

Affect Detection and an Automated Improvisational AI Actor in E-drama

Li Zhang¹, Marco Gillies², John A. Barnden³, Robert J. Hendley³,
Mark G. Lee³ and Alan M. Wallington³

¹School of Computing and Technology, University of East London, Dockland
Campus, 4-6 University Way, London, E16 4LZ

²Department of Computer Science, University College London, London, WC1E
6BT

³School of Computer Science, University of Birmingham, Birmingham, B15 2TT

L.Zhang@cs.bham.ac.uk

Abstract. Enabling machines to understand emotions and feelings of the human users in their natural language textual input during interaction is a challenging issue in Human Computing. Our work presented here has tried to make our contribution toward such machine automation. We report work on adding affect-detection to an existing e-drama program, a text-based software system for dramatic improvisation in simple virtual scenarios, for use primarily in learning contexts. The system allows a human director to monitor improvisations and make interventions, for instance in reaction to excessive, insufficient or inappropriate emotions in the characters' speeches. Within an endeavour to partially automate directors' functions, and to allow for automated affective bit-part characters, we have developed an affect-detection module. It is aimed at detecting affective aspects (concerning emotions, moods, value judgments, etc.) of human-controlled characters' textual "speeches". The work also accompanies basic research into how affect is conveyed linguistically. A distinctive feature of the project is a focus on the metaphorical ways in which affect is conveyed. Moreover, we have also introduced how the detected affective states activate the animation engine to produce gestures for human-controlled characters. The description of our approach in this paper is taken in part from our previous publications [1, 2] with new contributions mainly on metaphorical language processing (practically and theoretically), 3D emotional animation generation and user testing evaluation. Finally, Our work on affect detection in open-ended improvisational text contributes to the development of automatic understanding of human language and emotion. The generation of emotional believable animations based on detected affective states and the production of appropriate responses for the automated affective bit-part character based on the detection of affect contribute greatly to the ease and innovative user interface in e-drama, which leads to high-level user engagement and enjoyment.

Keywords: E-drama, affect detection, improvisational AI actor, emotional behaviour and 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 one of our industrial partners, 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. One main aspect of our project is the addition of types of intelligent automation.

In the *edrama* system, up to five virtual characters are controlled on a virtual stage by human users (“actors”), with characters’ (textual) “speeches” typed by the actors operating the characters. A director is also involved in a session. A graphical interface on each actor’s and director’s terminal shows the stage and characters. Speeches are shown as text bubbles. Actors choose their characters clothes and bodily appearance.

Currently, cartoon figures against backdrops of real-life photographic images are used. However, we are bringing in animated gesturing avatars and 3D computer-generated settings using technology from an industrial partner. Actors and the human director work through software clients connecting with the server. Clients communicate using XML stream messages via the server, which is usually remote from the terminals, which may themselves be remote from each other. Terminal-server communication is over the Internet using standard browsers.

The actors are given a loose scenario around which to improvise, but are at liberty to be creative. One scenario we have used is school-bullying where a schoolgirl Lisa is being bullied by her classmate Mayid. There are also roles for two friends and a teacher. Within these parameters, actors must improvise interesting interchanges.

The human director has a number of roles. S/he must constantly monitor the unfolding drama and the actors’ interactions, or lack of them, in order to check whether they are keeping to the general spirit of the scenario. If this is not happening, the director may then intervene. For example, a director may intervene when the emotions expressed or discussed by characters are not as expected (or are not leading consistently in a new interesting direction). The director may also feel the need to intervene if one character is not getting involved, or is dominating the improvisation.

Intervention can take a number of forms. The director can send messages to actors. However, another important means of directorial intervention is for the director to introduce and control a ‘bit-part’ character. This character will not have a major role in the drama, but might, for example, try to interact with a character who is not participating much in the drama or who is being ignored by the other characters. Alternatively, it might make comments intended to ‘stir up’ the emotions of those involved, or, by intervening, diffuse any inappropriate exchange developing.

Clearly, all this places a heavy burden on the director. In particular, playing the role of the bit-part character and interacting with other characters whilst keeping interventions limited so as to maintain the main improvisatory drama amongst the actors, makes it difficult to fully monitor the behaviour of all the other actors and send appropriate messages to them should they stray off topic or exhibit inappropriate emotions. The difficulty is particularly acute if the directors are novices, such as teachers trying to use e-drama in their lessons.

One major research aim is accordingly to automate some directorial functions, either to take some of the burden away from a human director, or to provide a fully automated (though necessarily very restricted) director. With a fully-automated director, even if highly restricted in what it could do, little or no human supervision might be required for at least minimally adequate improvisations, and *edrama* could, for example, be added to websites about certain topics allowing visitors to engage in on-line role-play germane to the topic. However, our main current work is on merely assisting a human director. The assistance is by

- (a) fully-automated control of an optionally-included bit-part character
- (b) sending of automated suggestions to the human director about the progress of the improvisation or about messages to send to the human actors.

Point (b) is addressed briefly below. Our main focus in this paper is on (a).

For purpose (a), we have created a simple automated actor, EMMA, which controls a bit-part character who is an acquaintance of the other characters. EMMA contains an affect-detection module, which tries to identify affect in characters' speeches, allowing the EMMA character to make responses that, it is hoped, will stimulate the improvisation, thus leading to less need for intervention by the human director. Within affect we include: basic and complex *emotions* such as anger and embarrassment respectively; *meta-emotions* (emotions about emotions) such as desiring to overcome anxiety; *moods* such as hostility; and *value judgments* (judgments of goodness, importance, etc.). Although merely detecting affect is limited compared to extracting the full meaning of characters' utterances, we have found that in many cases this is sufficient for the purposes of stimulating the improvisation.

