

Developments in Affect Detection in E-drama

Li Zhang, John A. Barnden, Robert J. Hendley, and Alan M. Wallington

School of Computer Science
University of Birmingham, UK

l.zhang@cs.bham.ac.uk

Abstract

We report work¹ in progress on adding affect-detection to an existing program for virtual dramatic improvisation, monitored by a human director. To partially automate the directors' functions, we have partially implemented the detection of emotions, etc. in users' text input, by means of pattern-matching, robust parsing and some semantic analysis. The work also involves basic research into how affect is conveyed by metaphor.

1 Introduction

Improvised drama and role-play are widely used in education, counselling and conflict resolution. Researchers have explored frameworks for e-drama, in which virtual characters (avatars) interact under the control of human actors. The springboard for our research is an existing system (*edrama*) created by Hi8us Midlands Ltd, used in schools for creative writing and teaching in various subjects. The experience suggests that e-drama helps students lose their usual inhibitions, because of anonymity etc. In *edrama*, characters are completely human-controlled, their speeches textual in speech bubbles, and their visual forms cartoon figures. The actors (users) are given a loose scenario within which to improvise, but are at liberty to be creative. There is also a human director, who constantly monitors the unfolding drama and can intervene by,

for example, sending messages to actors, or by introducing and controlling a minor 'bit-part' character to interact with the main characters. But this places a heavy burden on directors, especially if they are, for example, teachers and unpracticed in the directorial role. One research aim is thus partially to automate the directorial functions, which importantly involve affect detection. For instance, a director may intervene when emotions expressed or discussed by characters are not as expected. Hence we have developed an affect-detection module. It has not yet actually been used for direction, but instead to control a simple automated bit-part actor, EmEliza. The module identifies affect in characters' speeches, and makes appropriate responses to help stimulate the improvisation. Within affect we include: basic and complex emotions such as anger and embarrassment; meta-emotions such as desiring to overcome anxiety; moods such as hostility; and value judgments (of goodness, etc.). Although merely detecting affect is limited compared to extracting full meaning, this is often enough for stimulating improvisation.

Much research has been done on creating affective virtual characters in interactive systems. Emotion theories, particularly that of Ortony et al. (1988; OCC in the following), have been used widely. Prendinger & Ishizuka (2001) used OCC to reason about emotions. Mehdi et al. (2004) used OCC to generate emotional behaviour. Gratch and Marsella's (2004) model reasons about emotions. However, few systems are aimed at detecting affect as broadly as we do and in open-ended utterances. Although Façade (Mateas, 2002) included processing of open-ended utterances, the broad detection of emotions, rudeness and value judgments is not covered. Zhe & Boucouvalas (2002) demonstrated emotion extraction using a tagger and a chunker to help detect the speaker's own emotions. But it focuses only on emotional adjectives, considers only first-person emotions and

¹ This work is supported by grant RES-328-25-0009 from the ESRC under the ESRC/EPSRC/DTI "PACCIT" programme. We are grateful to Hi8us Midlands Ltd, Maverick Television Ltd, BT, and our colleagues W.H. Edmondson, S.R. Glasbey, M.G. Lee and Z. Wen. The work is also partially supported by EPSRC grant EP/C538943/1.

neglects deep issues such as figurative expression. Our work is distinctive in several respects. Our interest is not just in (a) the positive first-person case: the affective states that a virtual character X implies that it *has* (or *had* or *will have*, etc.), but also in (b) affect that X implies it *lacks*, (c) affect that X implies that *other* characters have or lack, and (d) *questions, commands, injunctions*, etc. concerning affect. We aim also for the software to cope partially with the important case of metaphorical conveyance of affect (Fussell & Moss, 1998; Kövecses, 1998).

Our project does not involve using or developing deep, scientific models of how emotional states, etc., function in cognition. Instead, the deep questions investigated are on linguistic matters such as the metaphorical expression of affect. Also, in studying how people understand and talk about affect, what is of prime importance is their *common-sense* views of how affect works, irrespective of scientific reality. Metaphor is strongly involved in such views.

2 A Preliminary Approach

Various characterizations of emotion are used in emotion theories. The OCC model uses emotion labels and intensity, while Watson and Tellegen (1985) use positive and negative affects as the major dimensions. Currently, we use an evaluation dimension (positive and negative), affect labels and intensity. Affect labels with intensity are used when strong text clues signalling affect are detected, while the evaluation dimension with intensity is used when only weak text clues are detected.

2.1 Pre-processing Modules

The language in the speeches created in e-drama sessions, especially by excited children, severely challenges existing language-analysis tools if accurate semantic information is sought. The language includes misspellings, ungrammaticality, abbreviations (such as in texting), slang, use of upper case and special punctuation (such as repeated exclamation marks) for affective emphasis, repetition of letters or words for emphasis, and open-ended onomatopoeic elements such as “grrr”. The genre is similar to Internet chat.

To deal with the misspellings, abbreviations and onomatopoeia, several pre-processing modules are used before the detection of affect starts using pattern matching, syntactic processing by means of the Rasp parser (Briscoe & Carroll, 2002), and subsequent semantic processing.

