# Disambiguating Dynamic Sentiment Ambiguous Adjectives

**Yunfang Wu**
"Key Laboratory of Computational
"'Linguistics (Peking University),
Ministry of Education"qh China

wuyf@pku.edu.cn

**Miaomiao Wen***
Department of Electrical Engineering and
Information Systems,
University of Tokyo

wenmiaomiao98@gmail.com

## Abstract

Dynamic sentiment ambiguous adjectives (DSAAs) like "large, small, high, low" pose a challenging task on sentiment analysis. This paper proposes a knowledge-based method to automatically determine the semantic orientation of DSAAs within context. The task is reduced to sentiment classification of target nouns, which we refer to sentiment expectation instead of semantic orientation widely used in previous researches. We mine the Web using lexico-syntactic patterns to infer sentiment expectation of nouns, and then exploit character-sentiment model to reduce noises caused by the Web data. At sentence level, our method achieves promising result with an f-score of 78.52% that is substantially better than baselines. At document level, our method outperforms previous work in sentiment classification of product reviews.

## 1 Introduction

In recent years, sentiment analysis has attracted considerable attention in the NLP community. It is the task of mining positive and negative opinions from natural language, which can be applied to many research fields. Previous work on this problem falls into three groups: opinion mining of documents, sentiment classification of sentences and polarity prediction of words.

Sentiment analysis both at document and sentence level rely heavily on word level.

The most frequently explored task at the word level is to determine the polarity of words, in which most work centers on assigning a prior polarity to words or word senses in the lexicon out of context. However, for some words, the polarity varies strongly with context, making it hard to attach each to a fixed sentiment category in the lexicon. For example, the word "low" has a positive orientation in "low cost" but a negative orientation in "low salary". We call these words like "low" dynamic sentiment ambiguous adjectives (DSAAs). Turney and Littman (2003) claim that DSAAs cannot be avoided in a real-world application. But unfortunately, DSAAs are discarded by most research concerning sentiment analysis.

In this paper, we are devoted to the challenging task of disambiguating DSAAs. The task is to automatically determine the semantic orientation (SO) of DSAAs within context. We limit our work to 14 frequently used adjectives in Chinese, such as "large, small, many, few, high, low", which all have the meaning of measurement. Although the number of such ambiguous adjectives is not large, they are frequently used in real text, especially in the texts expressing opinions and emotions. As demonstrated by the experimental results in this paper, the disambiguation of 14 DSAAs can obviously improve the performance of sentiment classification of product reviews.

The task of disambiguating DSAAs is reduced to sentiment classification of nouns. Previous studies classify nouns into three categories: positive, negative and neutral. In contrast, we propose two categories of sentiment expectation

---

of nouns: positive expectation and negative expectation. This paper presents a novel approach to automatically predict sentiment expectation of nouns. First, we infer the sentiment expectation of a noun by mining the Web with strongly-polar-steering lexico-syntactic patterns. Secondly, we derive the sentiment expectation of a noun from its component characters, which capture the semantic relationship between Chinese words and characters. Finally, a better performance is obtained by combing the two methods. We conduct two types of experiments: the experimental results at the sentence level validate the effectiveness of our approach; the experimental results at the document level confirm the significance of the problem we addressed.

## 2 Related Work

### 2.1 Word-level Sentiment Analysis

Recently there has been extensive research in sentiment analysis, for which Pang and Lee (2008) give an in-depth survey of literature. Closer to our study is the large body of work on automatic SO prediction of words (Hatzivassiloglou and McKeown, 1997; Turney and Littman, 2003; Kim and Hovy, 2004; Andreevskaia and Bergler, 2006), but unfortunately they all discard DSAAs in their research. In recent years, some studies go a step further, attaching SO to senses instead of word forms (Esuli and Sebastiani, 2006; Wiebe and Mihalcea, 2006; Su and Markert 2008), but their work is still limited in lexicon out of context.