Even limited types of affect detection can be useful. We do not purport to be able to make EMMA detect all types of affect under all ways affect can be expressed or implied, or to do it with a high degree of reliability. The spirit of the project is to see how far we can get with practical processing techniques, while at the same time investigating theoretically the nature of, and potential computational ways of dealing with, forms of affective expression that may be too difficult currently to handle in a usable implemented system.

Much research has been done on creating affective virtual characters in interactive systems. Picard's work [3] makes great contributions to building affective virtual characters. Also, emotion theories, particularly that of Ortony et al. [4] (OCC), have been used widely therein. Ortony et al. include a rule-based system for the generation of the 22 emotion types, which is widely used for emotion generation for the development of intelligent virtual agents. In Prendinger and Ishizuka's work [5], animated agents are capable to perform affective communication. They have defined eliciting condition rules for more frequently used OCC emotion types. Then the generated emotional states are filtered by two control states – personality and social role awareness – in order to achieve believable emotional expression. Wiltchko's *eDrama Front Desk* [6] is an online emotional natural language dialogue simulator with a virtual reception interface for pedagogical purposes. Natural language parser is used to analyze users' input texts which index into a list of phrases that are frequently used. Then the system's dialogue manager selects an output phrase for the computer character. The emotion model of this system is derived from OCC model based on

two factors: respect and power. Mehdi et al. [7] combined the widely accepted five-factor model of personality [8], mood and OCC in generating emotional behaviour for a fireman training application. Personality and mood have also played important roles in emotion generation and emotion intensity adaptation. Implementation of animated agents with emotional behaviour is their next step. Gratch and Marsella [9] presented an integrated model of appraisal and coping, in order to reason about emotions and to provide emotional responses, facial expressions and potential social intelligence for virtual agents. Their main contribution is the introduction of computational description of emotional coping for the first time. They have also extended the scope of discourse on appraisal theories by incorporating influence between cognition and appraisal to obtain emotional coping. Additionally, Egges et al. [10] have also provided virtual characters with conversational emotional responsiveness. Elliott et al. [11] demonstrated tutoring systems that reason about users' emotions. Aylett et al. [12] also focused on the development of affective behaviour planning for the synthetic characters. There is much other work in a similar vein.

Additionally, human computing focuses on the creation of innovative interfaces to facilitate users to engage in more natural and inspiring interaction with machines by providing automatic understanding of human affect, language and behaviour [32]. Emotion recognition in speech and facial expression has been studied extensively [32, 33, 34]. But very few research work has made an attempt to dig out the affect flavour in human open-ended linguistic textual input in online role-play, although the first interaction system based on natural language textual input, Weizenbaum's Eliza [35], was first developed back in 1966. Thus there has been only a limited amount of work directly comparable to our own, especially given our concentration on improvisation and open-ended language. However, *Façade* [13] included shallow natural language processing for characters' open-ended utterances. There are two language processing components in *Façade*. The first component did many-to-few mapping and transformed various user text input into a few discourse acts. Then the second language processing component produced several reactions according to the generated discourse acts. Finally, the highest priority reaction in the highest priority context is chosen as the response for the virtual agents. But in *Façade*, the detection of major emotions, rudeness and value judgements is not mentioned. Zhe and Boucouvalas [14] demonstrated an emotion extraction module embedded in an Internet chatting environment (see also [15]). It uses 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 focuses only on emotional adjectives, and does not address deep issues such as figurative expression of emotion. Also, the concentration purely on first-person emotions is narrow. We might also mention work on general linguistic clues that could be used in practice for affect detection (e.g. [16]).

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 [17, 18].

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. In studying how ordinary people understand and talk about affect in ordinary life, what is of prime importance is their *common-sense* views of how affect works, irrespective of scientific accuracy of those views. Metaphor is strongly involved in such views.

It should also be appreciated that this paper does not address the emotional, etc. states of the *actors* (or director, or any audience). Our focus is on the affect that the actors make their characters express or mention. While an actor may work him/herself up into, or be put into, a state similar to or affected by those in his/her own characters' speeches or those of other characters, such interesting effects, which go to the heart of the dramatic experience, are beyond the scope of this paper, and so is the possibility of using information one might be able to get about actors' own affective states as a hint about the affective states of their characters or vice-versa.

Various characterizations of emotion are used in emotion theories. The OCC model uses emotion labels (anger, etc.) and intensity, while Watson and Tellegen [19] use positivity and negativity of affect as the major dimensions. We have drawn ideas from several such sources. We use an evaluation dimension (negative-positive), affect labels, and intensity. The basic emotion labels (such as 'angry') we use are taken from Ekman [20], while other comparatively complex affect labels (such as 'approving') are taken from the OCC model. There are 25 affect labels used in our system currently. Affect labels plus intensity are used when strong text clues signalling affect are detected, while the evaluation dimension plus intensity is used when only weak text clues are detected. The description of our approach in this paper is taken in part from our previous publications [1, 2] with new contributions mainly on metaphorical figurative language processing (practically and theoretically), 3D animation generation and user testing evaluation.

Finally, machines have struggled to understand human language and emotion expressed in it. Expecting machines to engage in a meaningful role-play with human characters based on the affect understanding in drama improvisation is even unthinkable. Our work presented here made some efforts towards these goals. Although we have implemented a limited degree of affect-detection in an automated bit-part character in an e-drama application, the statistical analysis results in our user study, especially on user engagement and enjoyment, indicate that the improvisational AI actor, EMMA, performed as good as another secondary school student.

