Exploitation in Affect Detection in Open-Ended Improvisational Text

Li Zhang, John A. Barnden, Robert J. Hendley and Alan M. Wallington
School of Computer Science
University of Birmingham
Birmingham B15 2TT, UK
`l.zhang@cs.bham.ac.uk`

Abstract

We report progress on adding affect-detection to a program for virtual dramatic improvisation, monitored by a human director. We have developed an affect-detection module to control an automated virtual actor and to contribute to the automation of directorial functions. The work also involves basic research into how affect is conveyed through metaphor. The project contributes to the application of sentiment and subjectivity analysis to the creation of emotionally believable synthetic agents for interactive narrative environments.

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 one of our industrial partners, Hi8us Midlands, 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 an automated bit-part actor, EMMA (emotion, metaphor and affect). 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. Indeed, Picard's work (2000) makes great contributions to building affective virtual characters. Also, emotion theories, particularly that of Ortony, et al. (1988) (OCC), have been used widely therein. Egges et al. (2003) have provided virtual characters with conversational emotional responsiveness. 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 judgements 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 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 Our Current Affect Detection

Various characterizations of emotion are used in emotion theories. The OCC model uses emotion labels (anger, etc.) and intensity, while Watson and Tellegen (1985) use positivity and negativity of affect as the major dimensions. Currently, we use an evaluation dimension (negative-positive), affect labels, and intensity. Affect labels plus intensity are used when strong text clues signalling affect are detected, while the evaluation dimension plus intensity is used for weak text clues. Moreover, our analysis reported here is based on the transcripts of previous e-drama sessions. Since even a person's interpretations of affect can be very unreliable, our approach combines various weak relevant affect indicators into a stronger and more reliable source of information for affect detection. Now we summarize our affect detection based on multiple streams of information.

2.1 Pre-processing Modules

The language in the speeches created in e-drama sessions severely challenges existing language-analysis tools if accurate semantic information is sought even for the purposes of restricted affect-detection. The language includes misspellings, ungrammaticality, abbreviations (often as in text messaging), slang, use of upper case and special punctuation (such as repeated exclamation marks) for affective emphasis, repetition of letters or words also for affective emphasis, and open-ended interjective and onomatopoeic elements such as "hm" and "grrrr". In the examples we have studied, which so far involve teenage children improvising around topics such as school bullying, the genre is similar to Internet chat.

To deal with the misspellings, abbreviations, letter repetitions, interjections and onomatopoeia, several types of pre-processing occur before actual detection of affect.

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. We are also considering dealing with abbreviations, etc. in a more general way by including them as special lexical items in the lexicon of the robust parser we are using (see below).

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', 'grrrrrr', 'agh', 'aaaggghhh') 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 (http://aspell.net/metaphone/), which is based on pronunciation, works with the dictionary to locate the base forms of words with letter repetitions.

Finally, the Levenshtein distance algorithm (http://www.merriampark.com/ld.htm) with a contemporary English dictionary deals with spelling mistakes in users' input.

2.2 Processing of Imperative Moods

One useful pointer to affect is the use of imperative mood, especially when used without softeners such as 'please' or 'would you'. Strong emotions and/or rude attitudes are often expressed in this case. There are special, common imperative phrases we deal with explicitly, such as "shut up" and "mind your own business". They usually

indicate strong negative emotions. But the phenomenon is more general.

Detecting imperatives accurately in general is by itself an example of the non-trivial problems we face. We have used the syntactic output from the Rasp parser (Briscoe & Carroll, 2002) and semantic information in the form of the semantic profiles for the 1,000 most frequently used English words (Heise, 1965) to deal with certain types of imperatives.

Rasp recognises some types of imperatives directly. Unfortunately, the grammar of the 2002 version of the Rasp parser that we have used does not deal properly with certain imperatives (John Carroll, p.c), which means that examples like "you shut up", "Dave bring me the menu", "Matt don't be so blunt" and "please leave me alone", are not recognized as imperatives, but as normal declarative sentences. Therefore, further analysis is needed to detect imperatives, by additional processing applied to the possibly-incorrect syntactic trees produced by Rasp.

