What Computers Can Tell Us About Emotions – Classification of Affective Communication in Electronic Negotiations by Supervised Machine Learning

Michael Filzmoser(✉), Sabine T. Koeszegi, and Guenther Pfeffer

TU Wien, Institute of Management Science, Theresianumgasse 27, 1040 Vienna, Austria
{michael.filzmoser, sabine.koeszegi, guenther.pfeffer}@tuwien.ac.at

Abstract. Affective communication and emotions are an important part of negotiations. Negotiation support and negotiation support systems, however, tend to neglect this aspect given extant measurement difficulties. This study explores the possibilities of state of the art supervised machine learning techniques to classify emotions expressed in negotiation communication during electronic negotiation experiments. The affective content of the exchanged messages was determined by human coders and classified according to the circumflex model of affect. The output of this laborious activity, that can only be accomplished after a negotiation, which makes it irrelevant for negotiation support, was input to this study. Promising performance of some preprocessing and machine learning techniques was achieved. Especially the category of activating negative emotions, which is highly important in negotiations as it might reduce the prospects of reaching an agreement, was correctly classified quite often.

Keywords: Affect · Electronic negotiations · Machine learning

1 Introduction

Despite a focus on analytic aspects in negotiation research, the significant role of affect in negotiation processes and outcomes has been acknowledged among negotiation scholars (for overviews see e.g. [1, 2]). Particularly a research team around Van Kleef and De Dreu has analyzed the impact of various emotions such as anger, happiness, worry, guilt, regret, disappointment, etc. on negotiation processes and outcomes (e.g. [2–5]). However, in electronic negotiation support research, affect and emotions have been understudied so far [6–8].

In computer-mediated communication, affectivity is the sensitivity to attitudes toward the communication partner or the subject matter in a communication and denotes the inclusion of affective components in a (text) message. Affect comprises emotions, which are directed towards specific situational stimuli, of shorter duration

D. Bajwa et al. (Eds.): GDN 2016, LNBIP 274, pp. 113–123, 2017.
DOI: 10.1007/978-3-319-52624-9_9

and higher intensity and moods, which lack the quality of directedness but are more enduring and pervasive [2]. High affective complexity is associated with relational oriented obstacles such as mistrust and affective disruptions and therefore needs to be considered when deciding on the communication and negotiation strategy [9]. Especially in text-based negotiations affectivity can not only be indicative – i.e. consistent with, and thereby revealing, the affective state of a person – but also instrumental and therefore used strategically in the negotiation – e.g. expression of anger to elicit concessions [2].

Affect needs to be encoded or contextualized differently in computer-mediated communication compared to face-to-face communication with available non- and para-verbal cues. One possibility to contextualize emotions in text messages are emoticons (standing for emotion and icon) which are referred to relation icons, visual cues or pictographs and serve as surrogates for non-verbal communication to express emotion [10]. Additionally, communicators can use contextualization cues such as non-standard spelling, letter and punctuation mark repetition (e.g. '???') or lexical surrogates ('hmmm') and the like as linguistic form to express affect. All these cues contribute to the signaling of "contextual presuppositions" that allow for inferences about the meanings communicators intend to convey in a specific situation [11]. However, a substantial proportion of affective content is encoded implicitly in factual statements by communicators' lexical and syntactical choices. Not only what negotiators convey in their messages (content or substantial dimension) but also how they express themselves (affective dimension) substantially impacts the relationship and trust building between negotiators [12].

Te'ini therefore suggests a computerized support of communication strategies through e.g. templates of appropriate affectivity and feedback on current messages (e.g. language checks) [9]. Also Broekens et al. call for the development of negotiation support systems that also consider the affective dimensions [6]. The knowledge of the affective content of messages by negotiation support systems (NSS) or software agents would enable novel ways of supporting and automating negotiations. NSS could for example make the user aware of the affective content of own messages and messages of the opponent and thereby support the negotiator in a similar way to offer evaluation and generation [13]. Software agents could react not only to the offer behavior but also to affect explicated in messages of their human counterparts in semi-automated negotiations [14, 15].