2 Our Current Affect Detection

Before any automatic recognition and response components could be built for use in our automated actor EMMA, a detailed analysis of the language used in edrama sessions was necessary. A small corpus of sessions was analysed by hand to identify the range of linguistic forms used and to provide insight for the automatic processing. In fact, this analysis is often very difficult and unreliable but it does reveal some important observations.

- The language used is often complex and idiosyncratic. It is almost invariably ungrammatical, it uses abbreviations, it contains mis-spellings and it borrows heavily from the language of text-messaging (*textese*) and chat-rooms. Compared to the language normally analysed in computational linguistics it provides significant additional challenges.
- The literal meaning of the statements is often less important to its interpretation than the affect that it is expressing. The content of a statement is still important to building an understanding and to responding appropriately, but understanding the affective state being expressed is critical.
- The language contains a large number of weak cues to the affect that is being expressed. These cues may be contradictory or they may work together to enable a stronger interpretation of the affective state. In order to build a reliable and robust analyser of affect it is necessary to undertake several diverse forms of analysis and to enable these to work together to build stronger interpretations.

This leads to a system where the emphasis is moved away from building a representation of the meaning of the statement to one where more weight is given to building a robust representation of affective connotations.

The results of this affective analysis are then used to:

- Control an automated actor (EMMA) that operates a character in the improvisation: see (a) in Section 1.
- Independently of this, help create the directorial suggestions mentioned in Section 1, in point (b).
- Additionally, drive the animations of the avatars in the user interface so that they react bodily in ways that is consistent with the affect that they are expressing, for instance by changing posture or facial expressions.

The overall architecture is shown in Fig. 1.

Within the affect detection component we need to undertake several analyses of any given utterance by another character. These will each build representations which may be used by other components (e.g. syntactic structure) and will construct (possibly weak) hypotheses about the affective state. The architecture adopted is a blackboard based one. Each knowledge source undertakes its processing and writes its results to a central data structure (the blackboard) where they can be used by other knowledge sources and where hypotheses can be supported by multiple knowledge sources.

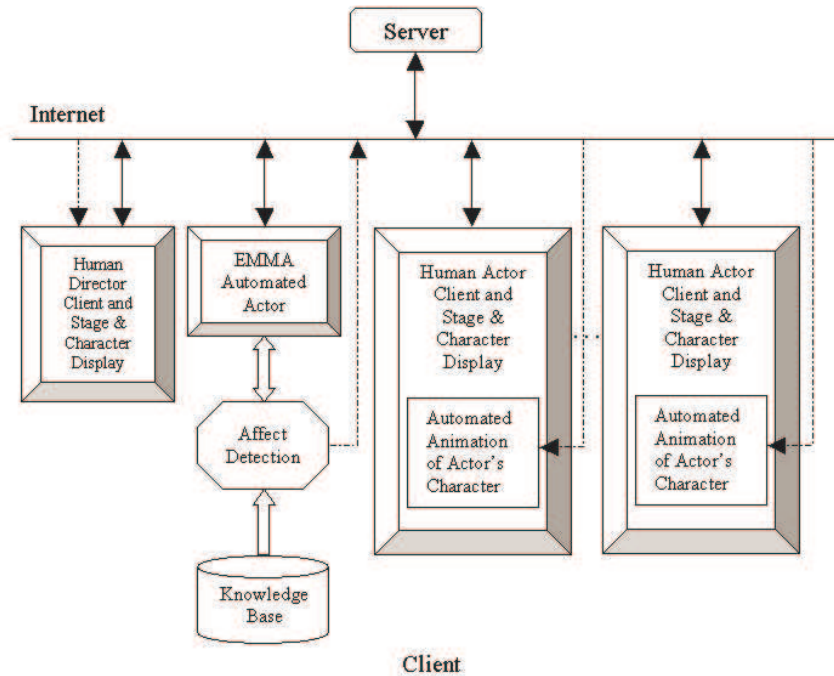


Fig. 1. Application architecture: Fat arrows show flow of information inside a program. Solid thin arrows show flow of character utterances and director messages formatted in XML. Dashed lines show flow of information about detected affect and flow of automated suggestions to the director.

A rule-based component takes these hypotheses and builds a single interpretation of the affective state being expressed by the utterance being analysed. This interpretation is then transmitted to the other components of the system: EMMA (the automated actor), the automated directorial-suggestion generator and the animation component in the software client that handles the actor whose utterance is being processed.

The response generation component of EMMA uses this interpretation to build its behaviour driven mainly by its role in the improvisation and the affect expressed in the statement to which it is responding.

Fig. 2. illustrates the overall structure of the language processing.