A lookup table has been used to deal with abbreviations e.g. ‘im (I am)’, ‘c u (see you)’ and ‘l8r (later)’. It includes abbreviations used in Internet chat rooms and others found in an analysis of previous edrama sessions. We handle ambiguity (e.g., “2” (to, too, two) in “I’m 2 hungry 2 walk”) by considering the POS tags of immediately surrounding words. Such simple processing inevitably leads to errors, but in evaluations using examples in a corpus of 21695 words derived from previous transcripts we have obtained 85.7% accuracy, which is currently adequate.

The iconic use of word length (corresponding roughly to imagined sound length) as found both in ordinary words with repeated letters (e.g. ‘seeeee’) and in onomatopoeia and interjections, (e.g. ‘wheee’, ‘grr’, ‘grrrrr’, ‘agh’, ‘aaagghhh’) normally implies strong affective states. We have a small dictionary containing base forms of some special words (e.g. ‘grr’) and some ordinary words that often have letters repeated in e-drama. Then the Metaphone spelling-correction algorithm, which is based on pronunciation, works with the dictionary to locate the base forms of words with letter repetitions.

Finally, the Levenshtein distance algorithm with a contemporary English dictionary deals with misspelling.

2.2 Affect Detection

In the first stage after the pre-processing, our affect detection is based on textual pattern-matching rules that look for simple grammatical patterns or phrasal templates. Thus keywords, phrases and partial sentence structures are extracted. The Jess rule-based Java framework is used to implement the pattern/template-matching rules. This method has the robustness to deal with ungrammatical and fragmented sentences and varied positioning of sought-after phraseology, but lacks other types of generality and can be fooled by suitable syntactic embedding. For example, if the input is “I doubt she’s really angry”, rules looking for anger in a simple way will output incorrect results.

The transcripts analysed to inspire our initial knowledge base and pattern-matching rules had independently been produced earlier from edrama improvisations based on a school bullying scenario. We have also worked on another, distinctly different scenario concerning a serious disease, based on a TV programme produced by Maverick Television Ltd. The rule sets created for one scenario have a useful degree of applicability to another, although some changes in the specific

knowledge database will be needed.

As a simple example of our pattern-matching, when the bully character says “Lisa, you Pizza Face! You smell”, the module detects that he is insulting Lisa. Patterns such as ‘you smell’ have been used for rule implementation. The rules work out the character’s emotions, evaluation dimension (negative or positive), politeness (rude or polite) and what response EmEliza might make. Although the patterns detected are based on English, we would expect that some of the rules would require little modification to apply to other languages.

Multiple exclamation marks and capitalisation of whole words are often used for emphasis in e-drama. If exclamation marks or capitalisation are detected, then emotion intensity is deemed to be comparatively high (and emotion is suggested even without other clues).

A reasonably good indicator that an inner state is being described is the use of ‘I’ (see also Craggs and Wood (2004)), especially in combination with the present or future tense. In the school-bullying scenario, when ‘I’ is followed by a future-tense verb, a threat is normally being expressed; and the utterance is often the shortened version of an implied conditional, e.g., “I’ll scream [if you stay here].” When ‘I’ is followed by a present-tense verb, other emotional states tend to be expressed, as in “I want my mum” and “I hate you”.

Another useful signal is the imperative mood, especially when used without softeners such as ‘please’: strong emotions and/or rude attitudes are often being expressed. There are common imperative phrases we deal with explicitly, such as “shut up” and “mind your own business”. But, to go beyond the limitations of the pattern matching we have done, we have also used the Rasp parser and semantic information in the form of the semantic profiles for the 1,000 most frequently used English words (Heise, 1965).

Although Rasp recognizes many simple imperatives directly, it can parse some imperatives as declaratives or questions. Therefore, further analysis is applied to Rasp’s syntactic output.

For example, if the subject of an input sentence is ‘you’ followed by certain special verbs or verb phrases (e.g. ‘shut’, ‘calm’, ‘get lost’, ‘go away’, etc), and Rasp parses a declarative, then it will be changed to imperative. If the softener ‘please’ is followed by a base forms of the verb, the inputs are also deemed to be imperatives. If a singular proper noun or ‘you’ is followed by a base form of the verb, the sentence is deemed to be imperative (e.g. “Dave bring me the menu”).

When ‘you’ or a singular proper noun is fol-

lowed by a verb whose base form equals its past tense form, ambiguity arises (e.g. “Lisa hit me”). For one special case of this, if the direct object is ‘me’, we exploit the evaluation value of the verb from Heise’s (1965) semantic profiles. Heise lists values of evaluation (goodness), activation, potency, distance from neutrality, etc. for each word covered. If the evaluation value for the verb is negative, then the sentence is probably not imperative but a declarative expressing a complaint (e.g. “Mayid hurt me”). If it has a positive value, then other factors suggesting imperative are checked in this sentence, such as exclamation marks and capitalizations. Previous conversation is checked to see if there is any recent question sentence toward the speaker. If so, then the sentence is taken to be declarative.

