal conversational short sentences for each emotional category (some are taken from previous user testing transcripts while the others are created by the authors) for the purpose of emotional speech recording. In order to justify the articulatory features discovered for different emotional speech, we have made neutral sentences not only recorded in a neutral way, but also recorded in the other five emotional ways. Similarly, we have also recorded all affective example sentences in a neutral way so that such speech samples could assist us to remove some of the recovered features from emotional speech data mainly caused by the syllables or phonemes used in some particular speech context. Thus we have recorded 1600 utterances as training data from 10 speakers age 18 – 27 with northeast British accent using standard sound studio. Each speaker contributes 100 emotional utterances – 20 for each category (including 10 neutral sentences spoken in that particular emotional way) and 60 neutral utterances – 10 for each category. We have also recorded 120 utterances as test data set by one chosen female speaker and one male speaker (60 from each speaker).

Speech processing tool, Praat [21], has been used to analyze the speech data. First of all, for each speech sample, Praat provides an automatic summarized voice report containing detailed information on articulatory features such as pitch (median pitch, mean pitch etc), pulses, voicing, jitter, shimmer, harmonicity of the voiced parts (mean noise-to-harmonics ratio etc). After a careful study of the voice reports and other articulatory features of the emotional and non-emotional speech samples from all the speakers, we have chosen 9 articulatory features (mean pitch, median pitch, standard pitch deviation, minimum and maximum pitch, pulses per second, mean intensity, minimum and maximum intensity) to carry out further analysis for both male and female groups. A further analysis and comparison between the generated voice reports of all the speech samples from each category, we further extend the 9 articulatory features to 69 features for male speakers and 52 features for female speakers. Table 1 contains some example features recovered for mean intensity for each emotional category for the male group.

Table 1. Mean intensity value ranges for the male group

Range	Emotion	Feature
] ... , 56]	Happy Fear	Ihf
]56 , 58]	Fear Sad	Ifs
]58 , 65]	Fear Sad Angry	Ifsa
]70 , 76]	Angry Happy Surprise	Iahsu
]76 , ...]	Happy Surprise	Ihsu

In our application, naïve Bayes classifier has been used to recognize emotion from speech. Equation 1 has been used to calculate the probabilities of different emotional states for any given test speech sample. The emotional state with the highest probability is regarded as the most probable affective state implied in that instance.

$$V_{max} = \text{argmax}_{v_j \text{ in } V}\ P(v_j) * P(a_1|v_j) * P(a_2|v_j) * \ldots * P(a_n|v_j) \qquad \textbf{(1)}$$

In equation 1, a_1, a_2, … and a_n represent the articulatory features recovered for each speech training data, such as features for mean pitch, median pitch, mean intensity etc. We assume that these 9 general features are all independent. Each training speech sample is represented by the set of 9 articulatory features with different values. M-estimate has been adopted to produce the probability of an attribute value given any emotional or neutral classification. A Java application has been implemented based on the above discussion to recognize emotion from speech. The training input data file contains distinctive 477 utterances with average 82 from each emotional category and 45 neutral utterances. The 120 test speech samples have also been represented in a similar format, but with totally different sets of values of the 9 articulatory features. We report the evaluation of this affect sensing component in section 4.

4 System Architecture and Emotional Animation

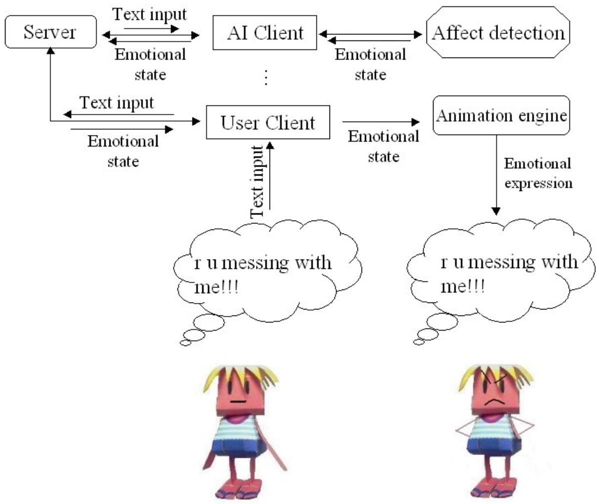

Fig. 1. An example of real-time interaction

In this section, we report the framework of our application and how to employ the detected affect from text to activate emotional animation for users' avatars.

