EmoHeart: Automation of Expressive Communication of Emotions in Second Life

Alena Neviarouskaya[1], Helmut Prendinger[2], and Mitsuru Ishizuka[1]

[1] University of Tokyo, Department of Information and Communication Engineering, Japan
`lena@mi.ci.i.u-tokyo.ac.jp, ishizuka@i.u-tokyo.ac.jp`
[2] National Institute of Informatics, Japan
`helmut@nii.ac.jp`

Abstract. In this paper, we describe lexical rule-based approach to affect sensing from text, and application of the developed Affect Analysis Model in 3D virtual world Second Life. To enrich user experience in virtual environment, to automate emotional behaviour of avatar and to avoid thus manual control by user, we developed EmoHeart object that, driven by the result of Affect Analysis Model, triggers animations of avatar facial expressions and visualizes emotion by the heart-shaped textures.

Keywords: Affective computing, affective user interface, avatar, emotions, online communication, language parsing and understanding, text analysis.

1 Introduction and Motivation

The ability to express emotions in face-to-face communication is very important for establishing a social and friendly atmosphere. Trends show that people communicating through online media often try to enrich their interaction, introducing affective symbolic conventions into text (emoticons, capital letters etc.), or manually controlling the expressiveness of graphical representations of users (avatars). In order to achieve truly natural communication in virtual worlds, we set a two-fold focus in our research: (1) recognition of affective content conveyed through text; (2) automatic visualization of emotional expression of avatars.

Recently computational linguists demonstrate an increased interest in the tasks of text classification as subjective or of factual nature, of determination of orientation and strength of sentiment, and of recognition of attitude type expressed in text at various grammatical levels. A variety of approaches have been proposed to determine the polarity of distinct terms [1,2], phrases/sentences [3], and documents [4]. To analyse contextual sentiment, rule-based approaches [5,6], and a machine-learning method using not only lexical but also syntactic features [7] were proposed. Advanced approaches targeting textual affect recognition at the sentence level are described in [8,9,10,11].

One of the first attempts to study effects of conveying emotional expressions through communication in computer-mediated environment was done by Rivera et al.

[12]. The results of their experiment indicated that subjects allowed to use emoticons were more satisfied with the system than those subjects having conversations without these symbolic emotional expressions. The user study of ExMS [13], messaging system allowing its users to concatenate and annotate avatar animations, showed that interplay between pure text and animation significantly improved the expressiveness of messages, and that users felt pride of being identified with their embodied representation. However, the need for users to manually control their avatars can distract them from communication process. The twenty-person user study conducted on AffectIM [14], affective Instant Messenger (IM), showed that the IM system with automatic emotion recognition function was successful at conveying users' emotional states during communication online, thus enriching expressivity and social interactivity of online communications.

In this paper, we describe the application of Affect Analysis Model, responsible for affect sensing from text, in 3D virtual world Second Life. To automate emotional behaviour of avatar and to avoid thus manual control by user, we developed EmoHeart object that, driven by the result of Affect Analysis Model, triggers animations of facial expressions and visualizes emotion by the heart-shaped textures.

2   Recognition of Fine-Grained Emotions from Text

2.1   Basis for Affective Text Classification

For affective text classification, we decided to use the subset of emotional states defined by Izard [15]: 'anger', 'disgust', 'fear', 'guilt', 'interest', 'joy', 'sadness', 'shame', and 'surprise'. Besides emotions, we defined five communicative functions that are frequently observed in online conversations ('greeting', 'thanks', 'posing a question', 'congratulation', and 'farewell'). In the paper we mainly focus on the recognition of emotional state, and did not discuss the task of detection of communicative functions.

In order to support the handling of abbreviated language and the interpretation of affective features of lexical items, the Affect database was created (details are given in [16]). The Affect database includes the following tables: Emoticons, Abbreviations, Adjectives, Adverbs, Nouns, Verbs, Interjections, and Modifiers. Affective lexicon was mainly taken from WordNet-Affect [17]. Emotion categories with intensities were manually assigned to the emotion-related entries of the database by three independent annotators. Emotion intensity values range from 0.0 to 1.0. Emoticons and abbreviations were transcribed and related to named affective states (with intensity), whereby each entry was assigned to only one category (for ex., emoticon ':-S' [worry] was related to 'fear' with intensity 0.4). Considering the fact that some affective words may express more than one emotional state, annotators could relate words to more than one category (for ex., final annotation for noun '*enthusiasm*' is 'interest:08, joy:0.5'). Two annotators gave coefficients for intensity degree strengthening or weakening (from 0.0 to 2.0) to the adverbs of degree, and the result was averaged (for ex., coeff('*significantly*') = 2.0).