If Rasp outputs a subject, 'you', followed by certain verbs (e.g. 'shut', 'calm', etc) or certain verb phrases (e.g. 'get lost', 'go away' etc), the sentence type will be changed to imperative. (Note: in "you get out" the "you" could be a vocative rather than the subject of "get", especially as punctuation such as commas is often omitted in our genre; however these cases are not worth distinguishing and we assume that the "you" is a subject.) If a softener 'please' is followed by the base forms of a verb, then the input is taken to be imperative. If a singular proper noun is followed by a base form of the verb, then this sentence is taken to be an imperative as well (e.g. "Dave get lost"). However, when a subject is followed by a verb for which there is no difference at all between the base form and the past tense form, then ambiguity arises between imperative and declarative (e.g. "Lisa hit me").

There is an important special case of this ambiguity. If the object of the verb is 'me', then in order to solve the ambiguity, we have adopted the evaluation value of the verb from Heise's (1965) compilation of semantic differential profiles. In these profiles, Heise listed values of evaluation, activation, potency, distance from neutrality, etc. for the 1,000 most frequently used English words. In the evaluation dimension, positive values imply goodness. Because normally people tend to use 'a negative verb + me' to complain about an unfair fact to the others, if the evaluation value is negative for such a verb, then this sentence is probably not imperative but

declarative (e.g. "Mayid hurt me"). Otherwise, other factors implying imperative are checked in this sentence, such as exclamation marks and capitalizations. If these factors occur, then the input is probably an imperative. Otherwise, the conversation logs are checked to see if there is any question sentence directed toward this speaker recently. If there is, then the input is conjectured to be declarative.

There is another type of sentence: 'don't you + base form of verb' that we have started to address. Though such a sentence is often interrogative, it is also often a negative version of an imperative with a 'you' subject (e.g. "Don't you dare call me a dog," "Don't you call me a dog"). Normally Rasp regards it as a question sentence. Thus, further analysis has also been implemented for such a sentence structure to change its sentence type to imperative. Although currently this has limited effect, as we only infer a (negative) affective quality when the verb is "dare", we plan to add semantic processing in an attempt to glean affect more generally from "Don't you …" imperatives.

2.3 Affect Detection by Pattern Matching

In an initial stage of our work, affect detection was based purely on textual pattern-matching rules that looked for simple grammatical patterns or templates partially involving lists of specific alternative words. This continues to be a core aspect of our system but we have now added robust parsing and some semantic analysis. Jess, a rule-based Java framework, is used to implement the pattern/template-matching rules in EMMA.

In the textual pattern-matching, particular keywords, phrases and fragmented sentences are found, but also certain partial sentence structures are extracted. This procedure possesses the robustness and flexibility to accept many ungrammatical fragmented sentences and to deal with the varied positions of sought-after phraseology in speeches. However, it lacks other types of generality and can be fooled when the phrases are suitably embedded as subcomponents of other grammatical structures. For example, if the input is "I doubt she's really angry", rules looking for anger in a simple way will fail to provide the expected results.

The transcripts analysed to inspire our initial knowledge base and pattern-matching rules were derived independently from previous *edrama* improvisations based on a school bullying scenario. We have also worked on another, distinctly different scenario, Crohn's disease, based on a TV programme by another of our industrial

partners (Maverick TV). The rule sets created for one scenario have a useful degree of applicability to other scenarios, though there will be a few changes in the related knowledge database according to the nature of specific scenarios.

The rules, as we mentioned at the beginning of this section, conjecture the character's emotions, evaluation dimension (negative or positive), politeness (rude or polite) and what response EMMA should make.

Multiple exclamation marks and capitalisation are frequently employed to express emphasis in e-drama sessions. If exclamation marks or capitalisation are detected in a character's utterance, then the emotion intensity is deemed to be comparatively high (and emotion is suggested even in the absence of other indicators).