Our system adopts client/server architecture for implementation. The conversation AI agent and other human-controlled characters consist of clients. The server broadcasts messages sent by one client to all the other clients. Thus user's text input from normal user client is sent to the AI agent client via the server. Then the AI agent, who plays a minor role in the improvisation with other human-controlled characters, analyzes the user's text input and derives the affective implication out of the text. Then the AI agent also searches its knowledge base to provide a suitable response to the human players using the detected affective states. We have particularly created the AI agent's responses in a way which could stimulate the improvisation by generating sensitive topics of the storyline. Then an XML stream composed of the detected affective state from one user input and the AI agent's response is dynamically created and broadcasted to all other clients by the server. The users' clients parse the XML stream to obtain the information of the previous "speaker's" emotional state and the current AI character's response. An animation engine has embedded in each user client which updates the user avatars' emotional facial and gesture animation on each user's terminal. Therefore, if the previous human-controlled character expresses 'anger' affective state by saying "r u messing with me!!!", the animation engine in each user client updates emotional animation of that character on each terminal using cross behavior via simple facial and gesture animation (see Fig. 1). In each session, up to four characters are engaged in. Fig. 2 displays the architecture of the overall framework.

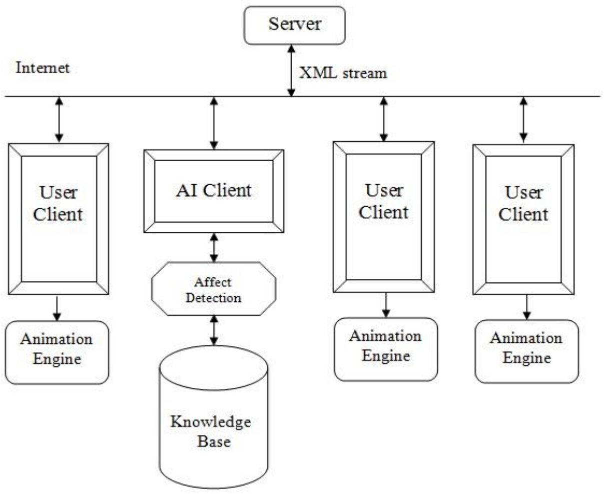

Fig. 2. The system architecture

We have adopted an approach of generating simple facial and gesture animation dynamically. We have assigned different lip shapes, eye brow shapes and arm positions dynamically to different emotional expression. Expressive animation has been considered for eight emotional states including ‘neutral’, five of Ekman’s basic emotions – ‘happy’, ‘sad’, ‘fear’, ‘angry’, ‘surprise’ – and another two complex emotions, ‘threatening’, ‘caring’, and an non-emotional state ‘greeting’. If the AI character derives an emotional state from a human-controlled character’s text input, emotional animation engine in each client updates the emotional expression of that user’ avatar on each client terminal. The overall system could provide the effects that the users’ avatars move in a way which is consistent with their emotional states implied in their text input in a real-time application. Although user avatars’ emotional animation is truly basic and simple, we obtained very positive feedback from the testing subjects based on the analysis of the post questionnaires and the discussion in the debriefing session for a new round of testing (see section 5).

Relationships between characters play an important role in how characters respond to one another. We have implemented a simple emotion appraisal model for the AI character. In the bullying scenario, there are four characters: the bully, the victim, and two close friends of the victim who try to stop the bullying. The AI agent plays a minor role – one male close friend of the bullied victim. If the AI character realizes that the bully seems being aggressive (‘rude’, ‘angry’ or ‘threatening’) implied in his text input, the AI character becomes ‘angry’ due to the fact that they have a negative relationship. Similarly, the AI character would become ‘caring’, if the bullied victim indicates ‘sad’ or ‘fear’ in the text input during the interaction. It also indicates that the AI character and the bullied victim have a positive relationship. In the meantime, the AI character’s responses also reflect its current emotional states.

