Using Morphological and Syntactic Structures
for Chinese Opinion Analysis

Lun-Wei Ku   **Ting-Hao Huang**   **Hsin-Hsi Chen**

Department of Computer Science and Information Engineering
National Taiwan University
No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan
{lwku,thhuang}@nlg.csie.ntu.edu.tw;hhchen@ntu.edu.tw

Abstract

This paper employs morphological structures and relations between sentence segments for opinion analysis on words and sentences. Chinese words are classified into eight morphological types by two proposed classifiers, CRF classifier and SVM classifier. Experiments show that the injection of morphological information improves the performance of the word polarity detection. To utilize syntactic structures, we annotate structural trios to represent relations between sentence segments. Experiments show that considering structural trios is useful for sentence opinion analysis. The best f-score achieves 0.77 for opinion word extraction, 0.62 for opinion word polarity detection, 0.80 for opinion sentence extraction, and 0.54 for opinion sentence polarity detection.

1 Introduction

Sentiment analysis has attracted much attention in recent years because a large scale of subjective information is disseminated through various platforms on the web. Sentiment information can be applied to a wide variety of fields, including product recommendation, review summarization, public polling, and so on.

Opinion dictionaries are important resources for identifying subjective information. Several approaches were proposed to collect such resources. Wiebe (2000) learned subjective adjectives from corpora. Takamura *et al.* (2005) extracted semantic orientations of words. Ku *et al.* (2007) measured sentiment degrees of Chinese words by averaging the sentiment scores of the composing characters. When the opinion words are available, the polarities of sentences and documents can be determined by them. Riloff and Wiebe (2003) learned the extraction patterns for subjective expressions. Kim and Hovy (2004) found the polarity of subjective expressions. Pang *et al.* (2002) and Dave *et al.* (2003) explored various techniques at document level.

Morphological information has been widely used in classifying words, telling the meanings, and doing other in-depth analysis (Tzeng and Chen, 2002). However, morphological information was seldom applied either in Chinese opinion extraction, or in solving the coverage problem of opinion dictionary. Instead of bag-of-characters approach (Ku *et al.*, 2007), this paper employs morphological structures of words to extract opinion words.

Relations between sentence segments are also defined by linguistics in the Chinese language. These are similar to morphological structures between Chinese characters. Based on parsing trees of sentences, we identify these relations and utilize them for opinion analysis on sentences.

As the experimental corpus, some researchers managed to generate annotated materials and gold standards under many constraints. Ku set a standard for generating final answers from annotations of multiple annotators (Ku *et al.*, 2007), and Somasundaran annotated discourse information from meeting dialogs to train a sentiment model (Somasundaran *et al.*, 2007). For multilingual issues, researchers concerned mainly about the applicability of corpus and algorithms from the native language to foreign languages (Banea *et al.*, 2008; Bautin *et al.*, 2008).

Several opinion analysis systems have been developed so far. OASYS (Cesarano *et al.*, 2007) and CopeOpi (Ku *et al.*, 2007) allow users input their queries and select preferred data sources,

and then track opinions in a time zone. For both systems, extracting opinions is the main focus, while holders and targets are identified implicitly when retrieving relevant documents. Carenini's team proposed a graphical user interface for evaluative texts (2006), in which color blocks were used to present the evaluations for components of products. Fair News Reader, a Japanese news Web system, incorporates sentiment information insensibly in an interesting way (Kawai *et al.*, 2007). It provides readers "balanced" reports by analyzing the sentiment in news articles which readers have read, and suggests them new articles according to the analysis results. It leads the application of opinion analysis to the direction of personalization.

2 Chinese Morphological Structures

In the Chinese language, a word is composed of one or more Chinese characters, and its meaning can be interpreted in terms of its composite characters. The morphological structures of Chinese words are formulated by three major processes in linguistics: *compounding*, *affixation*, and *conversion*. *Compounding* is a complex word-formation process. In most cases, two or more morphemes together are formed as a lexical item by this process. *Affixation* is a morphological process, by which grammatical or lexical information is added to a base form. By the *conversion* process, a word is changed from one part of speech into another without the addition or deletion of any morphemes.

Compounding is the most productive way to construct a Chinese word. Mostly, a Chinese character itself carries meanings, so that a morpheme can function as a character and has its own part of speech. In some cases, a Chinese morpheme may carry no specific meaning and just makes a word more readable. Cheng and Tian (1992) divided Chinese words into five morphological types based on the relations between the morphemes in *compounding* words.

(1) Parallel Type: Two morphemes play coordinate roles in a word. For example, the morphemes "財" (money) and "富" (wealth) are parallel in the word "財富" (money-wealth).