A reasonably good indicator that an inner state is being described is the use of 'I' (see also Craggs & 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 the affective state 'threatening' is normally being expressed; and the utterance is usually the shortened version of an implied conditional, e.g., "I'll scream [if you stay here]." Note that when 'I' is followed by a present-tense verb, a variety of other emotional states tend to be expressed, e.g. "I want my mum" (fear) and "I hate you" (dislike), I like you (liking). Further analysis of first-person, present-tense cases is provided in the following section.

2.4 Going Beyond Pattern Matching

In order to go beyond the limitations of simple pattern matching, sentence type information obtained from the Rasp parser has also been adopted in the pattern-matching rules. The general sentence structure information not only helps EMMA to detect affective states in the user's input (see the above discussion of imperatives), and to decide if the detected affective states should be counted, but also helps EMMA to make appropriate responses. Rasp will inform the pattern-matching rule with sentence type information. If the current input is a conditional or question sentence with affective keywords or structures in, then the affective states won't be valued. For example, if the input is "I like the place when it is quiet", Rasp works out its sentence type: a conditional sentence and the rule for structures containing 'like' with a normal declarative sentence label won't be activated. Instead, the rule for the keyword 'when' with a

conditional sentence type label will be fired. Thus an appropriate response will be obtained.

Additionally, as we discussed in section 2.2, we use Rasp to indicate imperative sentences, such as when Mayid (the bully) said "Lisa, don't tell Miss about it". The pseudo-code example rule for such input is as follows:

(defrule example_rule
?fact <- (any string containing negation and the sentence type is 'imperative') =>
(obtain affect and response from knowledge database)

Thus the declarative input such as "I won't tell Miss about it" won't be able to activate the example rule due to different sentence type information. Especially, we have assigned a special sentence type label ('imp+please') for imperatives with softener 'please'. Only using this special sentence type label itself in the pattern-matching rule helps us effortlessly to obtain the user's linguistic style ('polite') and probably a polite response from EMMA as well according to different roles in specific scenarios.

Aside from using the Rasp parser, we have also worked on implementing simple types of semantic extraction of affect using affect dictionaries and electronic thesauri, such as WordNet. The way we are currently using WordNet is briefly as follows.

2.5 Using WordNet for a First Person Case

As we mentioned earlier, use of the first-person with a present-tense verb tends to express an affective state in the speaker, especially in discourse in which affect is salient, as is the case in scenarios such as School Bullying and Crohn's Disease. We have used the Rasp parser to detect such a sentence. First of all, such user's input is sent to the pattern-matching rules in order to obtain the speaker's current affective state and EMMA's response to the user. If there is no rule fired (i.e. we don't obtain any information of the speaker's affective state and EMMA's response from the pattern-matching rules), further processing is applied. We use WordNet to track down the rough synonyms of the verb (possibly from different WordNet "synsets") in the verb phrase of the input sentence, in order to allow a higher degree of generality than would be achieved just with the use of our pattern-matching rules. In order to find the closest synonyms to the verb in different synsets, the semantic profiles of the 1,000 most frequently used English words (Heise, 1965) have been employed, especially to find the evaluation values of every synonym of the original verb. We transform positive and negative evaluation values in Heise's dic-

tionary into binary 'positive' and 'negative' only. Thus if any synonym has the same evaluation value ('positive' or 'negative') as that of the original verb, then it will be selected as a member of the set of closest synonyms. Then, we use one closest synonym to replace the original verb in the user's input. This newly built sentence will be sent to the pattern-matching rules in order to obtain the user's affective state and EMMA's response. Such processing (using a closest synonym to replace the original verb and sending the newly built sentence to the pattern-matching rules) continues until we obtain the speaker's affective state and appropriate response.

2.6   Responding Regimes

EMMA normally responds to, on average, every Nth speech by another character in the e-drama session, where N is a changeable parameter (currently set to 3). However, it also responds when EMMA's character's name is mentioned, and makes no response if it cannot detect anything useful in the utterance it is responding to. The one-in-N average is achieved by sampling a random variable every time another character says something. As a future development, we plan to have N dynamically adjustable according to how confident EMMA is about what it has discerned in the utterance at hand.