5 Evaluation

Since the component of affect sensing from speech hasn't been embedded into the AI agent yet, we have evaluated its performance individually based on the testing of 120 utterances from one male and one female speaker (60 from each speaker). The emotion recognition results are in the following. For the male speaker, utterances with emotional state surprise (100%), neutral (90%), anger (60%), and fear (50%), have been recognized well. It needs improvements on the emotion recognition of utterances with sadness (10%) and happiness (20%) implication. For the female speaker, the affect sensing component performs well for the utterances with emotional state neutral (90%), sadness (60%) and happiness (60%). For the utterances with emotional state surprise (40%), fear (10%) and anger (20%), the recognition performance became worse. A further detailed analysis indicated that for both male and female speakers, an emotional sentence labeled with one negative affective state tends to be recognized to contain another negative affect implication because of the resemblance of the articulatory features in these two emotional categories. E.g. speech samples with 'sadness' implication have been mis-interpreted to contain the affective state, 'fear'. Similarly, speech data with 'happiness' implication have been mis-regarded to contain 'surprise' taste, because they have also showed much similarity as those with 'good surprise' indication. These results indicate that our affect sensing component may have extracted some underlying generalization in the recognition of the general positive and negative affective states from the training data, but further improvement is needed in order to effectively distinguish one positive/negative affective state from another. We also aim to extend the evaluation by using more speech samples from several other speakers.

Moreover, we have also conducted a new round of user testing with 40 school children age from 11 to 15 to test the overall system with the textual affect detection and the AI agent under the improvisation of school bullying (SB) and skin cancer (SC) scenarios. We have classified the testing subjects based on their age into three groups. Group 1 has the children age 11 – 13, group 2 composed of children age 13 – 14 and group 3 with school children age 14 – 15. Generally, our statistical results based on the collected questionnaires indicate that the involvement of the AI character has not made any statistically significant difference to users' engagement and enjoyment with the emphasis of users' notice of the AI character's contribution throughout.

In the following, we especially report some results on the performances of the AI character and textual affect sensing from the testing. In general, we have compared the performance of the AI minor character in SB with that of the human-controlled minor character in the same scenario. Fig. 3 & 4 respectively show to what extent the AI and human-controlled minor characters have said strange things and to what extent the AI and human-controlled minor characters have contributed to the improvisation usefully based on the analysis of the questionnaires filled up by the three groups of young people respectively. The AI agent seemed haven't impressed the first group users with the scores for 'useful contribution' AI vs Human, 57% - 63% and the scores for 'saying strange things' AI vs Human, 51% - 47%. For the other two groups of users, the AI character has scored slightly better than the human-controlled minor character with the scores for 'useful contribution' AI vs Human respectively, 56% - 47% (Group 2) and 60% - 53% (Group 3) and the scores for 'saying strange things' AI vs Human respectively, 42% - 60% (Group 2) and 53% - 67% (Group 3). Some pupils from Group 3 expressed they felt more relaxed when doing improvisation online than in real-life situations. The group 1 users stated that they felt the scenarios (especially, the skin cancer scenario) were too heavy for them. They indicated that they preferred a light-hearted scenario embedded with entertainment.