(2) Substantive-Modifier Type: A modified morpheme follows a modifying morpheme. For example, the morpheme "哭" (cry) is modified by "痛" (bitterly) in the word "痛哭" (bitterly-cry).

(3) Subjective-Predicate Type: One morpheme is an expresser and the other one is described. The structure is like a subject-verb sentence condensed in one word. For example, the morpheme "心" (heart) is a subject of the predicate "疼" (hurt) in the word "心疼" (heart-hurt).

(4) Verb-Object Type: The first morpheme is usually a verb which governs the second one, making this word similar to a verb followed by its object. For example, the morpheme "控" (control) serves as the object of the verb "失" (lose) in the word "失控" (lose-control).

(5) Verb-Complement Type: The first morpheme is usually a verb but sometimes can be an adjective, and the second morpheme explains the first one from different aspects. For example, the morpheme "清" (clearly) expresses the aspects of the action "看" (look).

Chinese words constructed by *affixation* process can be one of the two cases – say, *morpheme* and *morpheme*, or *morpheme* and *non-morpheme*. In the case of *morpheme* and *morpheme*, the affixation word belongs to one of the above 5 types if the prefix and the suffix are neither negations nor confirmations. Types 6 and 7 defined below represent the affixation words whose prefix or suffix is a negation or a confirmation. The affixation words whose prefix or suffix characters are not morphemes are classified into type 8.

(6) Negation Type: There is at least one negation character in words of this type. For example, the prefix "無" (no) is the negation morpheme in the word "無法" (no-method).

(7) Confirmation Type: There is at least one confirmation character in words of this type. For example, the prefix "有" (do) is a confirmation in the word "有賴" (do-depend on).

(8) Others: Those words that do not belong to the above seven types are assigned to this type, such as words whose meanings are not a function of their composite characters, words whose composite characters are not morphemes, such as "姪子" (nephew-suffix) and "薄荷" (peppermint).

3 Opinion Scores of Chinese Words

The bag-of-characters approach proposed by Ku *et al.* (2007) considers the observation probabilities of characters in Chinese opinion words. It calculates the observation probabilities of characters from a set of seeds first, then dynamically enlarges the set and adjusts their probabilities. In

this approach, the opinion score of a word is determined by the combination of the observation probabilities of its composite characters defined by Formulas (1) and (2).

$$P(C, pos) = \frac{f(C, pos)/\sum_{i=1}^{n} f(C_i, pos)}{f(C, pos)/\sum_{i=1}^{n} f(C_i, pos) + f(C, neg)/\sum_{i=1}^{m} f(C_i, neg)} \quad (1)$$

$$P(C, neg) = \frac{f(C, neg)/\sum_{i=1}^{m} f(C_i, neg)}{f(C, pos)/\sum_{i=1}^{n} f(C_i, pos) + f(C, neg)/\sum_{i=1}^{m} f(C_i, neg)} \quad (2)$$

$$S(C) = P(C, pos) - N(C, neg) \quad (3)$$

$$S(C_1 C_2 ... C_l) = \frac{1}{l}\sum_{i=1}^{l} S(C_i) \quad (4)$$

where C is an arbitrary Chinese character, $f(C, polarity)$ counts the observed frequency of C in a set of Chinese words whose opinion polarity is positive (*pos*) or negative (*neg*); $P(C, pos)$ and $P(C, neg)$ denote the observation probabilities of C as a positive and a negative character, and n and m denote total number of unique characters in positive and negative words. The difference of $P(C, pos)$ and $P(C, neg)$ in Formula (3) determines the sentiment score of character C, denoted by $S(C)$. Formula (4) computes the opinion score of a word of l characters $C_1 C_2 … C_l$ by averaging their scores.

Instead of counting the weights as in the bag-of-characters approaches, we consider the word structures and propose a scoring function for each morphological type. According to the Frequency Dictionary of Modern Chinese, 96.5% of Chinese words are unigrams and bigrams (Chen, *et al.*, 1997). In the following functions, $S(C_1 C_2)$ computes the opinion scores of words with characters C_1 and C_2. $SIGN(s)$ returns -1 if polarity degree s is smaller than 0, i.e., negative, and returns 1 when positive.