The most relevant work is Ding et al. (2008), in which DSAAs are named as context dependant opinions. They argue that there is no way to know the SO of DSAAs without prior knowledge, and asking a domain expert to provider such knowledge is scalable. They adopt a holistic lexicon-based approach to solve this problem, which exploits external information and evidences in other sentences and other reviews. On the contrary in this paper, we obtain the prior knowledge of a product by mining the web, and then use such knowledge to determine the SO of DSAAs. The prior knowledge of a product, which is closer to the sentiment expectation of nouns described in this paper, is

an important research issue in itself and has many applications in sentiment analysis, as discussed in section 3.2.

### 2.2 Phrase-level Sentiment Analysis

The disambiguation of DSAAs can also be considered as a problem of phrase-level sentiment analysis. Wilson et al. (2004) present a two-step process to recognize contextual polarity that employs machine learning and a variety of features. Takamura et al. (2006, 2007) propose latent variable model and lexical network to determine SO of phrases, focusing on "noun+adjective" pairs. Their experimental results suggest that the classification of pairs containing ambiguous adjectives is much harder than those with unambiguous adjectives. The above mentioned approaches are all supervised, and need human labeled data for training. In contrast, our method is unsupervised and can overcome the data acquisition bottleneck. Moreover, we focus on the much harder task of disambiguating DSAAs in "noun+adjective" pairs.

### 2.3 Pattern-based Method

Previous studies have applied pattern-based method to sentiment analysis (Riloff and Wiebe, 2003; Wiebe et al., 2004; Riloff et al., 2005; Wiebe and Mihalcea, 2006; Andreevskaia and Berger; 2006). The differences with our method lie in two aspects: the used resources (corpus versus web) and the research target (subjective expressions versus sentiment expectation).

### 2.4 Character-based Method

Chinese characters carry semantic information that is indicative of semantic properties of words. Previous studies have exploited the character-based model to predict the semantic categories of Chinese unknown words (Chen, 2004; Lu, 2007). Yuen et al. (2004) presents a method to infer the SO of a Chinese word from its statistical association with strong-polarized characters rather than with strong-polarized words. The work by Ku et al. (2006) is similar to ours because they also define the sentiment score of a word by its composite characters. However, their algorithm is based only on frequency, while we exploit point mutual information that can capture the character-sentiment association.

# 3 Determining SO of Adjective by Target Noun

## 3.1 Classification of DSAAs

The frequently used DSAAs are given below. We group them into two categories: positive-like adjectives and negative-like adjectives. These adjectives are neutral out of context, but positive or negative emotion will be evoked when they co-occur with some target nouns, making it hard to assign each to a fixed sentiment category in lexicon.

(1) Positive-like adjectives (*Pa*) = { 大 da|large, 多 duo|many, 高 gao|high, 厚 hou|thick, 深 shen|deep, 重 zhong|heavy, 巨大 ju-da|huge, 重大 zhong-da|great}

(2) Negative-like adjectives (*Na*) ={小 xiao|small, 少 shao |few, 低 di|low, 薄 bao|thin, 浅 qian|shallow, 轻 qing|light}

## 3.2 Sentiment Expectation of Noun

The SO of most DSAAs can be determined by target nouns in noun-adjective phrases, as shown in Table 1. For example, the word "high" has a positive orientation when the target noun is "salary" but a negative orientation when the target noun is "price". Therefore, the task can be reduced to sentiment classification of nouns.

| Positive 潜力大|potential is great 工资高|salary is high | Negative 潜力小|potential is small 工资低|salary is low |
|---|---|
| Negative 压力大|pressure is big 价格高|price is high | Positive 压力小|pressure is small 价格低|price is low |

Table 1: The SO of DSAAs in noun-adjective phrases

In previous research, the SO of nouns is classified into three categories: positive, negative and neutral. Accordingly, "压力 ya-li|pressure" will be assigned as negative and "潜力 qian-li|potential" as positive, while "工资 gong-zi|salary" and "价格 jia-ge|price" will be assigned as neutral, as the two terms are objective and cannot evoke positive or negative emotion. Different from the traditional classification scheme, we propose sentiment expectation and classify nouns into two categories: positive expectation and negative expectation. For a positive expectation noun, people usually expect the thing referred to by the noun to be bigger, higher or happen frequently. On the contrary, for a negative expectation noun，people usually expect the thing referred to by the noun to be smaller, lower or don't happen . For example, "价格 jia-ge|price" is a negative expectation noun, as most people in most cases expect that the product prices become low, whereas "工资 gong-zi|salary" is a positive expectation noun, as most people in most cases expect that their salaries become high. The relationship between traditional SO and sentiment expectation can be defined as: positive (negative) terms correspond to positive (negative) expectation terms, but some neutral terms may also carry positive (negative) expectation.