This requires, first of all, the identification of affectivity in texts which is challenging because of the particularities of computer-mediated communications discussed above. In this paper we, therefore, focus on the identification and classification of affect in text-based negotiation messages by means of machine learning. The research question of this explorative study, therefore, is: "To what extent and in which quality are state of the art text preprocessing and supervised machine learning techniques able to classify affective communication in electronic negotiations?" To address this question we evaluate the performance of available techniques in supervised machine-learning, i.e. their ability to correctly assign electronic negotiation messages to the affective categories they were assigned to by human coders.

The remainder of this paper is structured as follows: Sect. 2 offers a brief theoretical background on affect classification, and Sect. 3 presents the data from electronic negotiation experiments applied in this data-driven approach. The negotiation case, the NSS applied and the coding and multi-dimensional scaling analysis to assign the messages to affective categories are also discussed in this section. Furthermore it introduces the preprocessing and supervised machine-learning techniques evaluated in this study as well as the experimental design. Section 4 presents and discusses the results of our study and derives suggestions for parameterization and algorithms for affect identification and classification in electronic negotiations. Section 5 concludes with a summary of the main findings and a discussion of future research.

2 Theoretical Background

An issue to resolve with regard to affect identification and classification is the potential complexity of emotion patterns. Even though [16] only differentiated between seven basic emotions (sadness, anger, happiness, contempt, fear, disgust, and surprise) hundreds of facets of emotions and emotion-related states have been identified in literature. We therefore suggest employing a dimensional model as suggested in [17]. In this two-dimensional perspective of affect, all emotions and emotion-related states can be represented by the two underlying bipolar affective dimensions of (i) valence (pleasure vs. displeasure) and (ii) degree of activation (high vs. low) [17–19] see Fig. 1 (adapted from [20: p. 141]).

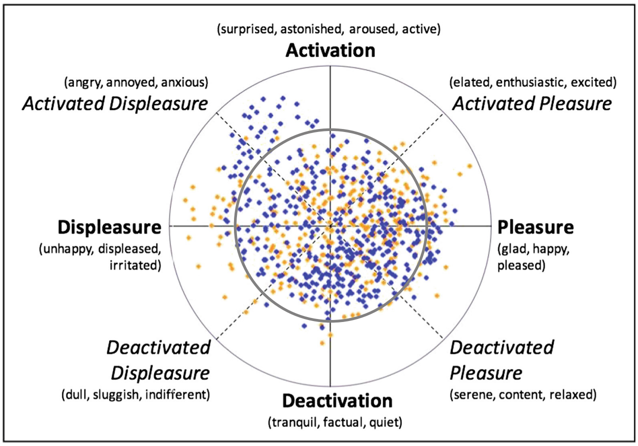

Fig. 1. Affectivity of messages in negotiations with (yellow) and without (blue) agreement (Color figure online)

In contrast to approaches based on discrete single emotions, which were often employed in previous work [21, 22], a dimensional approach provides a compact representation of the (implicit or explicit) "emotion quality" of each communication utterance in a two-dimensional Cartesian space and is preferable for the analysis of conversational settings [23, 24]. We therefore suggest identifying affect in negotiations by measuring the two dimensions of affective behavior, i.e. valence and arousal, manifest in communication behavior.

3 Data and Method

For our analyses, we used data from a previous negotiation experiment conducted with the NSS Negoisst [25], a web-based system that offers both decision and communication support. Participants in the negotiation experiment, students from negotiation courses of four European universities, represented either a Western European or an Eastern European company in, high conflict narrow zone of possible agreement, bilateral joint venture negotiations. The case contained seven issues with several continuous and discrete options and therefore was quite complex. The NSS recorded all exchanged offers and messages. A total of 57 negotiations between 114 negotiators were conducted from which 38 reached an agreement while 19 failed to reach an agreement. In all 57 negotiations a total of 730 messages were exchanged.

