Lexical Affect Sensing: Word Spotting Revisited

Alexander Osherenko, osherenko@informatik.uni-augsburg.de, University Of Augsburg

Abstract: Recently, there has been considerable interest in the recognition of affect from texts. In this paper, we revisit the word spotting technique for affect sensing in short texts, for which purpose, we extract words from different affect dictionaries and explore the performance of various strategies for sensing affect.

1 Introduction

Recently, there has been considerable interest in affect recognition from texts. Psychological studies reveal that a computer system is more likely to be accepted by the human user if it is able to recognise his or her emotions and respond accordingly. Affect recognition is important in dialogue applications, such as human-robot communication or tutoring systems.

Fully automatic statistical approaches as in [5] work reliably on long texts (empirically more than 200 words) and are not specialised for short texts. In contrast, reliable affect categorisation in short texts may be achieved by semantic approaches that analyse the meaning of particular words.

Following [4], we define word spotting as an approach that classifies a given text's emotion into affect categories based on the presence of emotional words in the text. Although word spotting has its drawbacks, for example, the grammatical weakness, it remains popular because of its accessibility and economy. In this paper, we study the applicability of the word spotting technique for affect sensing by extracting emotional words in analysed texts and recognising affect on the basis of these words.

The classified texts can contain contradictory emotional words or negations, such as the positive word *happy* and the negative word *sad* in the phrase: *I am neither happy nor sad*. To recognise affect in such phrases, we apply three strategies: first emotional word strategy, last emotional word strategy, and average vote strategy. The 'first word strategy' classifies a text according to the emotional meaning of the first emotional word, whereas the 'last word strategy' classifies a text according to the emotional meaning of the last emotional word. The 'average vote strategy' counts the number of emotional words in a text: if positive words prevail, the text is classified as positive, while a greater number of negative words classifies the text as negative; otherwise, the text is classified as neutral. For instance, according to the first emotional word strategy the phrase *I am neither happy nor sad* is classified as *positive* due to the first emotional word *happy* (considered as a positive word), or as *negative* according to the last word strategy due to the last emotional word *sad* (considered as a negative word), or according to the average vote strategy as *neutral* (there is no dominance of positive over negative words or vice versa).

In order to disambiguate words and simplify the analysis, we lemmatise words and add their parts of speech (POS) tags. Thus, for the example above, we modify the example text to *I PRP be VBP neither RB happy JJ nor CC sad JJ*.

2 Corpus

The Fifty Word Fiction corpus (FWF) consists of 759 English sentences manually annotated in terms of their sentiment and affect [6]. For instance, the corpus contains a sentence *We all laughed and ordered beers* annotated as *positive*. FWF was collected online and annotated by 49 annotators. Of the sentences, 82 sentences are annotated as *positive*, 171 as *negative*, and 506 as *unclassifiable*. The Kappa coefficient, a measure for inter-coder agreement, yielded a value of 65%. Although this value is lower than the desirable inter-annotator agreement (80% following [1]), the FWF corpus was used due to the lack of emotional corpora, which contain short emotional texts.

3 Affect Dictionaries

For our study, we collected emotional pairs, an emotional word and its POS tag, from different dictionaries. As several dictionaries may contain the same words, we first considered Levin verbs, words from the General Inquirer (GI), and words from WordNet-Affect.

[8] defines 34 additional emotional verbs that were not found in the GI. Stone and colleagues [7] introduce the GI dictionary, containing lemmas of different parts of speech in category *Positive* (1,914 positive word senses) and in category *Negative* (2,293 negative word senses). [10] introduces the Whissell's Dictionary of Affect that contains 8,742 words of different POS and inflections that are characterised by their emotional connotation along three dimensions: evaluation, activation, and imagery. The score values were determined by human judgment. Valitutti and colleagues [9] composed the WordNet-Affect Database, which makes use of 1,903 affect words of different inflection stored in the WordNet dictionary [1].

We collected 4,527 emotional pairs: 34 Levin verbs, 1,790 positive and 2,200 negative pairs from GI, and 503 pairs from WordNet-Affect.

4 Results

To evaluate the proposed strategies, we implemented a computer system that counts occurrences of emotional pairs in a text passage and classifies the conveyed affect. The POS tagging and lemmatisation is yielded by the Stanford parser [3].

Baseline (%)	*Strategy*	*R (%)*	*P (%)*
37.20	First word	47.30	43.90
	Last word	46.70	43.57
	Average vote	44.36	42.18

Table 1. Results of the word spotting

Table 1 shows the results of word spotting. The *Baseline* column contains the highest recall value calculated by the statistical approach in [5]. The *Strategy* column shows the strategies of word spotting. The *R (%)* and *P (%)* columns present recall and precision values, respectively, yielded by the corresponding strategy. All *Baseline*, *Precision*, and *Recall* values are averaged over classes.

5 Discussion

Calculated results are higher than the baseline value (47.30% vs. 37.20%). The results for particular strategies are very similar, thus, it is impossible to recommend a particular strategy. Furthermore, although the recall value of the first-word vote strategy yields numerically the highest value (47.30%), the revealed trends must be also tested on other corpora.

6 References

1. Craggs, R. 2004. Annotating emotion in dialogue. 7th Annual CLUK Research Colloquium.
2. Fellbaum, C., editor. 1998. WordNet An Electronic Lexical Database. The MIT Press.
3. Klein, D., Manning, C. D. 2003. Accurate Unlexicalized Parsing. *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, pp. 423-430.
4. Liu H., Lieberman H., Selker T. 2003. A Model of Textual Affect Sensing using Real-World Knowledge. ACM Conference on Intelligent User Interfaces, Miami, pp. 125-132.
5. Osherenko, A. André, E. 2007. Lexical Affect Sensing: Are Affect Dictionaries Necessary to Analyze Affect? In Proceedings of Affective Computing and Intelligent Interaction (ACII), Springer.
6. Read, J. 2004. Recognising affect in text using pointwise-mutual information. University of Sussex.
7. Stone, P. J., Dunphy, D. C., Smith, M.S., Ogilvie, D. M. 1966. The General Inquirer: A Computer Approach to Content Analysis. MIT Press.
8. Levin, Beth. 1993. English verb classes and alternations. The University of Chicago Press.
9. Valitutti, A., Strapparava, C., Stock, O. 2004. Developing Affective Lexical Resources. PsychNology Journal. Volume 2, Number 1, 61-83
10. Whissell, C.M. 1989. The dictionary of affect in language. In Robert Plutchik and Henry Kellerman (Ed.), *Emotion: Theory, Research, and Experience* (pp. 113–131). New York: Academic Press.