Affect Detection Exploration on Metaphor

Li Zhang
School of Computing
University of Teesside, UK
+44 (0) 1642 218121- 3719
l.zhang@tees.ac.uk

ABSTRACT

We report new developments of an affect detection component on the processing of several different types of metaphorical affective expression. The component has been embedded in a conversational AI agent interacting with human users under loose scenarios. Evaluation for the affect detection component is also provided. Our work contributes to the workshop themes on affect recognition via text and evaluation of affective interaction.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing – *language parsing and understanding, discourse and text analysis*

General Terms

Algorithms, Measurement, Experimentation, Human Factors

Keywords

Metaphorical affective language, affect detection

1. INTRODUCTION

We have reported an affect detection component on detecting simple and complex emotions, meta-emotions, etc in our previous work [1]. The work presented here reports further developments on affect detection for several particular metaphorical expressions with affect implication, which include affect as physical objects metaphor ("anger ran through me"), food metaphor ("X is walking meat"), animal and size metaphor. We report affect and food metaphor particularly in detail. The affect detection component has been embedded in a conversational AI agent, engaged in a drama improvisation with human users under loose scenarios (school bullying and Crohn's disease). We have also analyzed affect detection performance based on the collected transcripts from 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.

2. RELEVANT WORK

Textual affect sensing is a rising research branch for natural language processing. ConceptNet [2] is a toolkit to provide practical textual reasoning for six basic emotions. 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 in real-time application.

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 analyse 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.

Our work is distinctive in the following aspects: (1) affect detection in metaphorical expression; (2) real-time affect sensing for basic and complex affects in improvisational role-play situations; (3) and emotional animation activated by the detected affective states.

3. FURTHER DEVELOPMENT

3.1 Affect Detection on Affect Metaphor

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 [12]. One category of such metaphorical expression is 'Ideas/Emotions as Physical Objects" [13], e.g. "joy ran through me", "my anger returned in a rush" 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. 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) [14] 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'. Various user inputs could possess such syntactic forms, e.g. "the girl cries", "the big bully is running 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 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 sucks 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 cries", "the boy is sweeping 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 could not only be useful to detect affect from user input, but also provide a useful way to recognize affect metaphor in which emotions are used as external entities.

3.2 Affect Detection on Food Metaphor

Food has been used extensively as metaphor for social position, group identity, religion, etc. E.g. food could be used as a metaphor for national identity. British have been called 'roastbeefs' by the French, while French have been referred to as 'frogs' by the British. It has also been used to indicate social hierarchy. In our school bullying scenario, the big bully has called the bullied victim (Lisa) names, such as "u r a pizza", "Lisa has a pizza face" to exaggerate that fact that the victim has acne. Another most commonly used food metaphor is to use food to refer to a specific shape. E.g. body shape could be described as 'banana', 'pear' and 'apple' (http://jsgfood.blogspot.com/2008/02/food-metaphors.html). In our application, "Lisa has a pizza face" could also be interpreted as Lisa has a 'round (shape)' face. Therefore, insults could be conveyed in such food metaphorical expression. We especially focus on the statement of 'second-person/a singular proper noun + present-tense copular form + food term' to extract affect. A special semantic dictionary has been created by providing semantic tags to normal English lexicon. The semantic tags have been created by using Wmatrix [15], which facilitates the user to obtain corpus annotation with semantic and part-of-speech tags to compose dictionary. The semantic dictionary created consists mainly of food terms, animal names, measureable adjectives (such as size) etc with their corresponding semantic tags due to the fact they have the potential to convey affect and feelings.

In our application, rasp informs the system the user input with the desired structure - 'second-person/a singular proper noun + present-tense copular form + noun phrases' (e.g. "Lisa is a pizza", "u r a hard working man", "u r a peach"). The noun phrases are examined in order to recover the main noun term. Then its corresponding semantic tag is derived from the composed semantic dictionary if it is a food term, or an animal-name etc. E.g. "u r a peach" has been regarded as "second-person + present-tense copular form + [food-term]". WordNet has been employed in order to get the synset of the food term. If among the synset, the food term has been explained as a certain type of human being, such as 'beauty', 'sweetheart' etc. Then another small slang-semantic dictionary collected in our previous study containing terms for special person types (such as 'freak', 'angle') and their corresponding evaluation values (negative or positive) has been adopted in order to obtain the evaluation value of such synonyms. If the synonyms are positive ('beauty'), then we conclude that the input is affectionate expression with a food metaphor (e.g. "u r a peach"). The processing procedures are listed in the following for the input "u r a peach":