Su and Markert (2008) argue that polarity can also be attached to objective words. The difference with our scheme is that, for example, "价格 jia-ge|price" is attached to negative expectation in our scheme while is still neutral in Su and Markert's method.

The distinction between positive and negative expectation nouns is vital to determine the SO of some phrases. Using it to disambiguate DSAAs is a good example. Another application is the phrase containing verbs with the meaning of status change. For example, "工资上涨了|salary has been raised" will evoke positive emotion, while "价格上涨了 jiage-shangzhang-le|prices have gone up" will evoke negative emotion. As far as we are aware, this is the first sentiment analysis scheme that tries to exploit people's expectation towards nouns.

## 3.3 Determination of DSAAs

The SO of DSAAs in a given phrase can be calculated by Eq. (1).

$$C(a) = \begin{cases} 1 & \text{if } a \text{ is positive-like} \\ -1 & \text{if } a \text{ is negative-like} \end{cases} \quad (1)$$

$$C(n) = \begin{cases} 1 & \text{if } n \text{ is positive expectation} \\ -1 & \text{if } n \text{ is negative expectation} \end{cases}$$

$$SO(a) = C(a) * C(n)$$

If adverb="不 bu|not"，$SO(a)= -SO(a)$

Where C(a) denotes the category of DSAAs; C(n) denotes the sentiment expectation of nouns; SO(a) is the SO of DSAAs in a give noun-adjective phrase. When the adverb is the negation term "不 bu|not", the SO is reversed.

## 4 Predicting Sentiment Expectation of Noun

### 4.1 Pattern-based Prediction Using a Web Search Engine

In natural language, there are some lexico-syntactic patterns that people frequently use when they express their opinion about something. For example:

(3) 工资有点低 | Salary is *a little* low.

(4) 价格有点高 | Price is *a little* high.

The pattern "<n> 有点 <a>" carries a strong negative association in Chinese language. When a man is saying "工资有点低| Salary is a little low", it indicates that he wishes his "工资 |salary" to be raised. On the contrary, when a man is saying "价格有点高| price is a little high", it indicates that he wishes "价格 |price" to go down. As a result, "工资 |salary" has positive expectation while "价格 |price" has negative expectation.

With the rapid development and expansion of the internet, Web has become an important medium for people to post their ideas. The opinions expressed on the Web reflect the common cognition shared by collection of people in a culture. Therefore, using a Web search engine with the strong-polar-steering lexico-syntactic patterns as queries, we can infer the sentiment expectation of a noun, by calculating its statistical association with positive and negative hits.

As an example, using the search engine Baidu[2] with the pattern "<n> 有点 <a>" as queries, we obtain the following hits:

(5) 工资有点低 | Salary is a little low. (2890 hits)
    工资有点高 | Salary is a little high     (67 hits)

(6) 价格有点高 | Price is a little high.  (19400 hits)
    价格有点低 | Price is a little low.    (1080 hits)

The more than 40 times more numerous hits for "工资有点低 |Salary is a little low" indicate that that "工资|salary" is a positive expectation noun. For the same reason, we can infer that "价格 |price" has negative expectation.

DSAAs are classified into two opposite sets *Pa* and *Na*, as listed in (1) and (2) respectively.