EMMA sometimes makes a random response from several stored response candidates that are suitable for the affective quality it has discerned in an utterance it is responding to. In addition, EMMA sometimes reflects back, in modified form, part of the user's input string as a subcomponent of its response. Notice here that, because the pre-processing module reported in section 2.1 expands abbreviations and corrects misspellings, it helps to obscure the fact that part of EMMA's response is only a reflection. E.g.:

<Character> I'm ur (*your*) best m8 (*mate*).

<EMMA> Why are you my best mate?

On the other hand, we also, on a random basis, purposely put phenomena such as abbreviations and Internet-chat features in EMMA's responses in order to make EMMA express itself more like a typical school-student user. E.g.:

<Character> You get lost!

<EMMA> Hey, coool it folks. (notice the lengthening of 'cool')

<Character> He didn't do it.

<EMMA> No1 believes you, m8. (where 'no1' stands for 'no one')

It should be borne in mind that EMMA's responses are not aimed at engaging with the de-

tailed meaning of the utterance, but simply to stimulate the improvisation in a way that is somewhat sensitive to affect being expressed. Furthermore, in order to make the EMMA character's responses push the improvisation forward, the character will not only ask scenario related questions to the main characters, but also introduce new topics closely related to the scenario in the improvisation. In a recent user-testing debrief session, secondary school students mentioned that the human bit-part character did not stay in character and said pointless things, while in another session one student, who played a main character, believed that the EMMA character was the only one that stuck to scenario related topics. The directors reported that, even when a main character was silent and the director did not intervene very much, the EMMA character led the improvisation on the right track by raising new topics other characters were concerned about.

3   Affect via Metaphor

In the introduction we commented on two functions of metaphor. Metaphor is often used to convey affect and it also partly underlies folk theories of how affect and emotion work. As an example of the latter, folk theories of anger often talk about, and appear to conceive of, anger as if it were a heated fluid possibly exerting a strong pressure on its containing body. This motivates a wide range of metaphorical expressions both conventional such as "he was boiling with anger and about to blow his top" and more creative variants such as "the temperature in the office was getting higher and this had nothing to do with where the thermostat was set" (modified, slightly from a Google™ search). Passion, or lack of, is also often described in terms of heat and the latter example could in certain contexts be used in this manner. So far, examples of actors reflecting or commenting on the nature of their or others emotions, which would require an appropriate vocabulary, have been infrequent in the e-drama transcripts, although we might expect to find more examples as more students participate in the Crohn's disease scenario.

However, such metaphorically motivated folk models often directly motivate the terminology used to convey affect, as in utterances such as "you leave me cold", which conveys lack of interest or disdain. This use of metaphor to motivate folk models of emotions and, as a consequence, certain forms of direct expression of

emotion has been extensively studied, albeit usually from a theoretical, linguistic, perspective (Fussell & Moss, 1998; Kövecses, 1998).

Less recognised (although see Barnden et al., 2004; Wallington et al., 2006) is the fact that metaphor is also frequently used to convey emotion more indirectly. Here the metaphor does not describe some aspect of an emotional state, but something else. Crucially, however, it also conveys a negative or positive value judgement which is carried over to what is being described and this attitude hints at the emotion. For example to say of someone's room that "it is a cess-pit" allows the negative evaluation of 'cess-pit' to be transferred to 'the room' and we might assume an emotion of disgust. In our transcripts we find examples such as "smelly attitude" and "you buy your clothes at the rag market" (which we take to be not literally true). Animal insults such as "you pig" frequently take this form, although many are now highly conventionalised. Our analysis of e-drama transcripts shows that this type of metaphor that conveys affect indirectly is much more common than the direct use.

It should be apparent that even though conventional metaphorical phraseology may well be listed in specialised lexicons, approaches to metaphor and affect which rely upon a form of lexical look-up to determine the meaning of utterances are likely to miss both the creative variants and extensions of standard metaphors and also the quite general carrying over of affectual evaluations from the literal meaning of an utterance to the intended metaphorical meaning.