There is another type of sentence: ‘don’t you + (base form of verb)’, which is often a negative version of an imperative with a ‘you’ subject (e.g. “Don’t you call me a dog”). Normally Rasp regards such strings as questions. Further analysis has also been implemented for such sentence structure, which implies negative affective state, to change the sentence type to imperative.

Aside from imperatives, we have also implemented simple types of semantic extraction of affect using affect dictionaries and WordNet.

3 Metaphorical Expression of Affect

The *explicit* metaphorical description of emotional states is common and has been extensively studied (Fussell & Moss, 1998). Examples are “He nearly exploded”, and “Joy ran through me.” Also, affect is often conveyed *implicitly* via metaphor, as in “His room is a cess-pit”, where affect associated with a source item (cess-pit) is carried over to the corresponding target item.

Physical size is often metaphorically used to emphasize evaluations, as in “you are a big bully”, “you’re a big idiot”, and “you’re just a little bully”, although the bigness may be literal as well. “Big bully” expresses strong disapproval (Sharoff, 2005) and “little bully” can express contempt, although “little” can also convey sympathy. Such examples are not only practically important but also theoretically challenging.

We have also encountered quite creative use of metaphor in e-drama. For example, in a school-bullying improvisation that occurred, Mayid had already insulted Lisa by calling her a ‘pizza’, developing a previous ‘pizza-face’ insult. Mayid then said “I’ll knock your topping off, Lisa” – a theoretically intriguing spontane-

ous creative elaboration of the ‘pizza’ metaphor. Our developing approach to metaphor handling in the affect detection module is partly to look for stock metaphorical phraseology and straightforward variants of it, and partly to use a simple version of the more open-ended, reasoning-based techniques taken from the ATT-Meta project (Barnden et al., 2002; 2003; 2004). ATT-Meta includes a general-purpose reasoning engine, and can potentially be used to reason about emotion in relation to other factors in a situation. In turn, the realities of metaphor usage in e-drama sessions are contributing to our basic research on metaphor processing.

4 Conclusion

We have implemented a limited degree of affect-detection in an automated actor by means of pattern-matching, robust parsing and some semantic analysis. Although there is a considerable distance to go in terms of the practical affect-detection that we plan to implement, the already implemented detection is able to cause reasonably appropriate contributions by the automated character. We have conducted a two-day pilot user test with 39 secondary school students. We concealed the involvement of an earlier version of EmEliza in some sessions, in order to test by questionnaire whether its involvement affects user satisfaction, etc. None of the measures revealed a significant effect. Also, judging by the group debriefing sessions after the e-drama sessions, nobody found out that one bit-part character was sometimes computer-controlled. Further user testing with students at several Birmingham schools will take place in March 2006.

References

Barnden, J.A., Glasbey, S.R., Lee, M.G. & Wallington, A.M., 2002. Reasoning in metaphor understanding: The ATT-Meta approach and system. In *Proceedings of the 19th International Conference on Computational Linguistics*.

Barnden, J.A., Glasbey, S.R., Lee, M.G. & Wallington, A.M., 2003. Domain-transcending mappings in a system for metaphorical reasoning. In *Proceedings of the Research Note Sessions of the 10th Conference of EACL*.

Barnden, J.A., Glasbey, S.R., Lee, M.G. & Wallington, A.M. 2004. Varieties and Directions of Inter-domain Influence in Metaphor. *Metaphor and Symbol*, 19(1), pp.1-30.

Briscoe, E. & J. Carroll. 2002. Robust Accurate Sta-

tistical Annotation of General Text. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, Las Palmas, Gran Canaria. pp.1499-1504.

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*.

Fussell, S. & Moss, M. 1998. Figurative Language in Descriptions of Emotional States. In S. R. Fussell and R. J. Kreuz (Eds.), *Social and cognitive approaches to interpersonal communication*. Lawrence Erlbaum.

Gratch, J. & Marsella, S. 2004. A Domain-Independent Framework for Modeling Emotion. *Journal of Cognitive Systems Research*. Vol 5, Issue 4, pp.269-306.

Heise, D. R. 1965. Semantic Differential Profiles for 1,000 Most Frequent English Words. *Psychological Monographs* 79, pp.1-31.

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, pp.127-151.

Mateas, M. 2002. Ph.D. Thesis. Interactive Drama, Art and Artificial Intelligence. School of Computer Science, Carnegie Mellon University.

Mehdi, E.J., Nico P., Julie D. & Bernard P. 2004. Modeling Character Emotion in an Interactive Virtual Environment. In *Proceedings of AISB 2004 Symposium: Motion, Emotion and Cognition*. Leeds, UK.

Ortony, A., Clore, G.L. & Collins, A. 1988. *The Cognitive Structure of Emotions*. CUP

Prendinger, H. & Ishizuka, M. 2001. Simulating Affective Communication with Animated Agents. In *Proceedings of Eighth IFIP TC.13 Conference on Human-Computer Interaction*, Tokyo, Japan, pp.182-189.

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.

Watson, D. & Tellegen, A. 1985. Toward a Consensual Structure of Mood. *Psychological Bulletin*, 98, pp.219-235.

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.