(1) Parallel Type: Since the two composite characters of a word of this type are homogeneous, the opinion score is the average score of two characters' opinion scores.

$$S(C_1 C_2) = \frac{S(C_1) + S(C_2)}{2} \quad (5)$$

(2) Substantive-Modifier Type: The first morpheme of a word of this type modifies the second one, so that its opinion weight comes from the absolute opinion score of the first character, while the opinion polarity is determined by the occurrence of negative opinion characters. If at least one negative opinion character appears, the

word is negative, else it is positive. For example, the word "痛哭" (bitterly cry) is composed of "痛" (bitterly, negative) and "哭" (cry, negative). Negative characters make this word negative and its opinion strength, i.e., the absolute value of the score, is decided by the first character for the degree of crying.

if $(S(C_1) \neq 0$ and $S(C_2) \neq 0)$ then
    if $(S(C_1) > 0$ and $S(C_2) > 0)$ then $S(C_1 C_2) = S(C_1)$
    else $S(C_1 C_2) = -1 \times |S(C_1)|$ $\quad (6)$
else $S(C_1 C_2) = S(C_1) + S(C_2)$

(3) Subjective-Predicate Type: The first morpheme of a word of this type is a subject and the second morpheme is the action it performs, so that the action decides the opinion score of the word. If the action is not an opinion or it is neutral, the subject determines the opinion score of this word. For example, the word "山崩" (mudslide, negative) is composed of "山" (mountain, non-opinion) and "崩" (collapse, negative). Its opinion score depends only on the second character "崩" (collapse) since the first character is a subject and usually bears no opinions.

if $(S(C_2) \neq 0)$ then $S(C_1 C_2) = S(C_2)$
else $S(C_1 C_2) = S(C_1)$ $\quad (7)$

(4) Verb-Object Type: The first morpheme of words of this type acts upon the second morpheme. The effect depends not only on the action but on the target. The weight is determined by the action, but the polarity is the multiplication of the signs of the two morphemes. For example, the word "避暑" (to go away for the summer, positive) is composed of "避" (hide, negative) and "暑" (hot summer, negative). Its strength depends on the strength of "避" (hide) and polarity is positive from the multiplication of two negatives.

if $(S(C_1) \neq 0$ and $S(C_2) \neq 0)$
    then $S(C_1 C_2) = |S(C_1)| \times SIGN(S(C_1)) \times SIGN(S(C_2))$ $\quad (8)$
    else $S(C_1 C_2) = S(C_1) + S(C_2)$

(5) Verb-Complement Type: The scoring function for words of this type is defined the same as that of a Subjective-Predicate type in Formula (7). The complement morpheme is the deciding factor of the opinion score. For example, the word "提高" (raise, positive) is composed of "提" (carry or lift, non-opinion) and "高" (high,

positive). The complement morpheme "高" (high) describes the resulting state of the verb morpheme "提" (raise), so both strength and polarity depend on the morpheme "高" (high).

(6) Negation Type: A negative character specified in a predefined set *NC* has a negation effect on the opinion score of the other character. The strength depends on the modified morpheme while the polarity of the word is the negation of the polarity of the modified morpheme.

$$\text{if } (C_1 \in NC) \text{ then } S(C_1C_2) = (-1) \times S(C_2)$$
$$\text{else } S(C_1C_2) = (-1) \times S(C_1) \qquad (9)$$

(7) Confirmation Type: A positive character specified in a predefined set *PC* ensures that the opinion score of a word only comes from the other character. Therefore, the opinion score of this word is determined by the modified morpheme.

$$\text{if } (C_1 \in PC) \text{ then } S(C_1C_2) = S(C_2) \text{ else } S(C_1C_2) = S(C_1) \qquad (10)$$

(8) Others: Since words of this type contain no clear cues for their morphological structures, we postulate that both characters have the same contribution, and adopt Formula (5).

4 Identification of Morphological Types

To compute the opinion score of a word according to formulae in Section 3, we must know its morphological type from the morphological structure, i.e., the parts of speech of the composite morphemes. Currently, part of speech tagging is performed at the word level rather than the morpheme level, and morpheme-tagging corpus is not available. We consider an on-line Chinese dictionary, Dictionary of Chinese Words by Ministry of Education, Taiwan (*MOEDCW*), as a corpus, and compute the statistics of each morpheme in it.

Two classifiers, *CRF* classifier and *SVM* classifier are proposed to recognize morphological types (1)-(5). Morphological types (6) to (8) are determined by rules such as whether two composite characters are morphemes; whether there are confirmation/negation morphemes; and so on.

4.1 MOEDCW Corpus

MOEDCW corpus provides possible parts of speech for each morpheme by treating it as a unigram word, and possible senses under each part of speech. In each entry, there are a sense defini-

tion and some example words. Figures 1 and 2 show the specifications of two morphemes "冒" and "汗". The morpheme "冒" has three parts of speech (verb, adverb and noun) and includes 3, 1, and 1 senses. There are 3, 3, and 2 example words listed under the three verb senses.