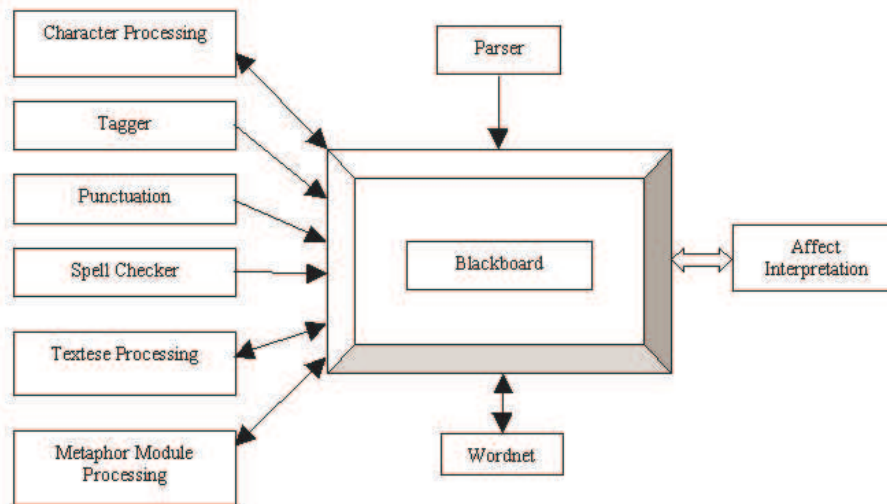


Fig. 2. Blackboard architecture

2.1 Pre-processing Modules

The language in the speeches created in *edrama* sessions severely challenges existing language-analysis tools if accurate semantic information is sought, even in the limited domain of restricted affect-detection. Aside from the complications noted above, the language includes slang, use of upper case and special punctuation (such as repeated exclamation marks) for affective emphasis, repetition of letters, syllables or words for emphasis, and open-ended interjective and onomatopoeic elements such as “hm”, “ow” 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 [21]. To deal with the misspellings, abbreviations, letter repetitions, interjections and onomatopoeia, several types of pre-processing occur before the main aspects of detection of affect. We have reported our work on pre-processing modules to deal with these language phenomena in detail in [1].

2.2 Affect Detection using Rasp, Pattern Matching & WordNet and Responding Regimes

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 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. Expression of the imperative mood in English is surprisingly various and ambiguity-prone, as illustrated below. We have used the syntactic output from the *Rasp* parser [22] and semantic information in the form of the semantic profiles for the 1,000 most frequently used English words [23] to deal with certain types of imperatives. Briefly, the grammar of the 2002 version of the *Rasp* parser that we have used incorrectly recognised certain imperatives (such as “you shut up”, “Dave bring me the menu” etc) as declaratives. We have made further analysis of the syntactic trees produced by *Rasp* by considering of the nature of the sentence subject, the form of the verb used, etc, in order to detect imperatives. We have also made an effort to deal with one special case of ambiguities: a subject + a verb (for which there is no difference at all between the base form and the past tense form) + “me” (e.g. ‘Lisa hit/hurt me’.). The semantic information of the verb obtained by using Heise’s [23] semantic profiles, the conversation logs and other indicators implying imperatives help to find out if the input is an imperative or not.

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 specific words or sets 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, including but going beyond the handling of imperatives discussed above.

A rule-based Java framework called Jess is used to implement the pattern/template-matching rules in EMMA allowing the system to cope with more general wording. 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 characters’ utterances. The rules conjecture the character’s emotions, evaluation dimension (negative or positive), politeness (rude or polite) and what response EMMA should make. 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 scenarios that we have used so far are school bullying and Crohn’s disease.

However, it lacks other types of generality and can be fooled when the phrases are suitably embedded as subcomponents of other grammatical structures. In order to go beyond certain such limitations, sentence type information obtained from the *Rasp* parser has also been adopted in the pattern-matching rules. This 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 (e.g. affects detected from conditional sentences won’t be valued), but also helps EMMA to make appropriate responses.

Additionally, the sentence type information can also help to avoid the activation of multiple rules, which could lead to multiple detected affect results for one user’s input. Mostly, it will help to activate only the most suitable rule to obtain the speaker’s affective state and EMMA’s response to the human character.