At the time of writing (early June 2006) little in the way of metaphor handling has been incorporated into the EMMA affect-detection module. However, certain aspects of metaphor handling will be incorporated shortly, since they involve extensions of existing capabilities. Our intended approach is partly to look for stock metaphorical phraseology and straightforward variants of it, which is the most common form of metaphor in most forms of discourse, including e-drama. However, we also plan to employ a simple version of the more open-ended, reasoning-based techniques described in the ATT-Meta project on metaphor processing (Barnden et al., 2004; Wallington et al., 2006).

As a first step, it should be noted that insults and swear words are often metaphorical. We are currently investigating specialised insult dictionaries and the machine-readable version of the OALD, which indicates slang.

Calling someone an animal of any sort usually conveys affect, but it can be either insulting or affectionate. We have noted that calling someone the young of an animal is often affectionate, and the same is true of diminutive (e.g., 'piglet') and nursery forms (e.g., 'moo cow'), even when the adult form of the animal is usually used as an insult. Thus calling someone 'a cat' or 'catty' is different from describing them as kittenish. Likewise, "you young pup" is different from "you dog". We are constructing a dictionary of specific animals used in slang and as insults, but, more generally, for animals not listed we can use WordNet and electronic dictionaries to determine whether or not it is the young or mature form of the animal that is being used.

We have already noted that in metaphor the affect associated with a source term will carry across to the target by default. EMMA already consults Heise's compilation of semantic differential profiles for the evaluation value of the verb. We will extend the determination of the evaluation value to all parts of speech.

Having the means to determine the emotion conveyed by a metaphor is most useful when metaphor can be reliably spotted. There are a number of means of doing this for some metaphors. For example, idioms are often metaphorical (Moon 1988). Thus we can use an existing idiom dictionary, adding to it as necessary. This will work with fixed idioms, but, as is often noted, idioms frequently show some degree of variation, either by using synonyms of standard lexis, e.g., '**constructing** castles in the air' instead of 'building castles in the air', or by adding modifiers, e.g., 'shut your **big fat** mouth'. This variability will pose a challenge if one is looking for fixed expressions from an idiom dictionary. However, if the idiom dictionary is treated as providing base forms, with for example the nouns being treated as the head nouns of a noun-phrase, then the Rasp parser can be used to determine the noun phrase and the modifiers of the head noun, and likewise with verbs, verb-phrases, etc. Indeed, this approach can be extended beyond highly fixed expressions to other cases of metaphor, since as Deignan (2005) has noted metaphors tend to display a much greater degree of fixedness compared to non-metaphors, whilst not being as fixed as what are conventionally called idioms.

There are other ways of detecting metaphors which we could utilise. Thus, metaphoricity signals (as in Goatly, 1997; Wallington et al., 2003) signal the use of a metaphor in some cases. Such

signals include phrases such as: *so to speak, sort of, almost, picture as*. Furthermore, semantic restriction violations (Wilks, 1978; Fass, 1997; Mason, 2004), as in "my car **drinks** petrol," often indicate metaphor, although not all metaphors violate semantic restrictions. To determine whether semantic restrictions are being violated, domain information from ontologies/thesauri such as WordNet could be used and/or statistical techniques as used by Mason (2004).

4   User Testing

We conducted a two-day pilot user test with 39 secondary school students in May 2005, in order to try out and a refine a testing methodology. The aim of the testing was primarily to measure the extent to which having EMMA as opposed to a person play a character affects users' level of enjoyment, sense of engagement, etc. We concealed the fact that EMMA was involved in some sessions in order to have a fair test of the difference that is made. We obtained surprisingly good results. Having a minor bit-part character called "Dave" played by EMMA as opposed to a person made no statistically significant difference to measures of user engagement and enjoyment, or indeed to user perceptions of the worth of the contributions made by the character "Dave". Users did comment in debriefing sessions on some utterances of Dave's, so it was not that there was a lack of effect simply because users did not notice Dave at all. Also, the frequencies of human "Dave" and EMMA "Dave" being responded to during the improvisation (sentences of Dave's causing a response divided by all sentences said by "Dave") are both roughly around 30%, again suggesting that users notice Dave. Additionally, the frequencies of other side-characters being responded to are roughly the same as the "Dave" character – "Matthew": around 30% and "Elise": around 35%.

