

Ontology-Driven Affective Chinese Text Analysis and Evaluation Method

Linhong Xu and Hongfei Lin

Department of Computer Science and Engineering,
Dalian University of Technology, Dalian, China 116024
qingniao1203@163.com, hflin@dlut.edu.cn
<http://ir.dlut.edu.cn>

Abstract. This paper studies text affective analysis from three different levels: words, sentences and discourses. Firstly, affective lexicon ontology is constructed by employing the manual and automatic classification methods referred to the present emotional classification. Then affective analysis of sentences and discourses is completed by emotional information of affective lexicon ontology. For sentential recognition, the lexical emotion information and semantic features are appended into Condition Random Fields, and the emotional chain of a text document is generated. Finally affective structure of a text document is evaluated by two different methods as single sentence evaluation and joint sentence evaluation. In analyzing the discourses, affective tone is acquired through Emotional Largest Chain method and Support Vector Machine method.

1 Introduction

In the year 2000, Pero Subasic presents that fuzzy semantic is employed to recognize affective text analysis, and exploit rather than reducing the word's ambiguity [1]. Three years later, Hugo Liu in MIT discusses visualizing the affective structure of a text document [2]. In 2005, Chunling Ma in the University of Tokyo employs the syntactical sentence-level processing and keyword spotting technique to analyze emotions [3], and, Fuji Ren in the University of Tokushima recognizes the emotion from text based on the constructed emotion thesaurus [4].

2 The Model of Affective Text Analysis

This paper studies text affective analysis from three different levels: words, sentences and discourses. Firstly, basic emotional information of words is provided by affective lexicon ontology which includes 10,200 entries now.

Secondly CRF is used to analyze sentences' emotion in literatures, and semantic features of emotional information of the sentence are chosen as follows: affective word feature; emotional class feature; words with negative meanings; the adversative conjunction; modal feature of sentence; whether it is the first sentence of the article; the length of the sentence; the type of sentence; lexical repetitive degree in a sentence.

Finally, this paper adopts two methods to acquire affective tone of the discourse. One is Emotional Largest Chain method. It chooses the one emotion which sustains the longest time in the emotional chain of the recognition of sentences, namely that one emotion occurs continuously in chain and has the longest sub chain. Another method firstly obtains the $tf*idf$ values of every affective words of the article accord to affective lexicon ontology, and then uses SVM to classify, finally gets affective tone of the discourse by the result of SVM.

Two evaluation methods of affective analysis are used. The first evaluation method (single sentence evaluation): the number of the correct tagged sentences divides the number of all sentences in an article; The second evaluation method (joint sentence evaluation): the emotional consistency is evaluated by the accuracy of the multivariable emotion co-occurrence.

3 Conclusion and Further Work

This paper studies text affective analysis from three different levels of words, sentences and discourses. Affective recognition of sentences is evaluated one by one, and then the emotion recognition of the discourse is completed. The result with CRFs is better than the result with tagging the emotion directly on the accuracy of micro average and the consistency of emotion. In the analysis of affective tone of the discourse, the Emotional Largest Chain method is better than Support Vector Machine on both total accuracy and respective accuracy of every emotion class, especially when the distribution of the corpus is not well-distributed.

Acknowledgments. This work is supported by grant from the Natural Science Foundation of China (No.60373095 and 60673039) and the National High Tech Research and Development Plan of China (2006AA01Z151).

References

1. Subasic, P., Huettner, A.: Affect Analysis of Text Using Fuzzy Semantic Typing, *Fuzzy Systems*, pp. 483–496 (2000)
2. Liu, H., Selker, T., Lieberman, H.: Visualizing the affective structure of a Text Document. *Conference on Human Factors in Computing Systems*, 740–741 (2003)
3. Ma, C., Helmut, P., Mitsuru, I.: Emotion estimation and reasoning based on affective textual interaction. *Affective computing and intelligent interaction*, 622–628 (2005)
4. Zhang, Y., Li, Z., Ren, F., Kuroiwa, S.: Semi-automatic Emotion Recognition from Textual Input Based on the Constructed Emotion Thesaurus. *Natural Language Processing and Knowledge Engineering*, 571–576 (2005)