For example, as we discussed in the above, we use *Rasp* to indicate imperative sentences, such as in the school bullying scenario, 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
  (any string containing negation and the sentence type is 'imperative')
  =>
  (obtain affect and response from knowledge base))
```

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.

Additionally, a reasonably good indicator that an inner state is being described is the use of ‘I’ (see also [16]), especially in combination with the present or future tense (e.g. ‘I’ll scream’, ‘I hate/like you’, and ‘I need your help’). We especially process ‘the first-person with a present-tense verb’ statements using WordNet. When we fail to obtain the speaker’s affective state in the current input by using Rasp and pattern matching, WordNet is used to find the synonyms of the original verb in the user’s input. These synonyms are then refined by using Heise’s [23] semantic profiles in order to obtain a subset of close synonyms. The newly composed sentences with the verbs in the subset respectively replacing the original verb, have extended the matching possibilities in the pattern-matching rules to obtain user’s affective state in the current input.

After the automatic detection of users’ affective states, EMMA needs to make responses in her role to the human characters during the improvisation. We have also created responding regimes for the EMMA character. Most importantly, EMMA can adjust its response likelihood according to how confident EMMA is about what it has discerned in the utterance at hand. Especially, in order to make contributions to the improvisation progression, EMMA also has a global view of the drama improvisation. Briefly, the knowledge base of EMMA (see Fig. 1.) provides scenario’s background knowledge for each human character. EMMA can raise various scenario-related topics in its role for the human characters according to the detected affective states and topics discussed in the text input by using the rule-based reasoning based on the knowledge base. Inspection of the transcripts collected in the user testing indicates that EMMA usefully pushed the improvisation forward on various occasions (see section 4). For example, in the school bullying scenario, if the victim character hasn’t shown the expected emotional states (such as fear or anger) in ones role-play (this also means that this character lacks of expected negative emotional states), then the EMMA character will deliberately raise scenario-related topics towards this character in order to push the improvisation forward, e.g. by asking why he/she feels sad.

Details of the work reported in this section can be found in [1]. The brief summaries here of our previous implementations and their capabilities aim to remind readers.

2.3 Metaphorical Language Processing in EMMA

The metaphorical description of emotional states is common and has been extensively studied [17, 18]. E.g.: “He nearly exploded” and “Joy ran through me,” where anger and joy are being viewed in vivid physical terms. Such examples describe emotional states in a relatively explicit if metaphorical way. But affect is also often conveyed

more implicitly via metaphor, as in “His room is a cess-pit”: affect (such as ‘disgust’) associated with a source item (cess-pit) gets carried over to the corresponding target item (the room). In other work, we have conducted research on metaphor in general (see, e.g. [24, 25]), and are now applying it to the e-drama application, and conversely using the application as a useful source of theoretical inspiration.

Our intended approach to metaphor handling in the EMMA 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 on metaphor processing [24]. As an example of stock phrase handling, insults in e-drama are often metaphorical, especially the case of animal insults (“you stupid cow”, “you dog”). Particularly the ‘second-person/a singular proper noun + present-tense copular form’ statements (such as ‘you’re a rat’, ‘Lisa is a pig’) and the second-person phrases (such as ‘you dog’) are often used to express insults. In the EMMA affect-detection module, we use *Rasp* to locate such user’s input. We have also employed an on-line animal-name dictionary (<http://dictionary.reference.com/writing/styleguide/animal.html>), including names of animals, animal groups, young animals, etc, since usually calling someone a baby animal name (e.g. “puppy”) may imply affection while calling someone an adult animal name could convey an insult. Then we use this animal-name dictionary to find out if there is any potential insulting/affectionate animal name present in the ‘second-person/a singular proper noun + present-tense copular form’ statements or in the second-person phrases. If there is, then we will use WordNet to analyze the animal name. If WordNet provides the semantic information of the animal name containing the description of the characteristics of a person/woman/men, such as ‘an adjective + person/woman/man’, then we will classify such user’s input as metaphorical language. Additionally, we also use another semantic profile developed by Esuli and Sebastiani [26] to obtain the evaluation value of the adjective preceding the word ‘person/woman/man’. If it’s negative (e.g. ‘a disagreeable person’, ‘a disgraceful person’), then we classify the user’s input as metaphorical insulting language. If it’s positive (e.g. ‘a lovely person’, ‘a famous man’), then the user’s input will be considered as metaphorical affectionate language. Otherwise, the user’s input will be regarded as metaphorical objective language.

Not only do animal names present in the above statements may convey affective states, but also use of special person-types (e.g. ‘freak’) or mythical being (e.g. ‘devil’, ‘angel’) in such statements could also imply insults or approbations. Thus if there is a noun, but not an animal name, present in the ‘second-person/a singular proper noun + present-tense copular form’ statements or in the second-person phrases, then we collect all of its synonyms using WordNet. If any of the special person-types or mythical being collected in previous e-drama transcripts is shown among the synonyms, then we conclude that it contains insulting/affectionate flavour.

Sometimes, adjectives instead of nouns are used to directly convey affective states, such as ‘Lisa is a stupid girl’, ‘you’re a good mum’ and ‘you stupid boy’. If there is no noun present in the above statements or we couldn’t obtain any semantic information by only analysing nouns in the above sentence structures, adjectives will be very helpful. We could find out the evaluation values of these adjectives, again using the semantic profile developed by Esuli and Sebastiani [26]. In this way, we may obtain at least the positive or negative flavour in the user’s input.