Here two-character adjectives ("巨大 |huge" and "重大 |great") are discarded. Four types of lexico-syntactic patterns, which are also classified into two opposite sets in consistent with *Pa* and *Na*, are used in this paper, as listed in Table 2. These patterns were manually designed, inspired by linguistic knowledge and after a deep investigation on the Web.

| Pos. expectation patterns | Neg. expectation patterns |
|---|---|
| 1) <n> 有点 *Na*<br>n is a little *Na* | 1) <n> 有点 *Pa*<br>n is a little *Pa* |
| 2) <n> 有点儿 *Na*<br>n is a little *Na* | 2) <n> 有点儿 *Pa*<br>n is a little *Pa* |
| 3) <n> *Na*, 怎么办<br>n is *Na*, what should we do? | 3) <n> *Pa*, 怎么办<br>n is *Pa*, what should we do? |
| 4) 嫌 <n> *Na*<br>n is too *Na* | 4) 嫌 <n> *Pa*<br>n is too *Pa* |

Table 2: The lexico-syntactic patterns

Here the noun (n) in these patterns was instantiated by 9,468 nouns in our collected data. A noun has together 48 patterns, 24 positive and 24 negative ones. For each noun, we obtain the hits of both positive and negative expectation patterns, using the search engine Baidu. The sentiment expectation of a noun is acquired by Eq. (2) and Eq. (3), where the magnitude of $PT\_SO(n)$ can be considered as the strength of sentiment expectation.

$$PT\_SO(n) = \sum_{b \in Na} \sum_{i=1}^{4} PositivePatternHit_i(n, b)$$

$$- \sum_{a \in Pa} \sum_{i=1}^{4} NegativePatternHit_i(n, a) \quad (2)$$

$$\text{n is} \begin{cases} \text{positive expectation if } PT\_SO(n) > 0 \\ \text{negative expectation if } PT\_SO(n) < 0 \\ \text{not predicted} \qquad \text{if } PT\_SO(n) = 0 \end{cases} \quad (3)$$

Table 3 gives some nouns with sentiment expectation predicted by the pattern-based method, descending (the left column) and ascending (the right column) by the absolute value of $PT\_SO(n)$. Most words (9 out of 10) are correctly predicted, demonstrating that the result of pattern-based method is promising. The only wrong predicted noun is "感觉 |feeling", due to the fact that most instances of it on the Web data are used as verb rather than noun, like "感觉有点大| I think it is large".

| Positive expectation | Negative expectation |
|---|---|
| Noun ( $PT\_SO(n)$ ) | Noun ( $PT\_SO(n)$ ) |
| 钱\|money (31349) | 温度\|temperature(-111576) |
| 工资\|wage (26311 ) | 噪音\|noise (-45790) |
| 感觉\|feeling (20102) | 价格\|price (-25653) |
| 收入\|income(19429) | 代价\|cost (-22051) |
| 官\|officer (10630) | 血压\|blood pressure (-21788) |

Table 3: Examples of nouns with sentiment expectation predicted by the pattern-based method

## 4.2 Character-based Derivation Using Sentiment Lexicons

But the sentiment expectation of some nouns cannot be predicted with the pattern-based method, mainly due to the reason that these nouns don't occur in the listed patterns in Table 2. An alternate way is to exploit the semantic knowledge of Chinese characters. It is assumed that there is a strong association between the sentiment category of a word and its component characters. For example, the three words "罪恶 zui'e|evil, 罪行 zuixing|crime, 罪过 zuiguo|fault", which all contain the character "罪 zui|sin" that carries negative meaning, are all negative expectation nouns.

First, we compute the character-word sentiment association by the following PMI formula, based on a sentiment lexicon:

$$PMI(c, Positive) = \log \frac{P(c, Positive)}{P(c)P(Positive)}$$

$$PMI(c, Negative) = \log \frac{P(c, Negative)}{P(c)P(Negative)} \quad (4)$$

$$SO(c) = PMI(c, Positive) - PMI(c, Negative)$$

Where $P(c, Positive)$ is the probability of a character $c$ in the positive category; $P(c)$ is the probability of a character $c$ in the sentiment lexicon; $P(Positive)$ is the probability of the positive category in the sentiment lexicon. $PMI(c, Negative)$ has the similar meaning. Probabilities are estimated according to the maximum likelihood principle.

The open language resources for Chinese sentiment analysis are quite limited. We selected the following two sentiment lexicons.

**Sentiment HowNet**. HowNet has published the Chinese vocabulary for sentiment analysis[3],

which was manually constructed. The positive category contains 4,566 words and the negative category contains 4,370 words.