1. Rasp recognizes the input has a structure of 'second-person (you) + present-tense copular form (are) + noun phrases (a peach)';
2. The newly composed semantic dictionary is used to obtain the semantic tag for the main noun term ('peach') in the noun phrase, 'peach' -> 'food term';
3. Thus the input is regarded as 'second-person (you) + present-tense copular form (are) + [food-term: peach]'. The system interprets that the input is a food metaphor;

4. The food term, ‘peach’, is sent to WordNet to get its synonyms. The synonyms include special person types, such as ‘beauty’ and ‘sweetheart’;
5. The special slang-semantic dictionary containing special person types and their corresponding evaluation values is used to obtain the evaluation values (positive) of special person types ‘beauty’ and ‘sweetheart’;
6. Since the evaluation values are positive, the system concludes the input expresses affectionate with a food metaphor.

However, in most of the cases, WordNet doesn’t provide any description of types of human beings when explaining a food term (e.g. ‘pizza’, ‘meat’ etc). According to the nature of the scenarios (e.g. bullying) we used, we simply conclude that the input implies insulting with a food metaphor when calling someone food terms (“u r walking meat”, “Lisa is a pizza”).

Another interesting phenomenon drawing our attention is food as shape metaphor. As mentioned earlier, food is often used as a metaphor to refer to body shapes (e.g. “you have a pear body shape”, “Lisa has a garlic nose”, “Lisa has a pizza face”). They might indicate literal truth, but most of which are potentially used to indicate very unpleasant truth. Thus they could be regarded as insulting. We extend our semantic dictionary created with the assistance of Wmatrix by adding terms of physiological human body parts, such as face, nose, body etc. For the user’s input with a structure of ‘second-person/a singular proper noun + have/has + noun phrases’ informed by rasp, the system provides a semantic tag for each word in the object noun phrase. If the semantic tag sequence of the noun phrase indicates that it consists of a food term followed by a physiological term (‘pizza face’), the system interprets that the input implies insulting with a food metaphor.

However, examples, such as “you have a banana body shape” and “you are a meat and potatoes man”, haven’t been used to express insults, but instead the former used to indicate a slim body and the latter to indicate a hearty appetite and robust character. Other examples such as “you are what you eat” could be very challenging theoretically and practically. They also indicate the direction of the future extension of our current system.

3.3 Other Processing on Metaphors

We have also implemented procedures to detect affects from animal and size metaphor (e.g. “u r a pig”, “u r a big idiot”, “shut ur big fat mouth”). We have reported animal metaphor processing using WordNet and a semantic profile developed by Esuli and Sebastiani [16] in our previous work [17]. Briefly we mainly intend to provide automatic processing on the user input with a structure of “second-person/a singular proper noun + present-tense copular form + [animal-name]”, which could convey affectionate (“u r a lion”) or insults (“Lisa is a pig”). WordNet has been used to analyse the animal name. If WordNet provides an ‘adjective + noun’ description of the characteristics of a person/woman/man (e.g. “a famous man” and “a disgraceful woman”) as one interpretation of the animal name, then Esuli and Sebastiani’s semantic profile is used to obtain the evaluation value of the adjective. If it’s positive (“a famous man”), then the user input is metaphorical affectionate expression, otherwise (“a disgraceful woman”) it is a metaphorical insulting expression. However, as stated in the food metaphor, WordNet has rarely provided descriptions of the characteristics of a human being as interpretation of an animal name, but only for those animal names with strong affect implication in culture background. Therefore, we have included animation names with corresponding semantic labels indicating if they are young or adult animal names in the created semantic dictionary mentioned above. Thus if WordNet strikes and the user input contains young animation names (such as “bunny”), then the user input probably implies affectionate since in common sense, calling someone a young animal name usually expresses affectionate. Also, we have adopted size metaphor (“u r a big fat pig (strong insulting)”, “u r a little idiot (weak insulting)”) [18] to recover intensities of emotions expressed.

Although the further processing only deals with four different types of metaphor, it could point out a good direction for affect detection in metaphorical language.

4. EMOTIONAL ANIMATION

The affect detection component has been integrated with a conversational AI agent who plays a minor character and interacts with human-controlled characters in role-play situations and the detected affect has been sent to the animation engine in order to provide real-time emotional gesture for human-controlled characters. The AI agent also intends to provide appropriate responses based on the detected affect from user inputs to simulate the improvisation.