One particular phenomenon of theoretical and practical interest is that 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.” The bigness is sometimes literal as well. “Big bully” expresses strong disapproval [27] and “little bully” can express contempt, although “little” can also convey sympathy or be used as an endearment. Such examples are not only important in practice but also theoretically challenging. We have also encountered surprisingly creative uses of metaphor in e-drama. For example, in the school-bullying scenario, Mayid is portrayed as having already insulted Lisa by calling her a ‘pizza’ (short for ‘pizza-face’). This figurative insult was given a theoretically intriguing, creatively metaphorical elaboration in one improvisation, where Mayid said “I’ll knock your topping off, Lisa.”

Our work on metaphor outside the e-drama research is focused on an approach and system called ATT-Meta [24]. This approach is heavily dependent on detailed utterance-meaning analysis and on rich knowledge bases and reasoning processes, and is currently unsuitable for direct use in the *edrama* system. However, examples arising in *edrama* transcripts such as the more creative ones above provide useful data guiding the further development of ATT-Meta and can pose useful challenges to current metaphor theory generally. Moreover, theoretical study of the affective metaphorical language in e-drama transcripts is presented in section 3.

Finally, the detected affective states in the user’s text input and EMMA’s responses to other human characters have been encoded in an xml stream, which is sent to the server by EMMA. Then the server broadcasts the xml stream to all the clients so that the detected affective states information can be picked up by the animation engine to contribute to the production of 3D gestures and postures for the avatars. Now we will discuss the generation of emotional believable animation in detail in the following section.

2.4 The Users Avatars and Emotional Animation

The topics discussed in the edrama scenarios are often highly emotionally charged and this is reflected in the animation of the characters. Each participant in edrama has their own animated graphical character (avatar). In order for the characters to enhance the interaction the characters all have emotionally expressive animations. Garau et al. [28] point out that avatars that do not exhibit appropriate emotional expression during emotionally charged conversation can be detrimental to an interaction. The problem with animated avatars is that they can be very complex to use if users have to directly control the avatars animation. Vilhjálmsson and Cassell [29]) have shown that users find controlling animated avatars difficult and their experience and interaction is improved if they use an avatar whose behaviour is controlled autonomously. We therefore have an autonomous model of affective animation for our avatars based on the affective states detected in users’ text input. These detected affective states control the animation of the user avatars using Demeanour expressive animation framework [30].



Fig. 3. Four actors in the green room

Demeanour makes it possible for our characters to express the affective states detected by EMMA. When EMMA 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. The animation system is based around a set of short animation clips, each of which is labelled with one or more affective states. Each clip only affects an individual part of the body (torso, legs, arms) and thus several clips can be easily combined at the same time. When a new affective state is received a new set of clips is chosen at random from the clips labelled with that state and these new clips are combined together to produce a new animation. Every few seconds the set of clips used is varied at random to produce a new animation, but one which has the same affective state as before. This allows us to produce varied behaviour over long time periods. The animation system also implements affective decay. Any affective state will eventually revert to a neutral state if it is not replaced by a new one.

Another feature of the animation system is that characters can produce affective responses to the states of other characters. If a character produces a strong affective state then other characters will also produce a milder response. Each character has a profile which specifies how it responds to the behaviour of each other character. This makes it possible to implement different responses for different characters. For example, two characters with a positive relationship may empathise with each other, when one is unhappy so is the other. On the other hand if two characters have a negative relationship then one might gloat at the other's unhappiness, and therefore display happiness.



Fig. 4. Emotional believable agents on stage

Demeanour generates a number of output affective states, which are used to select the animation clips. Each output state is a weighted sum of a number of input factors. The primary input factor is the affective state as detected from the input text, this always has weight 1. The inputs also include the states of other characters, with lower weights. So for example, the output state “happiness” depends on the input “happiness”, but also on the “happiness” and “sadness” values of other characters. The weights of the other characters states are contained in the character’s profile. The profile consists of a separate set of weights for each other character in the scenario. This makes it possible to respond differently to each character. For example, if two characters, A and B have a poor relationship and A is angry with B, B might respond by being angry back. On the other hand if B’s parents were angry then B might be sad or submissive. At any time each character has a single focus of social attention (which is itself another character), determined by the character’s direction of gaze (which is itself determined by an animated gaze model). In order to generate animation the first step is to update the input states based on any text typed in. Next the states of the current focus of attention are fetched. These are multiplied by the weights of given by the profile specific to the focus of attention and added to the input emotion to produce the output state.

3 Affect via Metaphor – Theoretical Analysis in E-drama Transcripts

The preliminary analysis of transcripts from the user-testing that we have performed (see below) has shown or suggested a number of things concerning metaphor, as follows.

It has provided additional evidence for the importance of one of the principles we have put forward in our theoretical work on metaphor concerning the meaning of metaphorical utterances. This concerns various types of information that we propose are transferred by default in metaphor generally, irrespective of the particular metaphorical view or views taken by the utterance (e.g., whether it views a mind as a physical container or a company as an animal). We have proposed in other work [24, 25] that – amongst various other types of information – emotional attitudes, value judgments and propositions concerning reduced functionality are transferred from source to target. We call the particular transfer principles involved view-neutral mapping adjuncts (VNMAAs). Much if not most of the discourse contribution of a metaphorical utterance can be conveyed by the action of VNMAAs, and this is borne out by many of the examples of metaphor in the e-drama transcripts.