The messages from the electronic negotiations were first grouped according to their affective similarity by 26 unbiased business students. Each student received up to 250 of the 730 messages and working instructions that indicated that they had to build decks (no limit on the number of decks was provided) with similar messages. The similarity between messages – measured in the number of times these messages occurred in the same deck – was input to a multi-dimensional scaling analysis. These analyses were part of another study [20] and build the base data set for the analysis of possibilities of machine learning to identify affect in electronic negotiations in this study. The multi-dimensional scaling data was used to derive five affect categories for the negotiation messages (neutral, activated pleasure, deactivated pleasure, activated displeasure and deactivated displeasure) based on the values of valence and activation in the circumplex model of affect [17] of the message. The base data set of these categories and messages, after necessary prepossessing of the text, in a last step was used to train and evaluate different machine learning techniques. The following subsections describe the steps of the study in detail.

3.1 Multi-dimensional Scaling

As already mentioned in Sect. 2 a dimensional approach provides a compact representation of the affect of each message in a two-dimensional Cartesian space. This is preferable for the analysis of conversations [23, 24] like negotiations to approaches based on discrete single emotions. The evaluation of the similarity of the affective content of messages by human raters in a three step multi-dimensional scaling procedure builds the basis for the analyses of the subsequent sections.

In a first step the input data for multi-dimensional scaling is generated. For this purpose human raters evaluated the affective similarity of the negotiation messages exchanged. 26 business students participated in this rating activity, they received no background information about the underlying study but detailed instructions to rate each up to 250 of the 730 messages. The task of the raters was to sort similar messages into the same deck. For this task the raters received no additional training or instructions, like coding schemes, number of decks, etc.

This data built the basis for multi-dimensional scaling based on the proximity of two messages, which was measured by the number of raters who assigned them to the same deck. The proximity matrix was processed by the multi-dimensional scaling software PERMAP 11.8a using nonparametric multi-dimensional scaling with Euclidean distances as distance measures, the preferable approach for proximity measures based on subjective judgments [26].

Goodness of fit (Stress-1) and the interpretation of possible dimensions indicated that a two-dimensional Cartesian space best fitted the data. Rotation of the axes lead to the two dimensions of valence and activation, which bring the results of the multi-dimensional scaling in accord with the circumplex model of affect [17]. A detailed description of the multi-dimensional scaling procedure can be found in [27].

For the categorization task of supervised machine-learning the data has to be distinguished and labeled into discrete classes which were determined according to the four sectors of the circumplex model of affect plus a neutral class which contains messages in the center and therefore of low affectivity. This resulted in approximately equal amounts of messages in all five classes as represented in Fig. 2.

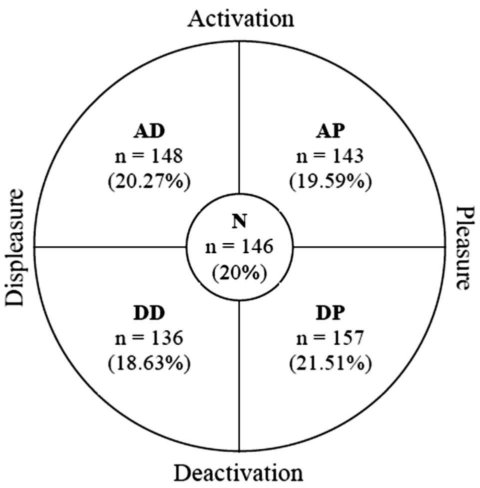

Fig. 2. Classes of affective communication and number of observations.

The adequacy of the existing variety of techniques for data preprocessing and supervised machine-learning for classification of affective content of electronic negotiation messages at present is unclear. There are some promising applications of machine-learning for sentiment analysis, i.e. the evaluation of whether there is a general positive or negative feeling towards an issue from blogs, newspapers, forums or stock reports [28]. However, there are significant differences between sentiment analysis and analysis of the affectivity of electronic negotiation messages, which hinder the direct application of these methods to the field of electronic negotiations. On the one hand the available amount of data from forums, blogs, etc. for machine-learning is considerably larger, on the other hand for sentiment analysis the classification need not be as detailed, a general tendency is sufficient while more concrete classifications of dimension of affectivity are considered in this study. This calls for a systematic comparison of the available options both for data preprocessing and machine-learning. We subsequently briefly describe these techniques, which are all implemented in WEKA, an open source software that implements various machine-learning techniques for (big-data) data-mining purposes, which was applied for the analyses of this study.