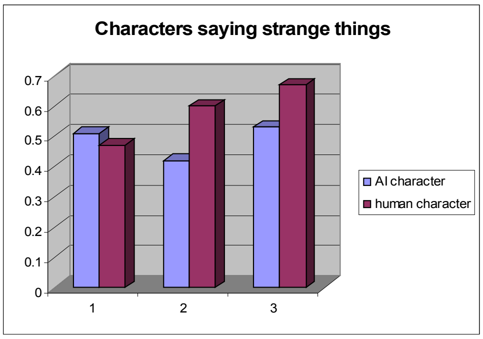

Fig. 3. The comparison of characters saying strange things during the improvisation between the AI and the human-controlled minor characters for the 3 groups

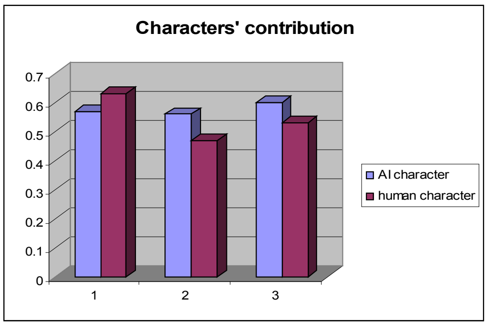

Fig. 4. The comparison of useful contribution to the improvisation between the AI and the human-controlled minor characters for the 3 groups

Moreover, we have provided *Cohen's Kappa* in order to evaluate the efficiency of the textual affect detection processing for the detection of 25 affective states, although simple emotional facial and gesture animation could only deal with 8 particular emotional states. Two human judges (not involved in any development stage) have been employed to annotate part of the recorded transcripts of the SB scenario (72 turn-taking user input) filed from the testing. The inter agreement between human judge A and B is 0.896. The agreements for human judge A/the AI agent and human judge B/the AI agent are respectively 0.662 and 0.729. Although improvement is needed, the AI agent's affect detection performance is acceptable and could achieve satisfactory level in good cases. Analysis results also indicate improvement is needed for negative affect detection (e.g. using context information). In some cases, when the two human judges both believed that user inputs carried negative affective states (such as angry, threatening, disapproval etc), the AI agent regarded them as neutral. The most obvious reason is that the context information used by the human judges to interpret emotions has been discarded by the AI agent due to the fact that our current processing is only based on the input of individual turn taking level rather than context level. However, an individual user input, regarded as neutral by itself in most cases by all human judges, could be interpreted as emotional with the consideration of the context profiles. Thus we aim to improve the detection performance by adopting context profile as one direction for future development for textual affect sensing.

6 Conclusions

First of all, we have made a step towards automatic affect sensing from textual metaphorical figurative language. However, there is still a long way to go in order to successfully process the rich diverse variations of metaphorical language and other figurative expressions, such as humor, lies, irony etc. Also, context information sometimes is very crucial for textual affect detection. These indicate in which our strength needs to lie in the future development.

We have also implemented a prototype system for affect sensing from speech using naïve bayes classifier. Although there is room for further improvements, the current performance of the affect sensing component is acceptable and promising. We intend to integrate this component with another intelligent conversational agent who interacts with human users during online speech based interaction so that the intelligent agent would be capable of detecting bullying or other emotional situations automatically from users' speech via the affect sensing component reported here.

Overall, our work provides automatic improvisational agents for virtual drama improvisation situations. It makes a contribution to the issue of what types of automation should be included in human-agent interaction, and as part of that the issue of what types of affect in speech and text should be detected and how. It also provides an opportunity for the developers to explore how emotional issues embedded in the scenarios, characters and dialogue can be represented visually without detracting users from the learning situation. Finally, the automated conversational AI agent and the emotional animation may contribute to improving the perceived quality of social interaction.

References

1. Zhang, L., Barnden, J.A., Hendley, R.J., Lee, M.G., Wallington, A.M. and Wen, Z. 2008. Affect Detection and Metaphor in E-drama. Int. J. Continuing Engineering Education and Life-Long Learning, Vol. 18, No. 2, 234-252.
2. Liu, H. & Singh, P. 2004. ConceptNet: A practical commonsense reasoning toolkit. BT Technology Journal, Volume 22, Kluwer Academic Publishers.