Indeed, the conveying of emotional attitudes and the description of emotional states seems to be done more commonly in our genre via VNMAAs than by specific metaphorical views of emotion such as those discussed by Kövecses [18] (an example of such a view is ANGER AS HEATED LIQUID).

These points are highly significant in terms of metaphor theory, because

(a) the central role of transfer of emotion and value judgments in metaphor has often been recognized but has not previously been properly systematized in theories,

(b) VNMAAs factor out metaphorical-view-neutral transfers that other researchers have cast as being aspects of specific metaphorical views – we are therefore able to propose much simpler accounts of specific metaphorical views.

The types of metaphor that have arisen in e-drama transcripts sometimes resist pigeonholing neatly into the array of metaphorical views that have been suggested in the literature, so that new views are suggested. For example, we suggest a general view whereby INTERPERSONAL EFFECTS are viewed as PHYSICAL EFFECTS. Particular special cases of this have been suggested in the literature, and some cases of the new view fit the much more general CAUSES AS FORCES view [18], but our new intermediate view leads to a more focussed and appropriate analysis.

Many authors have struggled with the task of relating metaphor to, and distinguishing it from, other types of figurative language, such as metonymy and hyperbole. A number of e-drama examples lend weight to our suspicion that the distinction between metaphor and other figures is yet more troublesome than has been assumed so far. For example, many insults rely on casting someone as insane: this can potentially be analysed either as just hyperbole or as hyperbolic metaphor. Our work is leading to new light being thrown on such analytical difficulties, which have consequences for automated processing approaches as well as for theory.

Finally, the e-drama examples throw into relief the importance of further clarifying how size attributions, for instance via the adjectives “big” and “little”, contribute to metaphorical descriptions, as mentioned above. We are studying the extent to which the effects could be produced by VNMAAs, or whether on the other hand they rest on specific metaphorical views.

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 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 few users 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 has been conducted recently (up to September 2006). We have conducted an initial evaluation of the quality of EMMA's determinations about emotion during these school testing sessions, by comparing EMMA's determinations during one of the School Bullying improvisations with emotion labels later assigned offline by two members of our team (not centrally involved in the development of EMMA's algorithms). The humans were constrained to use the set of emotion labels that EMMA uses. It is still unclear how to conduct such an evaluation, because the conscious thoughts of a human annotator (labeller) about emotions revealed by an utterance in an offline labelling task may be different from the emotions the annotator would have unconsciously understood during an online e-drama session, and EMMA only surmises an emotion when there is strong evidence whereas a human labeller may proceed on a different level of evidence. Also, it transpires that there was mediocre agreement between the two human labellers, and the task is artificial for them because they might normally have assigned an emotion outside the prescribed set of 25. To compare the labelling of the two human labellers to each other and to the labelling by EMMA, we used the often-used *kappa* statistic of Carletta [31]. It is a measure of the pairwise agreement among a set of coders making category judgements, correcting for expected chance agreement. The statistic, *K*, is calculated as $K = (P(A) - P(E)) / (1 - P(E))$ where *P*(*A*) is the proportion of times two coders agree and *P*(*E*) is the proportion of times we would expect them to agree if they categorized randomly. A value of at least 0.6 – 0.8 is generally required by researchers looking for good inter-annotator agreement. We calculated *K* for each pair among the three labellers (EMMA and two humans). The inter-human *K* was only 0.32, and so it is not surprising that the EMMA/human values were only 0.32 again and 0.23. However, we also performed a modified

comparison in which the emotion labels were conflated to three (positive, negative, neutral) by grossly lumping together, for example, the labels deemed positive by our team. We then got a human/human K of 0.65, and EMMA/human values of 0.55 and 0.42. The latter are not good values, but they at least give grounds for hope that with further refinement of our affect detection we can come near the rather low human/human level of agreement.

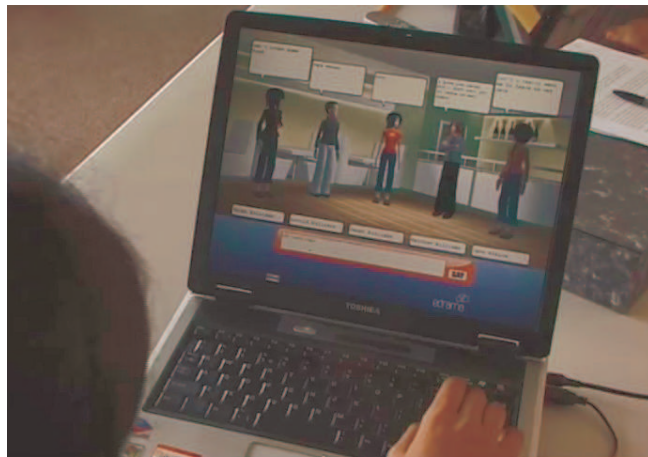


Fig. 5. E-drama user testing with school children

Inspection of the transcripts collected indicates that EMMA usefully pushed the improvisation forward on various occasions. Fig. 6. shows an example about how EMMA contributed to the drama improvisation in Crohn's disease scenario. Briefly, in Crohn's disease scenario, Peter has had a terrible bowel disease – Crohn's disease – since he was 15. Crohn's disease attacks the wall of the intestines and makes it very difficult to digest food properly. Peter has the option to undergo surgery (ileostomy) which will have a major impact on his life, but he will carry a bag outside his stomach for the rest of his life. So he needs his family (Dad: Arnold, Mum: Janet, brother: Matthew) and his close friend (Dave) to help him make a decision about having another operation or not. He decides to take them out to a restaurant to talk about it. His dad (Arnold) loves him but always feels uncomfortable and embarrassed when Peter talks about his disease in public. Peter's mum cares about him very much, sometimes worries too much about Peter. Matthew is totally against the operation idea and believes Peter will get bullied because of the side-effect of the operation. Dave, Peter's best mate, loves Peter and his family and tries to mediate the discussion. In the example transcript shown in Fig. 6, Dave was played by EMMA, which successfully led the improvisation on the desirable track. In another scenario (school bully scenario) that we have used for the user testing, example transcripts are also shown that the EMMA character has helped to push the improvisation forward.

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.

Example transcripts - Crohn's disease scenario

Dave - the AI character

Peter Williams: I'M THE ILL ONE

Dave Atkins: Arnold, Peter needs ur support and u can't just ignore it.

Janet Williams: nobody cares peter

Arnold Williams: we know peter now stop talking about it

Peter Williams: yes i need your support

Dave Atkins: Hey guys, lets all calm down a bit.

Arnold Williams: we will discuss it later

Janet Williams: well you have got mine

Peter Williams: help me daddy

Arnold Williams: not now son

Fig. 6. Part of an example transcript produced in the user testing in Swanshurst School, Birmingham. 28th June 2006