**Sentiment BaiduHit**. In our collected data, we extracted 9,468 nouns. Using the pattern-based method we acquired sentiment expectation of these nouns, where 2,530 ones were assigned as positive expectation, 1,837 ones as negative expectation and 5,101 ones were not predicted. It is assumed that most nouns are correctly predicted. These nouns with their sentiment expectation constitute the lexicon of Sentiment BaiduHit, which is automatically constructed.

**Combining HowNet and BaiduHit.** Most sentiment characters derived from HowNet have adjective property, since most words in Sentiment HowNet are adjectives. On the contrary, most sentiment characters derived from BaiduHit have noun property. Therefore, the combination of the two lexicons can cover more characters. As Sentiment HowNet is manually compiled, the sentiment characters derived from it should be more reasonable than those from BaiduHit. When combining the two lexicons in computing character polarity, we assign a high priority to HowNet. Only when a character is out of vocabulary in HowNet, we resort to BaiduHit.

Then, we acquire the sentiment category of a word by computing the following equation. Let a word consist of $n$ characters $w = c_1, c_2, \ldots c_n$, the sentiment category of the word is calculated by the average sentiment value of its component characters:

$$CH\_SO(w) = \frac{1}{n} \sum_{i=1}^{n} SO(c_i) \quad (5)$$

$$w \text{ is } \begin{cases} \text{positive expectation if } CH\_SO(w) > 0 \\ \text{negative expectation if } CH\_SO(w) < 0 \quad (6) \\ \text{neutral} \qquad\qquad \text{if } CH\_SO(w) = 0 \end{cases}$$

We acquired *sentiment expectation* of 9,468 nouns in our collected data, based on Sentiment HowNet, Sentiment BaiduHit, and the combination of the two lexicons, respectively.

Table 6 gives examples of nouns with sentiment expectation acquired by the character-based method combining the two lexicons of HowNet and BaiduHit, descending (the left column) and ascending (the right column) by the absolute value of $CH\_SO(w)$.

| Positive expectation | Negative expectation |
|---|---|
| Noun($CH\_SO(w)$) | Noun($CH\_SO(w)$) |
| 美称\|good name (3.23) | 灰\|ash (-3.22) |
| 健美\|health (3.06) | 毛\|gross (-2.93) |
| 香\|fragrance (3.05) | 税\|tax (-2.89) |
| 美方\|U.S.A (2.98) | 毛病\|fault (-2.84) |
| 职称\|title (2.64) | 毒\|poison (-2.82) |

Table 4: Example of nouns with sentiment expectation predicted by the character-based method

## 4.3 Integrating Pattern-based Prediction and Character-based Derivation

The two methods of pattern-based prediction and character-based derivation have complementary properties. The pattern-based method concentrates on a word's usage on the Web, whereas the character-based method focuses on the internal structure of a word. So the two methods can be integrated to get better performance. The results using pattern-based method are much better than character-based method, as illustrated in Table 3 and Table 4. So in the integrated scheme, we give a high priority to pattern-based method. The pattern-based approach is mainly used, and only when the value of $|PT\_SO(n)|$ is smaller than a threshold $r$, the character-based method is adopted. Because when the value of $|PT\_SO(n)|$ is very small, it could be caused by random noises on the Web. We set $r$ to 9 according to empirical analysis in the development data.

## 5 Experiments

### 5.1 Sentiment Analysis at Sentence Level

#### 5.1.1 Data

We collected data from two sources. The main part was extracted from Xinhua News Agency of Chinese Gigaword (Second Edition) released by LDC. The texts were automatically word-segmented and POS-tagged using the open software ICTCLAS[4]. In order to concentrate on the disambiguation of DSAAs, and reduce the noise introduced by the parser, we extracted sentences containing strings in pattern of (7), where the target noun is modified by the adjective in most cases.

(7) noun+adverb+adjective (adjective ∈ DSAAs)
e.g. 成本/n 较/d 低/a | the cost is low.

Another small part of data was extracted from the Web. Using the search engine Google[5], we searched the queries as in (8):

(8) 很| very+ adjective (adjective ∈ DSAAs )
From the returned snippets, we manually picked out some sentences that contain the strings of (7). Also, the sentences were automatically word-segmented and POS-tagged using ICTCLAS.