2.2 Affect Analysis Model

The affect sensing algorithm consists of five main stages: (1) symbolic cue analysis; (2) syntactical structure analysis; (3) word-level analysis; (4) phrase-level analysis; (5) sentence-level analysis. The Affect Analysis Model is capable of processing sentences of different complexity, including simple, compound, complex, and complex-compound sentences.

Symbolic Cue Analysis. In the first stage of the Affect Analysis Model, we test the sentence for emoticons, abbreviations, interjections, '?' and '!' marks, repeated punctuation, and capital letters. Several rules are applied to define the dominant emotion in cases when multiple emoticons and emotion-relevant abbreviations occur in a sentence. As interjections are added to sentence to convey emotion (e.g. '*Oh no*', '*wow*'), they are analysed as well. If there are no emotion-relevant emoticons or abbreviations in a sentence, we prepare the sentence for parser processing: emoticons and abbreviations relating to communicative function categories are excluded from the sentence; and non-emotional abbreviations are replaced by their proper transcriptions found in the database (e.g., '*I m* [am] *stressed bc* [because] *i have frequent headaches*'). In such a way, the issue of correct processing of abbreviated text by syntactical parser is resolved.

Syntactical Structure Analysis. The second stage is devoted to analysis of syntactical structure of sentences, and it is divided into two main subtasks: (1) sentence analysis by the Stanford Parser[1] [18] (in current research we employ it instead of commercial parser used in our previous work [16], as Stanford Parser is provided under GNU GPL) that returns word base forms (lemmas), parts of speech, and dependency functions representing relational information between words in sentences; (2) parser output processing. When handling the parser output, we represent the sentence as a set of primitive clauses (either independent or dependent). Each clause might include Subject formation (SF), Verb formation (VF) and Object formation (OF), each of which may consist of a main element (subject, verb, or object) and its attributives and complements. For the processing of complex or compound sentences, we build a so-called 'relation matrix', which contains information about dependences that the verbs belonging to different clauses have.

Word-Level Analysis. On the third stage, for each word found in our database, the affective features of a word are represented as a vector of emotional state intensities e=[anger, disgust, fear, guilt, interest, joy, sadness, shame, surprise] (e.g., e('*remorsefully*')=[0,0,0,0.8,0,0,0.5,0,0]). In the case of a modifier, the system identifies its coefficient (e.g., coeff('*barely*')=0.4). Our model also varies the intensities of emotional vectors of adjectives and adverbs in comparative or superlative forms (e.g., e('*glad*')=[0,0,0,0,0,0.4,0,0,0], e('*gladder*')= [0,0,0,0,0,0.48,0,0,0] and e('*gladdest*')=[0,0,0,0,0,0.56,0,0,0]).

---

[1] Stanford Parser: http://nlp.stanford.edu/software/lex-parser.shtml

Phrase-Level Analysis. The purpose of this stage is to detect emotions involved in phrases, and then in SF, VF, and OF. We have defined rules for processing phrases:

1. Adjective phrase: modify the vector of adjective (e.g., e('*extremely doleful*') = coeff('*extremely*') * e('*doleful*') = 2.0 * [0,0,0,0,0,0,0.4,0,0] = [0,0,0,0,0,0,0.8,0,0]).
2. Noun phrase: output vector with the maximum intensity within each corresponding emotional state in analysing vectors (for instance, e1=[0..0.7..] and e2=[0.3..0.5..] yield e3=[0.3..0.7..]).
3. Verb plus adverbial phrase: output vector with the maximum intensity within each corresponding emotional state in analysing vectors (e.g., e('*shamefully deceive*')=[0,0.4,0,0,0,0,0.5,0.7,0] where e('*shamefully*')=[0,0,0,0,0,0,0,0.7,0] and e('*deceive*')= [0,0.4,0,0,0,0,0.5,0,0]).
4. Verb plus noun phrase: if verb and noun phrase have opposite valences (e.g., '*break favourite vase*', '*enjoy bad weather*'), consider vector of verb as dominant; if valences are the same (e.g., '*like honey*', '*hate crying*'), output vector with maximum intensity in corresponding emotional states.
5. Verb plus adjective phrase (e.g., '*is very kind*', '*feel bad*'): output vector of adjective phrase.