3.2 Data Preprocessing Techniques

Identification of emotions in electronic negotiation messages is basically a text mining task. Text mining is a variant of data mining. However, the data used in data mining is usually more structured than communication transcripts. To make the established algorithms from data mining available to text mining the complexity and variety of the data has to be reduced. A variety of techniques exist for this purpose. Figure 3 gives an overview of the data preparation.

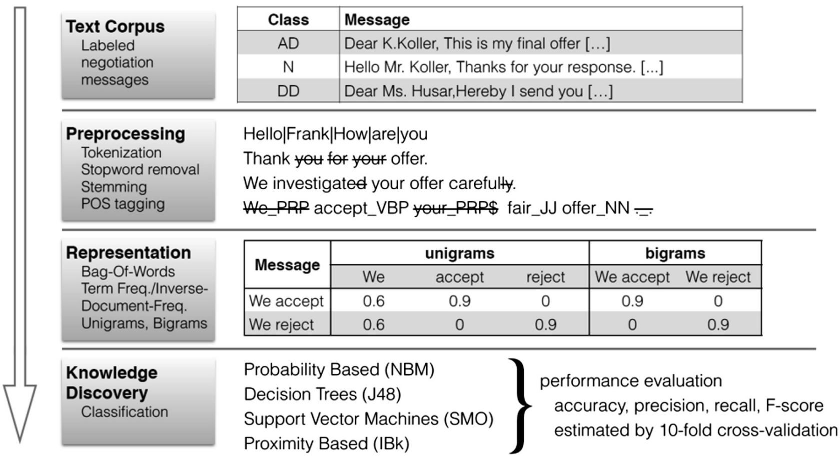

Fig. 3. Overview data preparation

Stemming is the reduction of different times and forms of words to their roots, like e.g. ‘is’, ‘was’ and ‘am’ to ‘be’. This complexity reduction technique thereby improves the performance of word mining algorithms as it increases similarities. Stopword removal eliminates the most often used words (‘the’, ‘a’, ‘and’, etc.) from the data set These stopwords are often equally present in all classes so that they do not add to the discriminative power of an algorithm but are rather noise, which can be filtered out with stopword removal. Stopwords can be taken from general lists (e.g. Swish-E) or be the most frequent (10, 50, 100, etc.) words in the data set. N-grams are word combinations, e.g. of two words like ‘hello dear’, ‘hello sir’, or ‘dear sir’ in the case of bi-grams, rather than just single words, like ‘hello’, ‘dear’ or ‘sir’ and are often more informative for classification algorithms than the single words. Part of speech (POS) tagging categorizes the original text into grammatical classes e.g. verbs, nouns and adjectives and replaces the words by these classes to reduce variance and facilitate pattern recognition.

3.3 Machine-Learning Techniques

After application of the data preprocessing techniques, to reduce variance and facilitate classification, machine-learning algorithms perform the actual classification task. Naive Bayes is a probabilistic classifier based on Bayes’ theorem. The hypothesis that an object belongs to a certain class is updated by learning and the actual assignment of an object to a class bases on the probability that the different hypotheses are true. Decision tree algorithms develop an internal hierarchy of nodes and arcs, where the former represent decision points and the latter stand for the classes. The trained model is the result of creating a tree with maximum discriminatory power of each decision node. The support vector machine approach fits during the training phase mathematical functions to the multi-dimensional feature space which are then used for classification. Proximity approaches like (k-) nearest neighbor are ‘lazy’ learning approaches. The (k) most similar objects from the – already classified – training data set to the focal object of the test set are identified and the test object is assigned to the most frequent class elicited this way. Machine-learning algorithms are trained on a training data set and then tested on a test data set. In this study we separate the total data set of 730 messages in ten subsets from which in ten runs each subset is used as test data set with the remaining nine data sets as training data set.

3.4 Experimental Design and Measurement

The data preprocessing techniques are combined to a total of 16 experimental settings (#01 to #16), illustrated in Table 1. The resulting data sets are then input to the four machine-learning techniques discussed above: Naïve Bayes (NBM), decision tree (J48), support vector machines (SMO) and nearest neighbor (lBk).