We can find the correct parts of speech of the composite characters of a word when it is an example word in the dictionary. However, not all words are listed in the corpus. Consider the word "冒汗" (sweat, verb). Figure 1 shows that "冒汗" (sweat) is an example word listed under the verb sense of the character "冒" (perspire), thus the character "冒" (perspire) in the word "冒汗" (sweat) functions as a verb. However, "冒汗" (sweat) is not an example for the character "汗" (sweat). Figure 2 show that there are two possible parts of speech, noun and verb, for the character "汗" (sweat). We then show how to identify its function in the word "冒汗".

verb	1	Goes out from the button to the top or from inside to outside. For example, **fume**, **smoking**, and **sweat**. 由下往上或往外透出、發散。如：「冒煙」、「冒氣」、「冒汗」。
	2	Burst into or regardless of. For example, **take risk**, **to offend**, and **offense**. 衝犯、不顧。如：「冒險」、「冒犯」、「衝冒」。…
	3	Fake or on the pretext of. For example, **personate** and **to pretend to be**. 假稱、假託。如：「冒名」、「假冒」。
ad-verb	1	Crude or rash. For example, **offensively** and **advance rashly**. 鹵莽、莽撞。如：「冒犯」、「冒進」。
noun	1	Family name. 姓。

Figure 1: Specification of "冒" in MOEDCW

noun	1	Sweat. For example, **cold sweat**, **night sweat**, **sweatiness**, and **to drip with sweat**. 由動物皮膚的毛細孔所排泄出的液體。如：「冷汗」、「盜汗」、「汗流浹背」、「揮汗如雨」。…
	2	Family name 姓。
verb	1	To sweat 流汗、使出汗。

Figure 2: Specification of "汗" in MOEDCW

$$T(C, POS) = NumberOfSenses(C, POS) \qquad (11)$$

The number of possible meanings one character can bear when it functions as a certain part of

speech is employed to estimate how often this part of speech is used. The function $T(C, POS)$ shown in Formula (11) defines the score of a character C functioning as a particular part of speech POS. Here, POS may be noun (N), adjective (ADJ), verb (V), adverb (ADV), auxiliary (AUX), conjunction (CONJ), pronoun (PRON), preposition (PREP), and interjection (INT). In Figure 2, $T(汗<sweat>, N) = 2$ and $T(汗<sweat>, V) = 1$.

4.2 Features for Classifiers

Features for training SVM and CRF classifiers include the pronunciation and the tone of the word, parts of speech of the first and the second characters of training words, and the position information of the composite characters. The tone of the word is acquired from MOEDCW. The parts of speech are estimated by Formula (11). $f(C, POS, k, start/end)$ counts the number of k-grams (k=2, 3, 4). In Figures 1 and 2, $f(冒, V, 2, start)$=6, $f(冒, V, 2, end)$= 2, $f(冒, ADV, 2, start) = 2$, and $f(冒, ADV, 2, end)$=0. This example shows that when the character "冒" functions as a verb or an adverb, it serves as the starting character more often than the ending character.

4.3 CRF and SVM Classifier

CRF and SVM are both common used algorithms for building classifiers (Lafferty *et al.*, 2001). We adopted $CRF++$[1] and $libSVM$ (Chang and Lin, 2001) to develop our classifiers. The features for training our CRF and SVM classifiers include the input word W, the tone of W, the first and the second characters C_1 and C_2, $T(C_1, POS)$, $T(C_2, POS)$, $f(C_1, POS, k, start)$, $f(C_1, POS, k, end)$, $f(C_2, POS, k, start)$, and $f(C_2, POS, k, end)$. POS denotes one of nine parts of speech in $MOEDCW$, and k equals to 2, 3 or 4.

Using SVM is straightforward. To classify a word into one of the morphological structure types, we construct the word's feature vector and input the vector into SVM. When using CRF, a different approach is taken. When predicting the classes of two successive instances, CRF takes the predicted class of the first instance into account when predicting the second instance's class. Here is how we exploit this capability. In a nutshell, we perform classification at the character level instead of the word level. Let W be a word composed of the two characters C_1 and C_2. Let v

[1] http://crfpp.sourceforge.net/

be the feature vector of W. Let t be the morphological structure type of W. We define C_1's feature vector to be composed of the features in v which are related to C_1, e.g., $T(C_1, verb)$. Similarly, C_2's feature vector is composed of the features in v which are related to C_2. C_1's class and C_2's class are defined as t_1 and t_2, respectively. Since t has five possible values, there are 10 character classes.