Furthermore, it surprised us that no user appeared to realize that sometimes Dave was computer-controlled. We stress, however, that it is not an aim of our work to ensure that human actors do not realize this. More extensive, user testing at several Birmingham secondary schools is being conducted at the time of writing this paper, now that we have tried out and somewhat modified the methodology.

The experimental methodology used in the testing is as follows, in outline. Subjects are 14-16 year old students at local Birmingham schools. Forty students are chosen by each school for the testing. Four two-hour sessions take place at the school, each session involving a different set of ten students. In a session, the main phases are as follows: an introduction to the software; a First Improvisation Phase, where five students are involved in a School Bullying improvisation and the remaining five in a Crohn's Disease improvisation; a Second Improvisation Phase in which this assignment is reversed; filling out of a questionnaire by the students; and finally a group discussion acting as a debrief phase. For each improvisation, characters are pre-assigned to specific students. Each Improvisation Phase involves some preliminaries followed by ten minutes of improvisation proper.

In half of the SB improvisations and half of the CD improvisations, the minor character Dave is played by one of the students, and by EMMA in the remaining. When EMMA plays Dave, the student who would otherwise have played him is instructed to sit at another student's terminal and thereby to be an audience member. Students are told that we are interested in the experiences of audience members as well as of actors. Almost without exception students have appeared not to have suspected that having an audience member results from not having Dave played by another student. At the end of one exceptional session some students asked whether one of the directors from Hi8us was playing Dave.

Of the two improvisations a given student is involved in, exactly one involves EMMA playing Dave. This will be the first session or the second. This EMMA-involvement order and the order in which the student encounters SB and CD are independently counterbalanced across students.

The questionnaire is largely composed of questions that are explicitly about students' feelings about the experience (notably enjoyment, nervousness, and opinions about the worth of the dramatic contributions of the various characters), with essentially the same set of questions being asked separately about the SB and the CD improvisations. The other data collected are: for each debrief phase, written minutes and an audio and video record; notes taken by two observers present during each Improvisation Phase; and automatically stored transcripts of the sessions themselves, allowing analysis of linguistic forms used and types of interactivity. To date only the non-narrative questionnaire answers have been subjected to statistical analysis, with the sole independent variable being the involvement or otherwise of EMMA in improvisations.

5 Conclusion and Ongoing Work

We have implemented a limited degree of affect-detection in an automated bit-part character in an e-drama application, and fielded the actor successfully in pilot user-testing. 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 also intend to use the affect-detection in a module for automatically generating director messages to human actors.

In general, our work contributes to the issue of how affect/sentiment detection from language can contribute to the development of believable responsive AI characters, and thus to a user's feeling of involvement in game playing. Moreover, the development of affect detection and sentiment & subjectivity analysis provides a good test-bed for the accompanying deeper research into how affect is conveyed linguistically.

References

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

Briscoe, E. & J. Carroll. 2002. Robust Accurate Statistical 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*.

Deignan , A. 2005. *Metaphor and corpus Linguistics*. John Benjamins.

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.

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.

Fass, D. 1997. *Processing metaphor and metonymy*. Greenwich, Connecticut: Ablex

Goatly, A. 1997. *The language of metaphors*. Routledge London and New York:

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.

Mason, Z.J. 2004. CorMet: a computational, corpus-based conventional metaphor extraction system. *Computational Linguistics 30:1*. pp. 23-44.

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

Moon, R. 1998. *Fixed idioms and expressions in English*. Clarendon Press: Oxford, U.K

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

Picard, R.W. 2000. Affective Computing. The MIT Press. Cambridge MA.

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.

Wallington, A.M., Barnden, J.A., Barnden, M.A., Ferguson, F.J. & Glasbey, S.R. 2003. Metaphoricity Signals: A Corpus-Based Investigation. Technical Report CSRP-03-5, School of Computer Science, The University of Birmingham, U.K.

Wallington, A.M., Barnden, J.A. Glasbey S.R. and Lee M. G. 2006. Metaphorical reasoning with an economical set of mappings. *Delta*, 22:1

Wilks, Y. (1978). Making preferences more active. *Artificial Intelligence*, 10, pp. 75- 97