DSAAs in the data were assigned as positive, negative or neutral, independently by two annotators. Since we focus on the distinction between positive and negative categories, the neutral instances were removed. Table 5 gives statistics of the data, and the inter-annotator agreement is in a high level with a kappa of 0.91. After cases with disagreement were negotiated between the two annotators, a gold standard annotation was agreed upon. In this paper, 3066 instances were divided randomly into three parts, 1/3 of which were used as the development data, and 2/3 were the test data.

Most of the data has been used as the benchmark dataset of SemEval-2010 task 18 "disambiguating sentiment ambiguous adjectives" (Wu and Jin, 2010), and so it can be downloaded freely for research.

|  | Pos# | Neg# | Total# |
|---|---|---|---|
| Pos# | 1280 | 58 | 1338 |
| Neg# | 72 | 1666 | 1738 |
| Total# | 1352 | 1724 | 3066 |

Table 5: The statistics of DSAAs data

#### 5.1.2 Baseline

We conducted two types of baseline.

**Simple Baseline**. Not considering the context, assign all positive-like adjectives as positive, and all negative-like adjectives as negative.

**HowNet Baseline**. Acquiring SO of nouns from Sentiment HowNet, the polarity of DSAAs is computed by Eq. (1).

#### 5.1.3 Methods

**Pattern-based method**. Acquiring sentiment expectation of nouns using the pattern-based method, the polarity of DSAAs is computed by Eq.(1).

---

**Character-based method**. Acquiring sentiment expectation of nouns using the character-based method, based on Sentiment HowNet, Sentiment BaiduHit and the combination of the two lexicons respectively, the polarity of DSAAs is computed by Eq.(1).

**Integrated method**. Acquiring sentiment expectation of nouns by integrating pattern-based and character-based methods, the polarity of DSAAs is computed by Eq. (1).

### 5.1.4 Results

Table 6 gives the experimental results at sentence level with different methods.

| Methods | Pre. | Rec. | F |
|---|---|---|---|
| Simple Baseline | 61.20 | 61.20 | 61.20 |
| HowNet Baseline | 97.58 | 9.88 | 17.94 |
| Pattern-based | 75.83 | 71.67 | 73.69 |
| Character-based (HowNet) | 69.89 | 69.37 | 69.63 |
| Character-based (BaiduHit) | 68.66 | 68.59 | 68.62 |
| Character-based (Combined) | 71.01 | 70.94 | 70.97 |
| **Integrated method** | **78.52** | **78.52** | **78.52** |

Table 6: The experimental results at sentence level

As for the simple baseline, both the precision and recall are low, suggesting that DSAAs cannot be neglected for sentiment analysis in a real-world application.

The HowNet baseline achieves a quite high precision of 97.58%, but a rather poor recall of 9.88%, suggesting that SO of nouns described in traditional sentiment lexicon, like HowNet, cannot effectively disambiguate DSAAs.

The proposed methods in this paper all yield results that are substantially better than two types of baseline. The pattern-based method, as straightforward as it is, achieves promising result with an f-score of 73.69%, which is 12.49% higher than the simple baseline. The pattern-based method outperforms the character-based method (combined) by 4.82% in precision and 0.73% in recall. The performance of the character-based method based on Sentiment BaiduHit is competitive with that based on Sentiment HowNet, which again proves the effectiveness of the pattern-based method. The character-based method combining the two lexicons outperforms each lexicon with small improvement. The approach integrating pattern-based and character-based methods outperforms each method in isolation, achieving an f-score of 78.52% that is 17.32% higher than the simple baseline and 60.58% higher than HowNet baseline.

## 5.2 Sentiment Analysis at Document Level

### 5.2.1 Data

We also investigated the impact of disambiguating DSAAs on the sentiment classification of product reviews. Following the work of Wan (2008), we selected the same dataset. The dataset contains 886 Chinese product reviews, which are manually annotated with polarity labels: positive or negative. Also, the files are automatically word-segmented and POS-tagged using ICTCLAS. We extracted the files that contain the following strings, where the nouns are modified by DSAAs in most cases.

(9)  noun+adjective            (adjective $\in$ DSAAs)
     noun+adverb+adjective
     noun+adverb+adverb+adjective.