To determine a word W's morphological structure type, we first apply CRF on W's constituent characters C_1 and C_2's feature vectors. For C_1, CRF will return a set of probabilities $P(C_1,t\_q)$, where $q \in \{1, 2\}$, indicating the likelihood of C_1 being an instance of class t_q. Similarly, a set of probabilities $P(C_2,t\_q)$ is returned for C_2. W's morphological structure type is defined as the value of t which maximizes the product of $P(C_1,t\_1)$ and $P(C_2,t\_2)$.

Though CRF is mostly used for sequential labeling, the idea of using CRF is to tail this classification questions into a labeling question in order to utilizing the position information of characters. As mentioned, if a word W of two characters $C1C2$ is of type 1, CRF will label $C1$ 1_1 (type1_1st char) and $C2$ 1_2 (type1_2nd char). The labeling of each character considers both the previous character's features and the next character's features. That is, if the current character is the first character, its previous character is an empty character (which is used for segmenting sequences in CRF); if the current character is the second character, its next character is an empty character. Hence the position information will be considered by CRF.

5 Experiments and Discussion

Experiments verify whether the morphological types benefit opinion polarity detection on words. The relation between the performance of morphological classifiers and opinion polarity detection is discussed.

5.1 Experimental Setup

To compare the bag-of-characters approach (Ku *et al.*, 2007) with our morphological structure approach, we adopt the same evaluation data set containing 836 words. To evaluate the performance of our two morphological classifiers, we prepare two sets of words, including the testing set of 836 words for word-level opinion prediction (abbreviated as *OP*), and a set of 8,186 words selected from words in *MOEDCW* corpus and news documents except those can be classi-

fied by patterns (abbreviated as *TRAIN set*), all with their morphological types annotated. Table 1 lists the distributions of morphological types in *OP* and *TRAIN* sets.

The polarity of words is predicted by their opinion scores ranging between -1 to 1. We set a positive threshold. Those words with scores above it are considered as positive while those below this threshold multiplied by (-1) are regarded as negative. The words with non-zero scores falling between the positive and negative thresholds are neutral. Fifty grids from 0 to 0.5 are searched for the best threshold. Since the opinion extraction at word level concerns only word structure, no retraining for the best threshold is need when domain shifts, which is a superiority of our method.

5.2 Morphological Type Classification and Polarity Detection

The performances of *CRF* and *SVM* classifiers on each morphological type are listed in Table 2. We perform four-fold cross validation on the *TRAIN* set. Results show that *CRF* classifier achieves better performance than *SVM* classifier in this task. The accuracy of *CRF* classifier (0.70) is 8% higher than that of *SVM* classifier (0.62). Note those type 8 words which could be extracted by rules are excluded from classification experiment. The remaining type 8 words are usually proper names. It is difficult for both classifiers to identify such words.

Table 3 further shows the performance of polarity prediction using morphological types de-termined by *CRF* classifier and *SVM* classifier. The performance of polarity detection is evaluated by the f-score defined in Formula (12).

The f-scores of polarity detection using *CRF* classified types and *SVM* classified types are 0.5806 and 0.5938, respectively. Both of them outperform baseline's f-score 0.5455, i.e., the bag-of-characters approach (Ku *et al.*, 2007). Experiments show that adopting morphological types annotated by two classifiers for polarity prediction has little difference. In other words, *CRF* and *SVM classifiers* have an 8% f-score difference in their best performance of classification, while the performance gap in word polarity prediction using morphological types provided by these two classifiers is around 1.3% only (0.5806 vs. 0.5938). The reason may be that we define scoring functions of each morphological type in a straightforward way. If they are not the best scoring functions, the benefit of considering the morphological type information could be restricted. Nevertheless, experimental results show that morphological type information is useful for word polarity detection (with p-value less than 0.05).

$$P = \frac{correct(opinion) \cap correct(polarity)}{proposed(opinion)},$$

$$R = \frac{correct(opinion) \cap correct(polarity)}{gold(opinion)}, \quad (12)$$

$$f - score = \frac{2 \cdot P \cdot R}{P + R}.$$

set/type	1	2	3	4	5	6	7	8
TRAIN	26.15	44.97	1.64	15.14	9.22	0	0	2.88
OP	45.8	24.4	1.3	7.9	8.0	2.3	0.5	9.8

Table 1: The Percentage of distribution for morphological types in *TRAIN* and *OP* sets

MorphoType	1	2	3	4	5	8	Accuracy
CRF	0.63	0.78	0.41	0.66	0.78	0.17	0.70
SVM	0.49	0.73	0.22	0.52	0.55	0	0.62

Table 2: The f-score of *CRF* and *SVM* classifiers

We further examine how well our polarity detection method works in combination with a word sentiment dictionary. We use the NTUSD[2] word sentiment dictionary. If a word appears in NTUSD, then the word's polarity is the one specified in NTUSD. If a word does not appear in NTUSD, then the word's polarity is determined using our morphological type method.