For the comparison of the performance of the machine-learning algorithms we apply the four measures suggested in [29]. The accuracy of an algorithm (1) is the percentage of correctly identified messages.

$$accuracy = \frac{correct}{total} \tag{1}$$

Accuracy, however, does provide little information about the characteristics of the classification errors. An algorithm can correctly (true) assign a message that belongs to the focal class to this class (true positive) or not to other classes (true negative), as well as it can incorrectly (false) assign a message that belongs to another class to the focal one (false positive) or that belongs to the focal class to other classes (false negative). Based on these correct and incorrect classifications and error types the precision (2) and recall (3) rations of the algorithm can be determined.

Table 1. Experiment settings – data preprocessing techniques.

Setting	Dataset	n-grams	Stopwords	Stemmer
#01	original	unigram	none	none
#02	original	unigram	none	Porter
#03	original	unigram	Swish-E	none
#04	original	unigram	Swish-E	Porter
#05	original	unigram	Top 50	none
#06	original	unigram	Top 50	Porter
#07	original	uni- & bigram	none	none
#08	original	uni- & bigram	none	Porter
#09	original	uni- & bigram	Swish-E	none
#10	original	uni- & bigram	Swish-E	Porter
#11	original	uni- & bigram	Top 50	none
#12	original	uni- & bigram	Top 50	Porter
#13	POS adjusted	unigram	none	none
#14	POS adjusted	unigram	Top 50	none
#15	POS adjusted	uni- & bigram	none	none
#16	POS adjusted	uni- & bigram	Top 50	none

$$precision = \frac{true\,positive}{true\,positive + false\,positive} \tag{2}$$

$$recall = \frac{true\,positive}{true\,positive + false\,negative} \tag{3}$$

Precision and recall are interdependent, therefore, the so-called f-score (4) is used as a harmonic mean of precision and recall, which is typically used as overall performance measure for machine-learning algorithms.

$$f - score = \frac{2 \times precision \times recall}{precision + recall} \tag{4}$$

As a conservative baseline for the evaluation of the accuracy of the machine-learning algorithms we define the percentage of the largest class (DP deactivation-pleasure) with 21.51% of the messages. This accuracy would be achieved by a plain algorithm that assigning all messages to the largest class, a random assignment to one of the five classes would lead to a slightly more generous baseline of 20% only.

4 Results

As can be seen from Fig. 4 all machine-learning techniques are significantly above the comparison baseline of 21.51% accuracy, which would be achieved by assigning all messages to the class DP which has the highest share. Moreover, on the one hand significant performance differences between machine-learning techniques exist and on

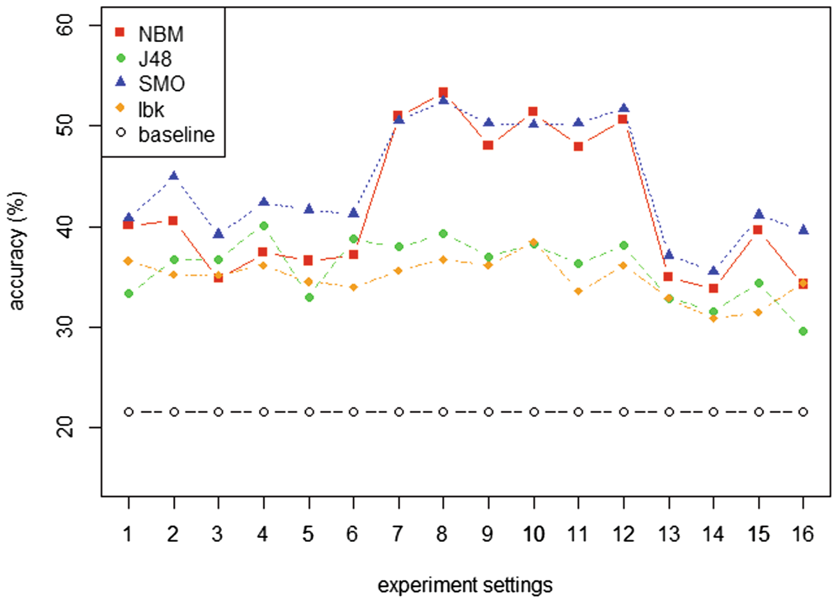

Fig. 4. Classification accuracy overview.

the other hand significant interaction effects exist between machine-learning and pre-processing techniques that have a major impact on performance.