We obtained 212 files, up to 24% of the overall data, suggesting again that DSAAs are frequently used in product reviews and cannot be avoided in a real-world application.

### 5.2.2 Methods

Our goal is not to propose a new method, but instead to test the performance gain by adding the disambiguation of DSAAs. We adopted the same algorithm with Wan (2008), and also used Sentiment-HowNet. But in our experiment, *Negation_Dic* contains only one term " 不 bu|not", for the sake of repeatable experiments.

The baseline algorithm is illustrated by the non-italic part in Figure 1, where we set the same parameters with Wan's approach: *PosValue*=1, *NegValue*=-2, *q*=2, *ρ*=2.

We added the disambiguation of DSAAs to the algorithm, as illustrated by the italic part in Figure 1. When a word is a DSAA, compute its SO with the proposed integrated method, rather than using its prior polarity specified in HowNet. For *Dy_PosValue* and *Dy_NegValue,* we first set *Dy_PosValue*=1 and *Dy_NegValue*=-2, just the same as *PosValue* and *NegValue.* In the second attempt, in order to further intensify the polarity of DSAAs*,* we set *Dy_PosValue*=1.5 and *Dy_NegValue*=-2.5. Other parameters were set the same as baseline.

**Algorithm Compute_SO:**

1. Tokenize document d into sentence set S, and each sentence s∈S is tokenized into word set Ws;

2. For any word w in a sentence s∈S, compute its value SO(w) as follows:

*1) if w∈DSAAs, compute SO(w) with the integrated method.*

   If SO(w)=1, SO(w)=Dy_PosValue;
   If SO(w)=-1, SO(w)=Dy_NegValue;

2) if w∈Positive_Dict, SO(w)=PosValue;

3) If w∈Negative_Dict, SO(w)=NegValue;

4) Otherwise, SO(w)=0;

5) Within the window of q words previous to w, if there is a term w'∈Negation_Dict,
   SO(w)= –SO(w);

6) Within the window of q words previous to w, if there is a term w'∈Intensifier_Dict,
   SO(w) =ρ×SO(w);

3. $S(d) = \sum_{s \in S} \sum_{w \in Ws} SO(w)$

Figure 1: Algorithm of computing SO of documents

### 5.2.3 Results

Adding the disambiguation of *DSAAs*, the performance of sentiment classification of 212 product reviews was significantly improved, as shown in Table 7.

|  |  | Baseline | DSAAs (1, -2) | DSAAs (1.5, -2.5) |
|---|---|---|---|---|
| Pos. | Pre. | 75.89 | 77.50 | 76.61 |
|  | Rec. | 78.70 | 86.11 | 87.96 |
|  | F | 77.27 | 81.58 | 81.90 |
| Neg. | Pre. | 87.01 | 88.46 | 87.06 |
|  | Rec. | 64.42 | 66.35 | 71.15 |
|  | F | 74.03 | 75.82 | 78.31 |
| Total | MacroF | 75.62 | 78.60 | **80.06** |
|  | Accu. | 71.70 | 76.42 | **79.72** |

Table 7: The experimental results at document level

As an example, the following review, which consists of only one sentence, is correctly classified as positive by DSAAs method, but is classified as negative by the baseline approach.

(10) 体积小，重量轻，携带很方便。
| Small size, light weight, and easy to carry.

According to HowNet, as shown in Table 8, the sentence contains two negative words " 小 |small" and "轻|light" and one positive word "方便 fangbian|easy", resulting the overall negative prediction. In our approach, "体积 tiji|size" and " 重量 zhongliang|weight" are assigned as negative expectation, and consequently both "体积小|small size" and "重量轻|light weight" have

positive meaning, resulting the overall positive prediction.

| Pos. | 大 |large, 高 |high, 厚 |thick, 深 |deep, 重|heavy, 重大|great |
|---|---|
| Neg. | 小 |small, 低 |low, 薄 |thin, 浅 |shallow, 轻|light |
| OOV | 多|many, 少|few, 巨大|huge |

Table 8: The SO of DSAAs described in HowNet

Adding the disambiguation of DSAAs, our method obviously outperforms the baseline by 4.44% in f-score and 8.02% in accuracy. The improvement in recall is especially obvious. When intensifying the polarity of DSAAs by setting *Dy_PosValue*=1.5 and *Dy_NegValue*=-2.5, the recall is improved by 9.26% for positive category and 6.73% for negative category.