After introducing a sentiment dictionary NTUSD[3], *CRF* and *SVM* classifiers both achieve the f-score 0.77 for opinion word extraction, and achieve f-scores 0.61 and 0.62 for polarity detection, respectively. Note that if only NTUSD is used to extract opinion words by string matching, the f-score is only 0.44.

---

[2] http://nlg18.csie.ntu.edu.tw:8080/opinion/

[3] http://nlg18.csie.ntu.edu.tw:8080/opinion/

Polarity f-score	Without NTUSD	With NTUSD
Ku	0.5455	0.5789
CRF type	0.5806	0.6100
SVM type	0.5938	0.6246

Table 3: Prediction with Morphological Types

We further analyze the improvement of polarity prediction for each morphological type. We find that the f-scores of polarity prediction of all morphological types are improved in different degrees, and among them the performance of type 2 words are improved the most. We have shown that our method can assign an opinion score to an arbitrary word without any word thesauri by considering its morphological information. Moreover, since the Substantive-Modifier (type 2) is the most common way to form a new word in the Chinese language (Cheng and Tian, 1992), the result presents the strength of our method in solving the coverage problem.

6 Syntactic Structure for Chinese Opinion Analysis

As mentioned, the relations introduced in Section 2 exist not only within words, but also between sentence segments. Relations between sentence segments are represented by structural trios hereafter and will be introduced in next section. We have already shown that morphological types are useful when extracting opinion words and would like to further testify whether structural trios also benefit the opinion analysis on sentences. We annotate these relations manually, propose a method to identify these relations, and compare results of experimental settings using structural trios with those not using structural trios.

6.1 Structural Trio

Each node in a parsing tree dominates a word string in a sentence. Linguistics have shown that there are also five relations between sentence segments: Parallel, Substantive-Modifier, Subjective-Predicate, Verb-Object, and Verb-Complement, same as morphological types (1) to (5). Because parsing trees have hierarchical structures, we define a structural trio to represent a relation between two nodes as follows:

(1) A structure trio contains two children nodes which bear a relation.
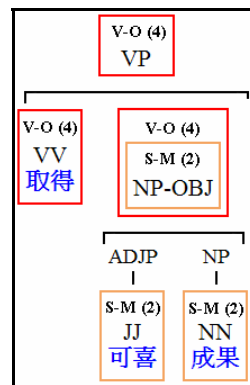(2) A structure trio contains one head node which is the nearest common parent of two children nodes in (1).

Figure 3: Example of structural trios

Figure 3 shows an example of a structure trio. It is a part of a parsing tree containing words "取得" (obtain), "可喜" (happy), "成果" (results). Two structural trios are shown in this example. The lower one contains two children nodes "可喜" (happy) and "成果" (results), and is labeled as Substantive-Modifier (S-M (2)) in their nearest common parent node, while the upper one contains two children nodes "取得" (obtain) and "可喜成果" (happy results), and is labeled as Verb-Object (V-O (4)).

6.2 Experimental Corpus

To experiment with structural trios, we need the parsing trees of all experimental sentences. For this purpose, we adopted Chinese Treebank 5.1[4] as the experimental materials. Chinese Treebank contains raw Chinese news documents together with their segmented, part of speech tagged, and parsed versions. The parsed documents are adopted in experiments utilizing structural trios, and the part of speech tagged documents are used in experiments not utilizing structural trios.

In Chinese Treebank, a unique ID is labeled on each sentence. For each sentence, we had three annotators label their opinions and then we generate the gold standard following NTCIR[5] MOAT protocol (Seki *et al.*, 2008). We also annotated structure trios in Chinese Treebank. A total of 17,159 sentences are obtained after dropping some faulty sentences such as empty sentences and sentences composed of more than one parsing tree. The statistics of opinion sentences and structural trios in the constructed experimental materials are shown in Table 4 and Table 5.

---

[4] http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T01
[5] http://research.nii.ac.jp/ntcir/index-en.html

	Opinion			Non-Opinion
	Positive	Neutral	Negative	
#	6,380	1,537	1,714	7,528
	9,631			
%	66.24	15.96	17.80	43.87
	56.13			

Table 4: Statistics of opinion sentences

Trio Type	Number	Percentage %
2	18,483	36.85
3	13,687	27.29
4	15,970	31.84
5	965	1.92
Others	1,054	2.10
Total	50,159	100.00

Table 5: Statistics of structural trios

6.3 Experiment Setup

The aim of our experiments is to know how opinion analysis approach performs when morphological and syntactic structures are incorporated. They are compared with the bag-of-character and bag-of-word approaches. We implemented the bag-of-word approach proposed by Ku *et al.* (2007) to show its performance on Chinese Treebank. In their approach, the opinion scores of words are summed to generate the opinion scores of sentences, and the negation words will negate the closest opinion words. Based on this approach, we further consider structural trios to experiment whether syntactic structures of sentences are beneficial for opinion analysis. Because the scoring functions may not be straight forward as those we have adopted for opinion word extraction, we did not design scoring functions for utilizing all types of structural trios. Instead, we emphasize their original opinion scores by multiplying a variable alpha to see whether these structures are important. In this paper, alpha equals five.

