Towards Semantic Affect Sensing in Sentences

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

Abstract. Recently, there has been considerable interest in the recognition of affect in written and spoken language. In this paper, we describe a semantic approach to lexical affect sensing in sentences that uses findings from linguistic literature and from empirical examples. The approach is evaluated using a corpus containing 759 English sentences.

1 INTRODUCTION

Lexical affect sensing is an important field of study whose results can be used in a wide range of applications, e.g. robotics or tutoring systems. Despite its illusory simplicity, the emotional analysis of texts presents a great challenge to computer scientists because of the manifoldness of expressed meanings in texts.

There are two types of approach aimed at solving this problem: statistical and semantic. Statistical approaches make use of data-mining methods, e.g. Support Vector Machines (SVM), and classify emotion in text, for instance by using word counts [8]. However, statistical approaches produce low classification results when classifying short texts.

In contrast, semantic approaches aim to classify affect in texts by using commonsense as well as linguistic information on emotional parts of analyzed texts. For instance, Prendinger and colleagues [9] classify the affective meaning of texts by using emotion words from [13] in word-level analysis, by using lexical modifiers of meaning and negations in phrase-level analysis, or by scrutinizing the grammatical structure of sentences in sentence-level analysis.

2 SYSTEM

We solve the introduced manifoldness problem by analyzing parts of studied texts: the whole text is split into sentences and the sentences into phrases. After the emotional meaning of each part is analyzed, the emotional meaning of the original text is deduced from the emotional meanings of the constituent phrases.

In order to test our idea, we implemented a computer system that uses two functionally complementary parsers: the SPIN parser and the Stanford parser.

The SPIN parser is a semantic parser for spoken dialogue systems, a rule-based framework that parses texts using order-independent word matching [2]. For instance, in text *I like this game* the SPIN parser finds the positive verb *like*. The probabilistic Stanford parser is used for determining parts of speech, lemmatizing words, and splitting text in parts [4]. For example, it takes the text *Finally, I was so angry that I could burst with rage* and splits it into a superordinate subsentence *I was so angry* and a subdominant sentence *that I could burst with rage*.

A text can contain several emotional phrases that have contradictory emotional meaning. We test three strategies for interpreting a text's emotional meaning: as defined by the first or by the last emotional part in the corresponding part (whole text, subsentences, phrases), or by the average meaning (the emotional meaning as defined by the majority of affective votes). For instance, in the sentence *I am happy and sad* the emotional word *happy* (considered a positive word) defines according to the strategy of the first phrase a positive meaning, the emotional word *sad* (considered a negative word) defines according to the strategy of the last phrase a negative meaning, and according to the strategy of an emotional average a neutral meaning (there is no emotional majority).

The affect recognition system classifies the emotional meaning into two stages: in the first stage (division) the system divides the text into parts of particular granularity (analyzes an unchanged text or splits it into subsentences or phrases) and scrutinizes the emotional meaning of each individual part, while in the second stage (consolidation) the system compiles the emotional meaning of the original text by composing it from the emotional meanings of the detected parts.

The proposed algorithm for semantic affect sensing (examined using the example of the emotional sentence *Finally, I was so angry that I could burst with rage*) is therefore as follows (depending on the chosen granularity of the analysis either *Whole text*, *Subsentences*, or *Phrases*):

a. *Whole text.* Apply the chosen classification strategy to the analyzed text (first phrase strategy – emotional meaning of word *angry*, last phrase strategy – emotional meaning of word *rage*, average strategy – average meaning of words, i.e. emotional meaning of words *angry* and *rage*).
b. *Subsentences*. Detect subsentences using the Stanford parser, classify their emotional meaning using the SPIN parser according to the chosen classification strategy (first phrase, last phrase, average), construct an auxiliary text (subsentence combination) out of the emotional meanings of subsentences, and classify the emotional meaning of an original sentence by analyzing the subsentence combination.