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 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 character successfully in pilot user-testing. Although there is a considerable distance to go in terms of the practical affect-detection that we plan eventually to implement, the already implemented detection is able to cause contributions by the automated character that are reasonably appropriate to the discourse.

Additionally, E-drama provides a platform for participants to engage in focused discussion around emotionally charged issues. This new prototype provides an opportunity for the developers to explore how emotional issues embedded in the scenarios, characters and dialogue can be represented visually without detracting from the learning situation.

Our project makes a contribution to the issue of what types of automation should be included in interactive narrative environments, and as part of that the issue of what types of affect should be detected (by directors, etc.) and how. Moreover, our project also makes a contribution to the development of automatic understanding of human language and emotion in human computing. The generation of emotional believable animations based on the detected affective states contributes to the ease and innovative user interface in e-drama, which leads to high-level user engagement and enjoyment.

Our remaining work on affect detection within the project will be on the metaphorical aspects of it, both in the sense of developing our theoretical ideas about metaphor further in the light of the e-drama data and in the sense of implementing certain limited forms of metaphor processing that can reveal useful extra hints about the affect that is present.

Future work could also include the exploration of automated bit-part characters to fully develop a non-human director. Additionally tools to enable participants to replay the role-plays have been considered. These could enable further reflection and group discussion, allowing for comparisons of sessions between different groups of learners. Replays could even be altered to adjust the emotional states of each character and

generate different online ‘performances’, which could create emotionally rich experiences for audiences as well as participants.

Acknowledgments. 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 Z. Wen, R. Aggeri, W.H. Edmondson and S.R. Glasbey. The work is also partially supported by EPSRC grant EP/C538943/1.

References

1. Zhang, L., Barnden, J.A., Hendley, R.J. & Wallington, A.M. 2006a. Exploitation in Affect Detection in Open-ended Improvisational Text. In *Proceedings of Workshop on Sentiment and Subjectivity at COLING-ACL 2006*, Sydney, July 2006.
2. Zhang, L., Barnden, J.A., Hendley, R.J. & Wallington, A.M. 2006b. Developments in Affect Detection in E-drama. In *Proceedings of EACL 2006, 11th Conference of the European Chapter of the Association for Computational Linguistics*, 2006, Trento, Italy. pp. 203-206.
3. Picard, R.W. 2000. *Affective Computing*. The MIT Press. Cambridge MA.
4. Ortony, A., Clore, G.L. & Collins, A. 1988. *The Cognitive Structure of Emotions*. Cambridge U. Press.
5. Prendinger, H. and Ishizuka, M. Simulating Affective Communication with Animated Agents. In *Proceedings of Eighth IFIP TC.13 Conference on Human-Computer Interaction*, Tokyo, Japan, pp.182-189.
6. Wiltschko, W. R. Emotion Dialogue Simulator. eDrama learning, Inc. eDrama Front Desk.
7. Mehdi, E. J., Nico P., Julie D. and Bernard P. 2004. Modeling Character Emotion in an Interactive Virtual Environment. *Proceedings of AISB 2004 Symposium: Motion, Emotion and Cognition*. Leeds, UK.
8. McCrae, R.R. and John, O.P. 1992. An Introduction to the Five Factor Model and Its Application. *Journal of Personality*, 60, 175-215.
9. Gratch, J. and Marsella, S. A Domain-Independent Framework for Modeling Emotion. *Journal of Cognitive Systems Research*. Vol 5, Issue 4, pp.269-306.
10. 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.
11. Elliott, C., Rickel, J. & Lester, J. 1997. Integrating Affective Computing into Animated Tutoring Agents. In *proceedings of IJCAI’97 Workshop on Intelligent Interface Agents*.
12. Aylett, R.S., Dias, J. and Paiva, A. (2006) An affectively-driven planner for synthetic characters. In *Proceedings of ICAPS 2006*.
13. Mateas, M. 2002. Ph.D. Thesis. Interactive Drama, Art and Artificial Intelligence. School of Computer Science, Carnegie Mellon University.
14. 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.
15. 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.
16. 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*.

17. 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.
18. 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.
19. Watson, D. & Tellegen, A. 1985. Toward a Consensual Structure of Mood. *Psychological Bulletin*, 98, 219-235.
20. Ekman, P. 1992. An Argument for Basic Emotions. In *Cognition and Emotion*, 6, 169-200.
21. Werry, C. 1996. Linguistic and Interactional Features of Internet Relay Chat. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*. Pragmatics and Beyond New Series 39. Amsterdam: John Benjamins, 47-64.
22. Briscoe, E. and 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. 1499-1504.
23. Heise, D. R. 1965. Semantic Differential Profiles for 1,000 Most Frequent English Words. *Psychological Monographs*. 70 8:(Whole 601).
24. Barnden, J., Glasbey, S., Lee, M. & Wallington, A. 2004. Varieties and Directions of Inter-Domain Influence in Metaphor. *Metaphor and Symbol*, 19(1), 1-30.
25. Barnden, J.A. Forthcoming. Metaphor, Semantic Preferences and Context-sensitivity. Invited chapter for a Festschrift volume. Kluwer.
26. Esuli, 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.
27. 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.
28. Garau, M., Slater, M., Bee, S. and Sasse, M.A. 2001. The impact of eye gaze on communication using humanoid avatars. *Proceedings of the SIG-CHI conference on Human factors in computing systems*, March 31 - April 5, 2001, Seattle, WA USA. 309-316.
29. Vilhjálmsson, H. & Cassell, J. 1998. BodyChat: Autonomous Communicative Behaviors in Avatars. *Proceedings of ACM Second International Conference on Autonomous Agents*, May 9-13, Minneapolis, Minnesota.
30. Gillies, M., Crabtree, I.B. and Ballin, D. 2006. Individuality and Contextual Variation of Character Behaviour for Interactive Narrative. In *Proceedings of the AISB Workshop on Narrative AI and Games*.
31. Carletta, J. 1996. Assessing Agreement on Classification Tasks: The Kappa statistic. *Computational Linguistics*, 22 (2), pp.249-254.
32. Pantic, M., Pentland, A., Nijholt, A. and Huang, T. 2006. Human Computing and Machine Understanding of Human Behavior: A Survey. In *Proc. Int'l Conf. Multimodal Interfaces*, pp. 239-248.
33. Cohn, J.F. 2007. Foundations of human-centered computing: Facial expression and emotion. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'07)* Hyderabad, India.
34. Nogueiras et al. 2001. Speech emotion recognition using hidden Markov models. In *Proceedings of Eurospeech 2001*, Denmark.
35. Weizenbaum, J. 1966. ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM* 9(1): 36-45.