We have shown that word morphological structures benefit the word opinion extraction. When we experiment on sentences, we also incorporate the word morphological structures to see whether they are also useful for opinion analysis on sentences. Five experimental settings are listed as below:

(1) bag[w]-bag[s]: structural information is not considered for both words and sentences. The bag-of-character approach is used to calculate the opinion scores of words, and the bag-of-word approach sentences.

(2) struc[w]-bag[s]: morphological structures are utilized to calculate word opinion scores, but structural trios are not considered. The bag-of-word approach is used to calculate the opinion scores of sentences.

(3) bag[w]-struc[s]: structural trios are considered for calculating sentence opinion scores, while the bag-of-character approach is used to calculate the opinion scores of words.

(4) struc[w]-(m)struc[s]: both word morphological structures and manually labeled structural trios are adopted.

(5) struc[w]-struc[s]: both morphological structure of words and system labeled structural trios are adopted.

As we have shown that NTUSD is beneficial to the opinion analysis at word level, it is used as described in section 5.2 by default.

Our system adopted *CRF* algorithm to label structural trios for setting (5). The content string and the part of speech of the current node, its parent node, its offspring nodes in the next three generations, together with the depth of the current node in the Chinese Treebank, are used as the features for each node in *CRF*. The co-occurrence of the current node and all its siblings are defined in *CRF*'s template file. *CRF* will label whether the current node is the first child or the second child of a certain relation in a structural trio, or it is not part of any structural trios. A four-fold experiment is performed for the learning and testing of this labeling process by *CRF*.

6.4 Results and Discussion

Table 6 shows the statistics of manually labeled structural trios in Chinese Treebank and identification performance of *CRF*. Table 7 shows the performance of five experiment settings described in Section 6.3. The experiment results show that the morphological structures of words do not have a large contribution for opinion sentence analysis (setting 1 vs. setting 2; setting 3 vs. setting 4). However, considering the structural trios improve the performance.

Trio Type	Number	Percentage	f-Score
2	18,483	36.85%	0.4883
3	13,687	27.29%	0.4944
4	15,970	31.84%	0.6360
5	965	1.92%	0.2034
Others	1,054	2.10%	
Total	50159	100%	

Table 6: Statistics and Results of Identifying
Structural Trios

Setting	Word [w]	Sentence [s]	f-Score (opinion)	f-Score (polarity)
1	bag	bag	0.7073	0.4988
2	struc	bag	0.7162	0.5117
3	bag	struc	0.8000	0.5361
4	struc	(m)struc	0.7922	0.5297
5	struc	struc	0.7993	0.5187

Table 7: Results of Opinion Extraction
on Chinese Treebank

By summarizing the experimental results in Section 5 and this section, we can conclude that considering the word morphological structures benefits the opinion polarity detection, but in the current approach its assistance to words does not propagate to sentences. Considering the syntactic structures, however, do help in opinion analysis both for the opinion sentence extraction and the polarity detection. The performance of opinion extraction boosts to an f-score 0.80 and the performance of polarity detection an f-score 0.54.

However, the utilization of structure trios needs the parsing tree of sentences as the prior knowledge. Hence these two kinds of structural information may be suitable for different applications: structural trios for well written sentences such as those in the news articles, while the morphological structures for casually written sentences such as those appear in SMS messages or articles with limit length on the Web.

Because there are no opinion experiments performed on Chinese Treebank, we mention the performance of Ku's approach (setting (1)) for opinion sentence extraction, f-score 0.6846, in NTCIR-7 MOAT task, on news articles, as a result for comparison. Their approach was ranked the second in this task, and the best team achieved an f-score 0.7453.

7   Conclusion and Future Work

This paper considers morphological and syntactic structures in analyzing Chinese opinion words and sentences. For morphological structures, eight Chinese morphological types are defined.

CRF classifier and SVM classifier for morphological type classification are proposed. Experiments show that CRF classifier achieves the best accuracy 0.70 in type classification, which is 8% better than SVM classifier. We further show that word morphological structures benefit the opinion word extraction significantly. With the help of the sentiment dictionary NTUSD, the f-score of opinion word extraction achieves 0.77 and the f-score of the word polarity detection achieves 0.62 when the word morphological types are provided by the SVM classifier. They are comparably better than bag-of-character approach and the dictionary based approach.