Support vector machines and Naïve Bayes perform approximately equal and better than decision tree or nearest neighbor approaches. Furthermore they are especially good when combined with bi-grams, stemming and no stopword removal or POS adjustment.

Besides this overall performance to correctly classify all five classes the detailed performance per class is also of interest. Especially the activating negative emotions are critical for negotiation success as they might lead to negotiation break-offs. Table 2 presents the detailed classification results for all five categories of the two best performing machine-learning techniques (Naïve Bayes and support vector machines) combined with the best performing data preprocessing approaches (#08: no POS adjustment, usage of uni- and bi-grams, no stopword exclusion and usage of a stemmer).

Table 2. Detailed results for algorithms.

	#08.NBM			#08.SMO		
Class	Precision	Recall	f-score	Precision	Recall	f-score
N	63.6%	43.2%	51.4%	50.7%	47.3%	48.9%
AP	53.9%	48.3%	50.9%	55.6%	51.7%	53.6%
AD	62.4%	56.1%	59.1%	62.2%	53.4%	57.5%
DD	44.4%	58.1%	50.3%	44.5%	53.7%	48.7%
DP	49.5%	60.5%	54.4%	51.8%	56.1%	53.8%
Weighted avg.	54.8%	53.3%	53.3%	53.1%	52.5%	52.6%

5 Conclusion

Emotions are crucial in negotiations. To further-develop existing NSS towards proactive negotiation support that also considers the affective component of communication the possibility to identify this affective component is a mandatory prerequisite. To automate this task the knowledge set derived from machine learning would be helpful. The aim of this paper, therefore, was to explore the possibilities to classify affective content of electronic negotiation messages by means of supervised machine-learning. For this purpose we compared extant text preprocessing and machine-learning techniques in an explorative study.

Our study found significant performance differences of text mining algorithms for the identification and classification of affect in electronic negotiation messages. Moreover performance relevant interaction effects between data preprocessing techniques and classification algorithms were identified. Naïve Bayes and support vector machines are the two approaches that seem better suited for this endeavor than available alternatives. Stemming and bi-grams are relevant data preprocessing techniques, while others are not suited for the purposes of affectivity classification (i.e. POS adjustment and stopword removal). The potential discriminatory power of even higher dimensional n-grams is one avenue of necessary further research.

The achieved classification accuracy of nearly 55% is a promising initial result, and the accuracy of over 62% for the important activating negative category even more so. However, the performance is still not satisfying for the actual implementation in NSS. Additional research is necessary to achieve this ultimate goal. Especially more coded data, which is laborious work, and more textual indicators for affect in negotiation messages are necessary to establish a convenient training data set. The 730 messages from 57 negotiations are a relatively small data set compared to the 'big data' problems for which machine-learning is usually applied. This data should also be more diverse, featuring different levels of conflict different types of negotiations etc. to avoid overfitting for a specific negotiation problem.

References

1. Kumar, R.: The role of affect in negotiations: an integrative overview. J. Appl. Behav. Sci. **33**(1), 84–100 (1997)
2. Barry, B., Fulmer, I.S., van Kleef, G.A.: I laughed, I cried, I settled: the role of emotion in negotiation. In: Gelfand, M.J., Brett, J.M. (eds.) The Handbook of Negotiation and Culture, pp. 71–94. Stanford University Press, Palo Alto (2004)
3. van Kleef, G.A., de Dreu, C.K.W., Manstead, A.S.R.: The interpersonal effects of anger and happiness in negotiations. J. Pers. Soc. Psychol. **86**(1), 57–76 (2004)
4. van Kleef, G.A., de Dreu, C.K.W., Manstead, A.S.R.: The interpersonal effects of emotions in negotiations: a motivated information processing approach. J. Pers. Soc. Psychol. **87**(4), 510–528 (2004)
5. van Kleef, G.A., van Lange, P.A.M.: What other's disappointment may do to selfish people: emotion and social value orientation in a negotiation context. Pers. Soc. Psychol. Bull. **34**(8), 1084–1095 (2008)