The rules for modifiers are as follows: (1) adverbs of degree multiply or decrease emotional intensity values; (2) negation modifiers such as 'no', 'never', 'nothing' etc. cancel (set to zero) vectors of the related words, i.e., 'neutralize the emotional content' (e.g., '*Yesterday I went to a party, but nothing exciting happened there*'); (3) prepositions such as 'without', 'except', 'against', 'despite' cancel vectors of related words (e.g., '*I climbed the mountain without fear*' is neutralized due to preposition).

Statements with prefixed words like 'think', 'believe', 'sure', 'know', 'doubt' or with modal operators such as 'can', 'may', 'would' etc. are neutralized by our system. Conditional clause phrases beginning with 'even though', 'if', 'unless', 'whether', 'when', etc. are neutralized as well (e.g., '*I eat when I'm angry, sad, bored…*').

Each of the Subject, Verb, or Object formations may contain words conveying emotional meaning. During this stage, we apply the described rules to phrases detected within formation boundaries. Finally, each formation can be represented as a unified vector encoding its emotional content.

Sentence-Level Analysis. The emotional vector of a simple sentence (or a clause) is generated from Subject, Verb, and Object formation vectors resulting from phrase-level analysis. The main idea here is to first derive the emotion vector of Verb-Object formation relation. It is estimated based on the 'verb plus noun phrase' rule described above. In order to apply this rule, we automatically determine valences of Verb and Object formations using their unified emotion vectors (particularly, non-zero-intensity emotion categories). The estimation of the emotion vector of a clause (Subject plus Verb-Object formations) is then performed in the following manner: (1) if valences of Subject formation and Verb formation are opposite (e.g., SF = '*my darling*', VF = '*smashed*', OF '*his guitar*'; or SF = '*troubled period*', VF = '*luckily comes to an end*'), we consider the vector of the Verb-Object formation relation as dominant; (2) otherwise, we output the vector with maximum intensities in corresponding emotional states of vectors of Subject and Verb-Object formations. Fig. 1 shows example of processing a simple sentence '*My darling smashed his favourite guitar without regret*'.

	word:	word-level:	phrase-level:
SF:	*my*	e^0=[0,0,0,0,0,0,0,0,0]	e^+=[0,0,0,0,0,0.7,0,0,0]
	darling	e^+=[0,0,0,0,0,0.7,0,0,0]	
VF:	*smashed*	e^-=[0,0,0.6,0,0,0,0.8,0,0]	e^-=[0,0,0.6,0,0,0,0.8,0,0]
	without	modif. coeff=0.0	e^0=[0,0,0,0,0,0,0,0,0]
	regret	e^-=[0,0,0,0.2,0,0,0.1,0,0]	
OF:	*his*	e^0=[0,0,0,0,0,0,0,0,0]	e^0=[0,0,0,0,0,0,0,0,0]
	favourite	e^+=[0,0,0,0,0,0.6,0,0,0]	e^+=[0,0,0,0,0,0.6,0,0,0]
	guitar	e^0=[0,0,0,0,0,0,0,0,0]	e^0=[0,0,0,0,0,0,0,0,0]

Phrase-level combined:
- SF: e^+=[0,0,0,0,0,0.7,0,0,0]
- VF: e^-=[0,0,0.6,0,0,0,0.8,0,0]
- OF: e^+=[0,0,0,0,0,0.6,0,0,0]

sentence-level:
1. (SF$^+$ and VF$^-$) yields domination of (VF and OF);
2. (VF$^-$ and OF$^+$) yields domination of VF;
3. e (sentence) = e (VF$^-$) = [0,0,0.6,0,0,0,0.8,0,0];
4. e (sentence) * coeff (tense:'past'; FPP:'yes') = [0,0,0.6,0,0,0,0.8,0,0] * 0.8 = [0,0,0.48,0,0,0,0.64,0,0];
5. result ('*My darling smashed his favourite guitar without regret*'): sadness:0.64.

Fig. 1. Example of affect sensing in a simple sentence

In Fig. 1 e = [anger, disgust, fear, guilt, interest, joy, sadness, shame, surprise]; the superscripts ^{0}, $^-$, $^+$ indicate 'neutral', 'negative', 'positive' valences, respectively.

In order to estimate the emotional vector of a compound sentence, first, we evaluate the emotional vectors of its independent clauses. Then, we define the resulting vector of the compound sentence based on two rules: (1) with comma and coordinate connectors 'and' and 'so' (e.g., '*It is my fault, and I am worrying about consequences*', '*Exotic birds in the park were amazing, so we took nice pictures*'), or with a semicolon with no conjunction: output the vector with the maximum intensity within each corresponding emotional state in the resulting vectors of both clauses; (2) with coordinate connector 'but' (e.g., '*They attacked, but we luckily got away!*'): the resulting vector of a clause following after the connector is dominant.