   For instance, the system detects the superdominant subsentence *Finally, I was so angry* and the subdominant subsentence *I could burst with rage* and constructs a subsentence combination *superord_high_neg subord_low_neg*, where *superord_high_neg* stands for the high negative meaning of the superordinate sentence and *subord_low_neg* for the low negative meaning of the subordinate sentence. The system classifies the original text as high negative (*high_neg*) by applying patterns for subsentences in Table 2.
c. *Phrases.* In contrast to step *b* above, run an additional intermediate step that facilitates the analysis of the emotional meanings of subsentences, not by using subsentences' text, but rather by using auxiliary texts – phrase combinations. Detect subsentences, then phrases that are contained in the detected subsentences, classify the emotional meaning of phrases according to the chosen

classification strategy (first phrase, last phrase, average), construct an auxiliary text for the emotional structure of the corresponding subsentence (phrase combination), classify the emotional meaning by applying patterns for phrases in Table 3, compile a subsentence combination, and calculate an emotional meaning of the original sentence by using patterns for subsentences in Table 2.

The system detects the dominant subsentence *Finally, I was so angry* and the subdominant subsentence *I could burst with rage*. In the dominant subsentence it extracts four phrases: adverb phrase (*finally*), noun phrase (*I*), verb phrase (*was*), and adjective phrase (*so angry*); and in the subdominant sentence three phrases: noun phrase (*I*), verb phrase (*could burst with*), and noun phrase (*rage*). The system constructs the phrase combination *phrase_null phrase_null phrase_null phrase_high_neg* for the dominant sentence (the phrase *so angry* is classified as *high_neg*), where *phrase_null* corresponds to a neutral meaning and *phrase_high_neg* to the high negative meaning of a phrase, and for the subdominant sentence the phrase combination *phrase_null phrase_null phrase_low_neg*. The system classifies affect in phrase combinations by applying patterns for phrases in Table 3, constructs a subsentence combination *superord_high_neg subord_low_neg that* (cf. step *b*) and classifies it as *high_neg* by applying patterns for subsentences in Table 2.

d. If necessary, calculate the majority vote on the basis of values yielded by the granularities above.

The system calculates the majority vote on the basis of values yielded by the granularities above. It takes from the *Whole text* granularity the *low_neg* value, from the *Subsentences* granularity the *high_neg* value, and from the *Phrases* granularity the *high_neg* value, and calculates the majority vote *high_neg*.

3 CORPUS

We chose in our experiments the Fifty Word Fiction corpus (FWF) containing 759 grammatically correct English sentences that are manually annotated in terms of their sentiment and affect as *positive*, *neutral*, or *negative* [11]. For instance, *We all laughed and ordered beers* is annotated as *positive*. The corpus was collected online and available to the general public for one month, during which some 3,301 annotations were made by 49 annotators. Of the sentences, 82 are annotated as *positive*, 171 as *negative*, and 506 as *unclassifiable*. The inter-coder agreement is 65% (less than 80% – a desirable agreement in line with [1]).

4 SOURCES OF AFFECT INFORMATION

Affect Sensing of Parts

Our system utilizes the following information to classify the affect of parts: information from affect dictionaries, grammatical lexical patterns from linguistic studies, and empirical lexical patterns from our own studies.

Information from Affect Dictionaries

We use emotion words from various affect dictionaries as the basis for our system: Levin verbs [6], GI [12], and WordNet-Affect [13]. We consider 4,527 words from affect dictionaries in our study: 503 words from WordNet-Affect, GI words (1,790 positive and 2,200 negative), and 34 Levin verbs.

Grammatical lexical patterns from linguistic studies

In our system, we use 11 grammatical patterns to scrutinize the emotional meaning of texts [5]:

1. Interjections (299), e.g. *Oh, what a beautiful present!*
2. Exclamations (300a), e.g. *What a wonderful time we've had!*
3. Emphatic *so* and *such* (300b), e.g. *I'm so afraid they'll get lost!*
4. Repetition (300c), e.g. *This house is 'far, 'far too expensive!*
5. Intensifying adverbs and modifiers (301), e.g. *We are utterly powerless.*
6. Emphasis (302), e.g. *How ever did they escape?*
7. Intensifying a negative sentence (303a), e.g. *She didn't speak to us at all.*
8. A negative noun phrase beginning with *not a* (303b), e.g. *We arrived not a moment too soon.*
9. Fronted negation (303c).
10. Exclamatory questions (304), e.g. *Hasn't she grown!*
11. Rhetorical questions (305), e.g. *What difference does it make?*

Empirical lexical patterns from our own studies

We used 25 empirical examples of emotional texts containing negations and intensifiers to build lexical patterns for analyzing emotional meanings. The patterns classify the emotional meanings of texts, facilitating a five-class scheme (*low positive*, *high positive*, *low negative*, *high negative*, *neutral*) using emotion words, negations (*not*, *never*, *any*), and 74 intensifiers of emotional meaning, e.g. *definitely* (Table 1).