3. Shaikh, M. A. M., Prendinger, H. & Mitsuru, I. 2007. Assessing sentiment of text by semantic dependency and contextual valence analysis. In Proceeding of ACII 2007, 191-202.
4. Mateas, M. 2002. Ph.D. Thesis. Interactive Drama, Art and Artificial Intelligence. School of Computer Science, Carnegie Mellon University.
5. Zhe, X. & Boucouvalas, A. C. 2002. Text-to-Emotion Engine for Real Time Internet Communication. In Proceedings of International Symposium on Communication Systems, Networks and DSPs, Staffordshire University, UK, 164-168.
6. Boucouvalas, A. C. 2002. Real Time Text-to-Emotion Engine for Expressive Internet Communications. In Being There: Concepts, Effects and Measurement of User Presence in Synthetic Environments. G. Riva, F. Davide and W. IJsselsteijn (eds.), 305-318.
7. Craggs, R. & Wood. M. 2004. A Two Dimensional Annotation Scheme for Emotion in Dialogue. In Proceedings of AAAI Spring Symposium: Exploring Attitude and Affect in Text.
8. Egges, A., Kshirsagar, S. & Magnenat-Thalmann, N. 2003. A Model for Personality and Emotion Simulation, In Proceedings of Knowledge-Based Intelligent Information & Engineering Systems (KES2003), Lecture Notes in AI. Springer-Verlag: Berlin, 453-461.
9. Elliott, C., Rickel, J. & Lester, J. 1997. Integrating Affective Computing into Animated Tutoring Agents. In Proceedings of IJCAI'97 Workshop on Intelligent Interface Agents, 113-121.
10. Aylett, R., Louchart, S., Dias, J., Paiva, A., Vala M., Woods, S. Hall, L. E. 2006. Unscripted Narrative for Affectively Driven Characters. IEEE Computer Graphics and Applications 26(3). 42-52.
11. Cavazza, M., Smith, C., Charlton, D., Zhang, L., Turunen, M. and Hakulinen, J. 2008. A 'Companion' ECA with Planning and Activity Modelling. In Proceedings of the 7th International Conference on Autonomous Agents and Multi-Agent Systems. Portugal, 1281-1284.
12. Murray, I.R. and Arnott, J.L. Implementation and testing of a system for producing emotion-by-rule in synthetic speech. Speech Communication (16), 369-390, 1995.
13. Nogueiras, A., Moreno, A., Bonafante, A. and Maririo. J. Speech Emotion Recognition Using Hidden Markov Models, Eurospeech 2001, 2679-2682, 2001.
14. Grimm, M., Kroschel, K., Harris, H., Nass, C., Schuller, B., Rigoll, G. and Moosmayr, T. On the Necessity and Feasibility of Detecting a Driver's Emotional State While Driving. In Proceedings of ACII 2007, Lisbon, ACM, Springer, 126-138, 2007.
15. Amir, N and Cohen. R. Characterizing Emotion in the Soundtrack of an Animated Film: Credible or Incredible? In Proceedings of ACII 2007, Lisbon, ACM, Springer, 148-158, 2007.
16. Cichosz, J. and Slot, K. Emotion recognition in speech signal using emotion-extracting binary decision trees. Doctoral Consortium. ACII 2007, Lisbon, ACM, Springer, 2007.
17. Oudeyer, P.Y. The production and recognition of emotions in speech: features and algorithms, International Journal in Human-Computer Studies, vol. 59/1-2, 157-183, Special Issue on Affective Computing. 2003.
18. ATT-Meta Project Databank: Examples of Usage of Metaphors of Mind. http://www.cs.bham.ac.uk/~jab/ATT-Meta/Databank/. July 2008.
19. Kövecses, Z. 1998. Are There Any Emotion-Specific Metaphors? In Speaking of Emotions: Conceptualization and Expression. Athanasiadou, A. and Tabakowska, E. (eds.), Berlin and New York: Mouton de Gruyter, 127-151.
20. Strapparava, C. and Valitutti, A. 2004. WordNet-Affect: An Affective Extension of WordNet, In Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004), Lisbon, Portugal, 1083-1086.
21. Praat, a speech processing tool. http://www.personal.rdg.ac.uk/~llsroach/phon2/freespeech.htm, 2008.