In order to process a complex sentence with a complement clause (e.g., '*I hope that Sam will not harass my dog*'), first we derive the emotional vector of the complement clause, then create Object formation for the main clause using this vector, and finally estimate the resulting emotional vector of the main clause with added Object formation. In brief, we represent such sentence as a simple one, using the following pattern: 'who-subject does-verb what-object', where object is represented as a complement clause. In our algorithm, the complex sentences containing adjective (relative) clauses are analysed in the following manner: (1) the emotional vector of adjective clause is estimated; (2) this emotional vector is added to the SF or OF of the main clause depending on the role of the word to which the adjective clause relates; and (3) the emotional vector of the whole sentence is estimated.

While processing complex-compound sentences (e.g., '*Max broke the china cup, with which Mary was awarded for the best song, so he regretted profoundly*'), first we generate emotional vectors of dependent clauses, then of complex sentences, and finally, we analyse the compound sentence formed by the independent clauses.

It is important to note that our system enables the differentiation of the strength of the resulting emotion depending on the tense of a sentence and availability of first person pronouns. The dominant emotion of the sentence is determined according to the emotion state with the highest intensity within the final emotional vector.

3   EmoHeart

Emotional expression is natural and very important for communication in real life, but currently rather cumbersome in 3D virtual world Second Life, where expressions have to be selected and activated manually. In order to automate the emotional expressiveness of graphical representations of users (avatars), we applied the developed Affect Analysis Model to textual chat in Second Life. The architecture of the system is presented in Fig. 2.
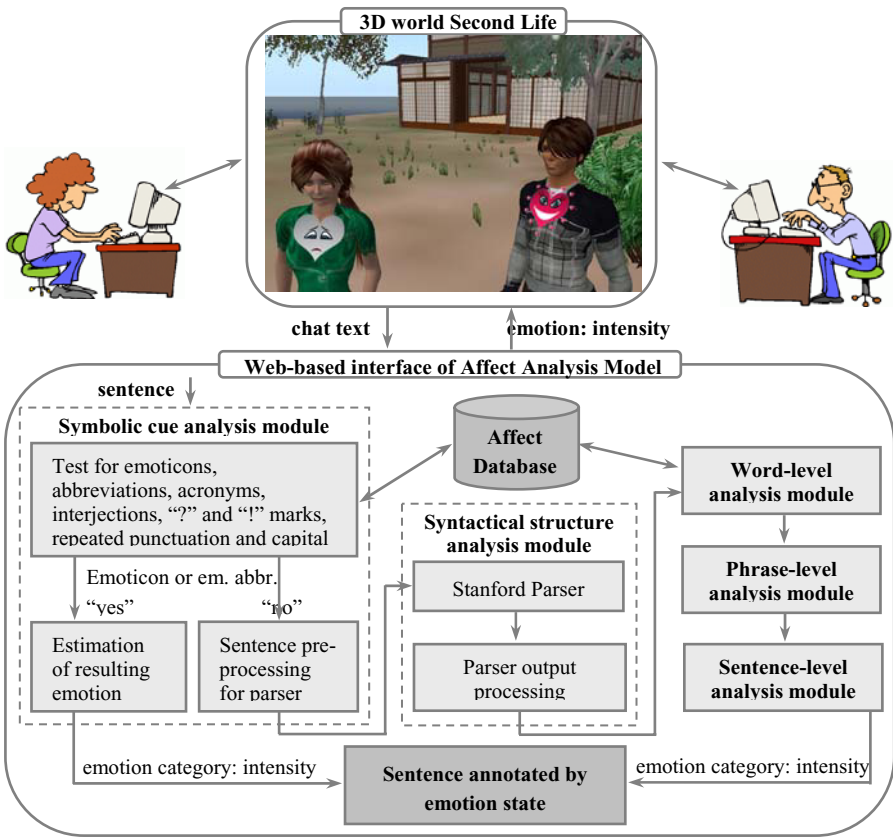