Example	*Pattern*
I am so happy.	<*Intensifier*> <*Emotional word+*>→ <*Result++*>
I am not happy.	<*Negation*> <*Emotional word+*> → <*Result−*>
I am not very happy.	<*Negation*> <*Intensifier*> <*Emotional word+*>→ <*Result−*>

Table 1. Example patterns for modifying affect

Table 1 shows example patterns for modifying affect. The *Pattern* column shows a pattern that matches the example text in the *Example* column. <*Intensifier*> denotes an intensifier word, <*Emotional word+*> a low positive emotional word, <*Result++*> the high positive result of affect sensing, and <*Result−*> the low negative result of affect sensing.

Patterns for Linking Parts

Phrases and subsentences divide the original sentence into parts, with each potentially having its own emotional meaning. For the purpose of compiling the meaning of the original text from constituent parts, the implemented system composes the emotional meaning of the original text out of the emotional meanings of constituent phrases and subsentences.

The proposed system contains 122 empirical patterns for linking subsentences and 19 empirical patterns for linking phrases.

Pattern for linking subsentences	*Example*
<Sup++> <Sup+>→ <Result++>	*It is a very good film and the acting is excellent.*
<Sup++> <Sub−>→ <Result+>	*It is a very good film although the acting seems at first to be not excellent.*

Table 2. Example patterns for linking subsentences

Table 2 shows sample patterns for linking subsentences. The *Pattern for linking subsentences* column shows a pattern that matches the text in the *Example* column. <*Sup*++> represents the high positive emotional meaning of the superdominant subsentence, <*Sup*+> the low positive meaning of the superdominant sentence, <*Sub*−> the low negative emotional meaning of the subdominant subsentence, <*Result*++> the high positive result of affect sensing, <*Result*+> the low positive result of affect sensing, and <*Result*−> the low negative result of affect sensing.

Table 3 shows sample patterns for linking phrases.

Example pattern for linking phrases	*Example*
<Phrase+> <Phrase0> → <Result+>	*exact and accurate*
<Phrase+><Phrase−> → <Result−>	*happy and depressing*

Table 3. Example patterns for linking phrases

Table 3 shows sample patterns for linking phrases. The *Pattern for linking phrases* column shows a pattern that matches the text in the *Example* column. <*Phrase*+> represents the positive emotional meaning of the phrase and <Phrase−> the low negative emotional meaning of the phrase.

5 RESULTS

The baseline for evaluating the proposed approach provides the best recall value, 37.20% averaged over classes, calculated via the statistical approach in [8] using word counts as features and a SVM classifier.

Table 4 shows the results for solving a three-class problem using the proposed approach with and without lexical patterns (using only emotional words). The R^a column represents the recall value averaged over classes and the P^a column the corresponding precision value averaged over classes. The $R^{a\text{-}lp}$ column represents the recall value averaged over classes when classifying texts without lexical patterns and the $P^{a\text{-}lp}$ column signifies the corresponding precision value averaged over classes. The *Gran.* column represents the granularity of the text division (the decision based on the majority vote – no division; the text as a whole; division into subsentences – abbreviated as *Subsent.*; division into phrases), and the *Strategy* column shows the strategy of semantic sensing (first phrase, last phrase, average vote).