We defined structural trios to represent the relations between sentence segments and also extract these relations using CRF algorithm. Results show that considering structural trios benefits the opinion analysis on sentences. An f-score 0.80 for opinion extraction and an f-score 0.54 for polarity detection are achieved, which is a great improvement.

The opinion scoring functions for morphological types and structural trios are critical for polarity detection, and scoring functions for words determine the scoring functions for sentences. Now we define these functions intuitively based on linguistic rules, but learning methods like regression will be investigated in the future. Examining the interaction of cues from word and sentence levels on the opinion sentence extraction and the opinion polarity detection is our next goal.

References

Banea, C., Mihalcea, R., Wiebe, J. and Hassan, S. 2008. Multilingual Subjectivity Analysis Using Machine Translation. In *Proceedings of Empirical Methods in Natural Language Processing (EMNLP 2008)*.

Bautin, M., Vijayarenu, L. and Skiena, S. 2008. International sentiment analysis for news and blogs. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*.

Carenini, G., Ng, R. T. and Pauls, A. 2006. Interactive Multimedia Summaries of Evaluative Text. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (pp. 124-131), Sydney, Australia.

Cesarano, C., Picariello, A., Reforgiato, D. and Subrahmanian, V.S. 2007. The OASYS 2.0 Opinion Analysis System. Demo in *Proceedings of International Conference on Weblogs and Social Media* (pp. 313-314), Boulder, CO USA.

Chang, Chih-Chung and Lin, Chih-Jen. 2001. LIBSVM: a library for support vector machines, http://www.csie.ntu.edu.tw/~cjlin/libsvm

Chen, A., Xu, L., Gey, F.C. and Meggs, J. 1997. Chinese Text Retrieval without Using a Dictionary. *ACM SIGIR Forum*, Volume 31, Issue SI (pp. 42-49).

Cheng, X.-H. and Tian, X.-L. 1992. *Modern Chinese*. Bookman Books Ltd.

Dave, K., Lawrence, S., and Pennock, D.M. 2003. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. In *Proc. of the 12th International WWW Conference* (pp. 519-528).

Kawai, Y., Kumamoto, T. and Tanaka, K. 2007. Fair News Reader: Recommending news articles with different sentiments based on user preference. In *Proceedings of Knowledge-Based Intelligent Information and Engineering Systems (KES), No. 4692 in Lecture Notes in Computer Science* (pp. 612–622).

Kim, S.-M. and Hovy, E. 2004. Determining the Sentiment of Opinions. In *Proc. of the 20th ICCL* (pp. 1367-1373).

Ku, L.-W. and Chen, H.-H. 2007. Mining Opinions from the Web: Beyond Relevance Retrieval. *Journal of American Society for Information Science and Technology*, Special Issue on Mining Web Resources for Enhancing Information Retrieval, 58(12), 1838-1850.

Lafferty, J., McCallum, A. and Pereira, F. 2001. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, In *Proc. of ICML* (pp.282-289).

Pang, B., Lee, L. and Vaithyanathan, S. 2002. Thumbs up? Sentiment Classification Using Machine Learning Techniques. In *Proc. of the 2002 Conference on EMNLP* (pp. 79-86).

Riloff, E. and Wiebe, J. 2003. Learning Extraction Patterns for Subjective Expressions. In *Proc. of the 2003 Conference on EMNLP* (pp. 105-112).

Seki, Y., Evans, D. K., Ku, L.-W., Sun, L., Chen, H.-H. and Kando, N. 2008. Overview of Multilingual Opinion Analysis Task at NTCIR-7. In *Proceedings of the 7th NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access.*

Somasundaran, S., Ruppenhofer, J. and Wiebe, J. 2007. Detecting arguing and sentiment in meetings. *Proceedings of the SIGdial Workshop on Discourse and Dialogue*, 2007.8.6

Takamura, H., Inui, T. and Okumura, M. 2005. Extracting Semantic Orientations of Words Using Spin Model. In *Proc. of the 43rd Annual Meeting of the ACL* (pp. 133-140).

Tzeng, H. and Chen, K.-J. 2002. Design of Chinese Morphological Analyzer. In *Proc. of the 1st SIGHAN Workshop on Chinese Language Processing*, vol.18, 1-7.

Wiebe, J. 2000. Learning Subjective Adjectives from Corpora. In *Proc. of the 17th National Conference on AAAI and Twelfth Conference on IAAI* (pp. 735-740).