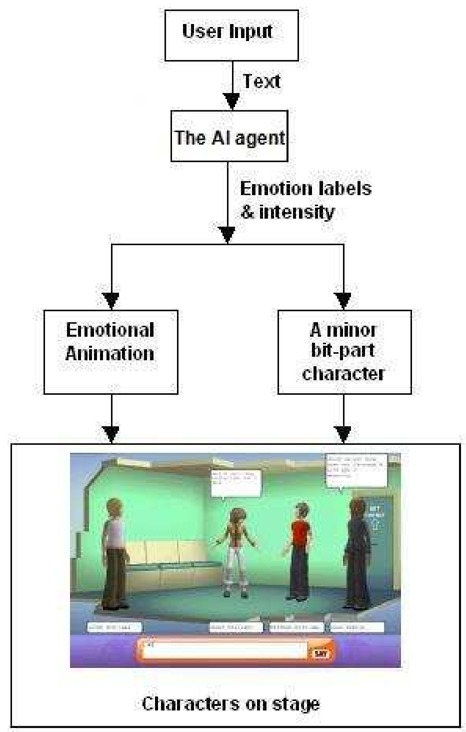

Figure 1. Affect detection and the control of characters

The detected affective states from users’ open-ended text input play an important role in producing emotional animation of human players’ avatars. The emotional animation mainly includes emotional gesture and social attention (such as eye gazing). The expressive animation engine, Demeanour [19], makes it possible for human-controlled characters to express the affective states detected by the conversational AI agent. When the AI agent detects an affective state in a user's text input, this is passed to the Demeanour system attached to this user's character and a suitable emotional animation is produced.

Figure 1 gives an overview of the control of the expressive characters. Users' text input is analyzed by the AI agent in order to detect affect in the text. The output is an emotion label with intensity derived from the text. This is then used in two ways. Firstly it is used by the minor bit-part character (played by the AI agent) to generate a response. Secondly the label and the intensity are sent to the emotional animation system (via an XML stream) where it is used to generate animation. More discussion on emotional animation could be found in Zhang et al. [17].

5. EVALUATION AND CONCLUSIONS

We carried out user testing with 180 secondary school students for school bullying and Crohn's disease scenarios. Generally, the statistical results based on the collected questionnaires indicate that the involvement of the AI agent has not made any statistically significant difference to users' enjoyment etc with the emphasis of users' notice of the AI agent's contribution. Two human judges (not engaged in any development) marked up affect of the users' input in some recorded transcripts for both scenarios in order to verify the efficiency of the affect detection component. *Cohen's Kappa* has been provided to compare the agreements for the detection of 25 affective states. The agreement for human judge A/B is 0.38. The poor human/human agreement also indicated the annotation of recorded transcripts using 25 emotional states is highly ambiguous and even human judges may interpret the emotions expressed in the same role-play situations differently. The values for human judge A/the AI agent, human judge B/the AI agent are respectively 0.38 and 0.30. Although a good agreement is within 0.6-0.8, the performance of the automatic affect detection is acceptable and could achieve human-to-human level agreement 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. One 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. Moreover, analogies and metaphors could also be used to convey humor. Thus a potential insulting phrase may become a joke or a positive message. Such expression will challenge the current affect detection component, but it also points out another interesting direction for future development.

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-robots interaction, and as part of that the issue of what types of affect 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.

6. REFERENCES

[1] Zhang, L., Barnden, J.A., Hendley, R.J., et al. 2008. Affect detection and metaphor in E-drama. Int. J. Continuing Engineering Education and Life-Long Learning, Vol. 18, No. 2, pp.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. ACII 2007. pp.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, pp 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.

[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.

[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) pp. 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. May 2008. Portugal.

[12] 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.

[13] ATT-Meta Project Databank: Examples of Usage of Metaphors of Mind. http://www.cs.bham.ac.uk/~jab/ATT-Meta/Databank/. July 2008.

[14] Strapparava, C. and Valitutti, A. 2004. WordNet-Affect: An Affective Extension of WordNet, 1083–1086, In Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004), Lisbon, Portugal, 2004.

[15] Rayson, P. 2003. Matrix: A statistical method and software tool for linguistic analysis through corpus comparison. Ph.D. thesis, Lancaster University.

[16] Esuli1, A. and Sebastiani, F. 2006. Determining Term Subjectivity and Term Orientation for Opinion Mining. In Proceedings of EACL-06, 11th Conference of the European Chapter of the Association for Computational Linguistics, Trento, IT. pp. 193-200.

[17] Zhang, L., Gillies, M. & Barnden, J.A. 2008. EMMA: an Automated Intelligent Actor in E-drama. In Proceedings of International Conference on Intelligent User Interfaces. 13th –16th Jan 2008. Canary Islands, Spain.

[18] Sharoff, S. 2005. How to Handle Lexical Semantics in SFL: a Corpus Study of Purposes for Using Size Adjectives. Systemic Linguistics and Corpus. London: Continuum.

[19] Gillies, M. & Ballin, D. 2004. Integrating autonomous behavior and user control for believable agents. In Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems. Columbia University, New York, July 2004.