Gran.	*Strategy*	R^a	$R^{a\text{-}lp}$	P^a	$P^{a\text{-}lp}$
Majority	*First phrase*	***47.20***	*45.02*	*44.09*	*42.76*
	Last phrase	***47.64***	*46.24*	*44.26*	*43.45*
	Average vote	*45.92*	*45.66*	*43.14*	*43.05*
Whole Text	*First phrase*	*45.41*	***47.30***	*42.90*	*43.90*
	Last phrase	*47.45*	*46.70*	*44.05*	*43.57*
	Average vote	*42.79*	*44.36*	*41.15*	*42.18*
Subsent.	*First phrase*	*47.20*	*45.22*	*44.08*	*42.88*
	Last phrase	*47.24*	*45.84*	*44.03*	*43.22*
	Average vote	*46.04*	*45.66*	*43.22*	*43.05*
Phrase	*First phrase*	*44.79*	*43.71*	*42.90*	*42.13*
	Last phrase	*45.21*	*44.54*	*43.13*	*42.65*
	Average vote	*44.22*	*44.16*	*42.41*	*42.40*

Table 4. Results of affect sensing for three classes

The results corresponding to the word spotting (see the definition in [7]) are shown in the rows of *Whole text* (hereafter referred to as the word-spotting values). Other alternatives, e.g. in the rows of *Phrase*, cannot be considered as word-spotting-processing, since additional patterns for processing combinations (phrase and subsentence combination) are necessary.

The *Majority* rows show the majority vote of the most entities (phrases, subsentences, utterance). If the majority vote cannot be calculated, i.e. classification results are pairwise different, the result of the *Subsentences* classification is taken as the basis.

6 DISCUSSION & FUTURE WORK

The proposed semantic approach is tested on a corpus with English sentences, and the applied patterns improve classification rates compared both with the word-spotting values and with the statistical baseline of 37.20% (Table 4). For instance, the *Majority*, *Last phrase* classification rate, 47.64%, is much higher than the statistical baseline 37.20% and also higher than the word-spotting value of 47.30% for *Whole text*, *First phrase*.

Moreover, the classification rates are higher for the majority evaluation using the full grammar set of applied patterns (47.64% for *Majority*, *Last phrase*). Furthermore, the results are significantly higher compared with the statistical baseline (47.64% vs. 37.20%). In addition, the average vote does not generally bring an enhancement of classification results, e.g. 47.64% for *Majority*, *Last phrase* vs. 45.92% for *Majority*, *Average vote*.

In future, we will revise our approach and collect new corpora containing short emotional texts, for instance through acquiring new data from the Internet.

REFERENCES

1. Craggs, R. 2004. Annotating emotion in dialogue – issues and approaches. 7th Annual CLUK Research Colloquium.
2. Engel, R. 2006. SPIN: A Semantic Parser for Spoken Dialog Systems. In *Proceedings of the Fifth Slovenian And First International Language Technology Conference (IS-LTC 2006)*, 2006.
3. Kollias, S. 2007. ERMIS Project. URL: http://www.image.ntua.gr/ermis/.

4. Klein, D., Manning, C. D. 2003. Accurate Unlexicalized Parsing. Proceedings of the 41st Meeting of the Association for Computational Linguistics, pp. 423-430.
5. Leech, G. N., Startvik, J. 2003. A communicative grammar of English. Third edition. Longman.
6. Levin, Beth. 1993. English verb classes and alternations. The University of Chicago Press.
7. Liu, H., Lieberman, H., Selker, T. 2003. A Model of Textual Affect Sensing Using Real-World Knowledge. Pages 125-132 of: Proceedings of IUI-03, the 8th international conference on intelligent user interfaces. Miami, US: ACM Press.
8. 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.
9. Neviarouskaya, A., Prendinger, Ishizuka, M. Textual Affect Sensing for Social and Expressive Online Communication, Affective Computing and Intelligent Interaction (A. Pavia, P. Prada and R. W. Picard (Eds.)), (Proc. Int'l Conf. ACII2007, Lisbon, Portugal), Springer LNCS 4738, pp.218-229 (2007.9).
10. Quirk, R., Greenbaum, S. 1988. A University Grammar of English. Longman Publishing Group; Abridged Ed (January 1, 1973).
11. Read, J. 2004. Recognising affect in text using pointwise-mutual information. Masters thesis, University of Sussex.
12. Stone, P. J., Dunphy, D. C., Smith, M.S., Ogilvie, D. M. 1966. The General Inquirer: A Computer Approach to Content Analysis. MIT Press.
13. Valitutti, A., Strapparava, C., Stock, O. 2004. Developing Affective Lexical Resources. PsychNology Journal. Volume 2, Number 1, 61-83.