Fig. 2. Architecture of the EmoHeart system

The control of the conversation is implemented through the object called EmoHeart[2]. (invisible in case of 'neutral' state) attached to the avatar's chest. The distributor of the EmoHeart object is located inside a (fictitious) Starbucks cafe[3] of the Second Life replica of National Center of Sciences building. in Tokyo. Once attached to the avatar, EmoHeart object (1) listens to each message of its owner, (2) sends it to the web-based interface of the Affect Analysis Model located on the server, (3) receives the result (dominant emotion and intensity), and visually reflects the sensed affective state through the animation of avatar's facial expression, EmoHeart texture (indicating the type of emotion), and size of the texture (indicating low, middle, or high level of emotion strength). If no emotion is detected in the text, the EmoHeart remains invisible and the facial expression remains neutral. With the heart-shaped object of EmoHeart, we provide an additional channel for visualizing emotions in a vivid and intense way. The examples of avatar facial expressions and EmoHeart textures are shown in Fig. 3.
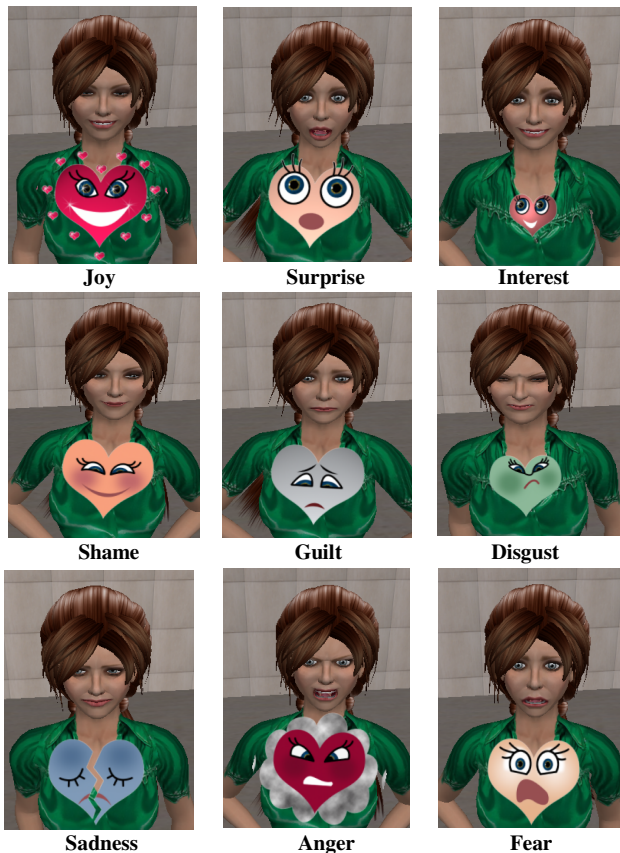

Fig. 3. Examples of avatar facial expressions and EmoHeart textures

While designing EmoHeart textures, we followed the description of main characteristic features of expressive means in relation to communicated emotion (Table 1).

Table 1. Emotional states and relevant expressive means (data partially taken from [19])

Emotion	Expressive means
anger	widely open eyes, fixated; pupils contracted; stare gaze; ajar mouth; teeth usually clenched tightly; rigidity of lips and jaw; lips may be tightly compressed, or may be drawn back to expose teeth
disgust	narrowed eyes, may be partially closed as result of nose being drawn upward; upper lip drawn up; pressed lips; wrinkled nose; turn of the head to the side quasi avoiding something
fear	widely open eyes; pupils dilated; raised eyebrows; open mouth with crooked lips; trembling chin
guilt	downcast or glancing gaze; inner corners of eyebrows may be drawn down; lips drawn in, corners depressed; head lowered
interest	eyes may be exaggeratedly opened and fixed; lower eyelids may be raised as though to sharpen visual focus; increased pupil size; sparkling gaze; mouth slightly smiling; head is slightly inclined to the side
joy	'smiling' and bright eyes; genuinely smiling mouth
sadness	eyelids contracted; partially closed eyes; downturning mouth
shame	downcast gaze; blushing cheeks; head is lowered
surprise	widely open eyes; slightly raised upper eyelids and eyebrows; the mouth is opened by the jaw drop; the lips are relaxed

4 Conclusion

The richness of the information cues is a salient aspect of effective interpersonal communication. This holds true also within the virtual environment context. Researchers argue that there is a positive relationship between the amount of IM use and verbal, affective, and social intimacy [20]. Marcel [21] characterized the intimacy in the following words: "Even if I cannot see you, if I cannot touch you, I feel that you are with me". In our work, we strive to provide vivid and expressive visual signals to enhance socially oriented online communication media.