6. Broekens, J., Jonker, C.M., Meyer, J.J.: Affective negotiation support systems. J. Ambient Intell. Smart Environ. **2**(2), 121–144 (2010)
7. Johnson, N.A., Cooper, R.B., Chin, W.W.: Anger and flaming in computer-mediated negotiation among strangers. Decis. Support Sys. **46**(3), 660–672 (2009)
8. Martinovski, B.: Emotion in negotiation. In: Kilgour, D.M., Eden, C. (eds.) Handbook of Group Decision and Negotiation, vol. 4, pp. 65–86. Springer, Netherland (2010)
9. Te'ini, D.: A cognitive-affective model of organizational communication for designing IT. Manag. Inf. Sys. Q. **25**(2), 251–312 (2001)
10. Walther, J.B., D'Addario, K.P.: The impacts of emoticons on message interpretation in computer-mediated communication. Soc. Sci. Comput. Rev. **19**, 324–347 (2001)
11. Darics, E.: Non-verbal signalling in digital discourse: the case of letter repetition. Discourse Context Media **2**, 141–148 (2013)
12. Griessmair, M., Koeszegi, S.T.: Exploring the cognitive-emotional fugue in electronic negotiations. Group Decis. Negot. **18**(3), 213–234 (2009)
13. Vetschera, R., Filzmoser, M., Mitterhofer, R.: An analytical approach to offer generation in concession-based negotiation processes. Group Decis. Negot. **23**(1), 71–99 (2014)
14. Filzmoser, M.: Simulation of Automated Negotiation. Springer, Vienna (2010)
15. Filzmoser, M.: Automated vs human negotiation. Int. J. Artif. Intell. **4**(10), 64–77 (2010)
16. Ekman, P., Cordaro, D.: What is meant by calling emotions basic. Emot. Rev. **3**(4), 364–370 (2011)
17. Barrett, L.F.: Feelings or words? Understanding the content in self-report ratings of experienced emotion. J. Pers. Soc. Psychol. **87**(2), 266–281 (2004)
18. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**(6), 1161–1178 (1980)
19. Watson, D., Tellegen, A.: Toward a consensual structure of mood. Psychol. Bull. **98**(2), 219–235 (1985)
20. Hippmann, P.: Multi-level dynamics of affective behaviors in text-based online negotiations: impacts on negotiation success and impacts of decision support. Doctoral thesis, University of Vienna (2014)
21. Lazarus, R.S., Smith, C.A.: Knowledge and appraisal in the cognition—emotion relationship. Cogn. Emot. **2**(4), 281–300 (1988)
22. Ortony, A., Clore, G.L., Foss, M.A.: The referential structure of the affective lexicon. Cogn. Sci. **11**(3), 341–364 (1987)
23. Burgoon, J.K., Hale, J.I.: The fundamental topoi of relational communication. Commun. Monogr. **51**, 193–214 (1984)
24. Frijda, N.H.: Emotions, individual differences and time course: reflections. Cogn. Emot. **23** (7), 1444–1461 (2009)
25. Schoop, M., Jertila, A., List, T.: Negoisst: a negotiation support system for electronic business-to-business negotiations in e-commerce. Data Knowl. Eng. **47**(3), 371–401 (2003)
26. Heady, R.B., Lucas, J.L.: Permap: an interactive program for making perceptual maps. Behav. Res. Methods Instrum. Comput. **29**(3), 450–455 (1997)
27. Filzmoser, M., Hippmann, P., Vetschera, R.: Analyzing the multiple dimensions of negotiation processes. Group Decis. Negot. **25**(6), 1169–1188 (2016). doi:10.1007/s10726-016
28. Devitt, A., Ahmad, K.: Is there a language of sentiment? An analysis of lexical resources for sentiment analysis. Lang. Resour. Eval. **47**(2), 475–511 (2013)
29. Weiss, S.M., Indurkhya, N., Zhang, N.: Fundamentals of Predictive Text Mining. Springer, London (2010)