This paper introduced the integration of the developed emotion recognition module, Affect Analysis Model, into 3D virtual world Second Life. The proposed rule-based algorithm to affect sensing from text enables analysis of nine emotions at various grammatical levels. For textual input processing, our Affect Analysis Model handles not only correctly written text, but also informal messages written in an abbreviated or expressive manner.

The control of the conversation in Second Life is implemented through the EmoHeart object attached to the avatar's chest. This object communicates with Affect Analysis Model located on the server, and visually reflects the sensed affective state through the animation of avatar's facial expression, EmoHeart texture, and size of the texture.

References

1. Turney, P.D., Littman, M.L.: Measuring Praise and Criticism: Inference of Semantic Orientation from Association. ACM Transactions on Information Systems 21(4), 315–346 (2003)
2. Andreevskaia, A., Bergler, S.: Mining WordNet for Fuzzy Sentiment: Sentiment Tag Extraction from WordNet Glosses. In: Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics, pp. 209–216 (2006)
3. Kim, S.-M., Hovy, E.: Determining the Sentiment of Opinions. In: Proceedings of Conference on Computational Linguistics, pp. 1367–1373 (2004)
4. Nadeau, D., Sabourin, C., De Koninck, J., Matwin, S., Turney, P.D.: Automatic Dream Sentiment Analysis. In: Proceedings of the Workshop on Computational Aesthetics at the Twenty-First National Conference on AI (2006)
5. Nasukawa, T., Yi, J.: Sentiment Analysis: Capturing Favorability using Natural Language Processing. In: Proceedings of Second International Conference on Knowledge Capture, pp. 70–77 (2003)
6. Moilanen, K., Pulman, S.: Sentiment Composition. In: Proceedings of the Recent Advances in Natural Language Processing International Conference, pp. 378–382 (2007)
7. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In: Proceedings of HLT/EMNLP Conference, pp. 347–354 (2005)
8. Boucouvalas, A.C.: Real Time Text-to-Emotion Engine for Expressive Internet Communications. In: Being There: Concepts, Effects and Measurement of User Presence in Synthetic Environments, pp. 306–318 (2003)
9. Liu, H., Lieberman, H., Selker, T.: A Model of Textual Affect Sensing using Real-World Knowledge. In: Proceedings of IUI 2003, pp. 125–132 (2003)
10. Mulder, M., Nijholt, A., den Uyl, M., Terpstra, P.: A Lexical Grammatical Implementation of Affect. In: Sojka, P., Kopeček, I., Pala, K. (eds.) TSD 2004. LNCS, vol. 3206, pp. 171–177. Springer, Heidelberg (2004)
11. Alm, C.O.: Affect in Text and Speech. Ph.D. diss., University of Illinois at Urbana-Champaign (2008)
12. Rivera, K., Cooke, N.J., Bauhs, J.A.: The Effects of Emotional Icons on Remote Communication. In: Electronic Proceedings of CHI (1996)
13. Persson, P.: ExMS: an Animated and Avatar-Based Messaging System for Expressive Peer Communication. In: Proceedings of SIGGROUP Conference, pp. 31–39 (2003)
14. Neviarouskaya, A., Prendinger, H., Ishizuka, M.: User Study of AffectIM, an Emotionally Intelligent Instant Messaging System. In: Prendinger, H., Lester, J.C., Ishizuka, M. (eds.) IVA 2008. LNCS, vol. 5208, pp. 29–36. Springer, Heidelberg (2008)
15. Izard, C.E.: Human emotions. Plenum Press, New York (1977)
16. Neviarouskaya, A., Prendinger, H., Ishizuka, M.: Textual Affect Sensing for Sociable and Expressive Online Communication. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 218–229. Springer, Heidelberg (2007)
17. Strapparava, C., Valitutti, A.: WordNet-Affect: an Affective Extension of WordNet. In: Proceedings of LREC 2004, pp. 1083–1086 (2004)
18. Klein, D., Manning, C.D.: Fast Exact Inference with a Factored Model for Natural Language Parsing. In: Advances in Neural Information Processing Systems, vol. 15, pp. 3–10. MIT Press, Cambridge (2003)
19. Izard, C.E.: The face of emotion. Appleton-Century-Crofts, New York (1971)
20. Hu, Y., Wood, J.F., Smith, V., Westbrook, N.: Friendships through IM: Examining the Relationship between Instant Messaging and Intimacy. Journal of Computer-Mediated Communication 10(1) (2004)
21. Marcel, G.: The Philosophy of Existence. Books for Libraries Press, Freeport (1969)