## 6   Conclusion and Future Work

This paper presents a knowledge-based unsupervised method to automatically disambiguate dynamic sentiment ambiguous words, focusing on 14 DSAAs. We exploit pattern-based and character-based methods to infer sentiment expectation of nouns, and then determine the polarity of DSAAs based on the nouns. For the sentiment analysis at sentence level, our method achieves promising result that is significantly better than two types of baseline, which validates the effectiveness of our approach. We also apply the disambiguation of 14 DSAAs to the sentiment classification of product reviews, resulting obvious improvement in performance, which proves the significance of the issue.

There leaves room for improvement. Our future work will explore more contextual information in disambiguating DSAAs. In addition, we will find out new methods to reduce noises when mining the Web to infer sentiment expectation of nouns. Discovering the lexico-syntactic patterns for sentiment expectation of nouns automatically or semi-automatically with bootstrapping method is also a challenging direction.

# References

Andreevskaia A. and Bergler S. 2006. Sentiment tagging of adjectives at the meaning level. *The 19th Canadian Conference on Artificial Intelligence.*

Andreevskaia, A. and Bergler, S. 2006. Mining WordNet for fuzzy sentiment: Sentiment tag extraction from WordNet glosses. *Proceedings of EACL 2006.*

Chen, C-J. 2004. Character-sense association and compounding template similarity: automatic semantic classification of Chinese compounds. *Proceedings of the 3rd workshop on Chinese language processing.*

Ding X., Liu B. and Yu, P. 2008. A holistic lexicon-based approach to opinion mining. *Proceedings of WSDM'08.*

Esuli, A. and Sebastiani, F. 2006. SentiWordNet: a publicly available lexical resource for opinion mining. *Proceedings of LREC'06.*

Hatzivassiloglou, V. and McKeown, K. 1997 Predicting the semantic orientation of adjectives. *Proceedings of ACL'97.*

Kim, S and Hovy, E. 2004. Determining the sentiment of opinions. *Proceedings of COLING'04.*

Ku, L, Liang Y. and Chen, H. 2006. Opinion extraction, summarization and tracking in news and blog corpora. *Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs.*

Lu X-F, 2007. Hybrid models for semantic classification of Chinese unknown words. *Proceedings of NAACL HLT'07..*

Pang, B. and Lee, L. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval.*

Riloff, E. and Wiebe, J. 2003. Learning Extraction Patterns for Subjective Expressions. *Proceedings of EMNLP'03.*

Riloff, E., Wiebe, J. and Phillips, W. 2005. Exploiting Subjectivity Classification to Improve Information Extraction. *Proceedings of AAAI'05.*

Su, F. and Markert, K. 2008. From words to senses: a case study of subjectivity recognition. *Proceedings of COLING'08.*

Takamura, H., Inui,T. and Okumura, M. 2006. Latent Variable Models for Semantic Orientations of phrases. *Proceedings of EACL'06.*

Takamura, H., Inui,T. and Okumura, M. 2007. Extracting Semantic Orientations of Phrases from Dictionary. *Proceedings of NAACL HLT '07.*

Turney, P. and Littman, M. 2003. Measuring praise and criticism: inference of semantic orientation from association. *ACM transaction on information systems.*

Wan, X. 2008. Using Bilingual Knowledge and Ensemble Techniques for Unsupervised Chinese Sentiment Analysis. *Proceedings of EMNLP'08.*

Wiebe, J. and Mihalcea, R. 2006. Word sense and subjectivity. *Proceedings of ACL'06.*

Wiebe, J., Wilson, T., Bruce, R., Bell, M. and Martin, M. 2004. Learning Subjective Language. *Computational Linguistics.*

Wilson, T., Wiebe, J. and Hoffmann, P. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. *Proceedings of HLT/EMNLP'05.*

Wu, Y. and Jin, P. 2010. SemEval-2010 task 18: disambiguating sentiment ambiguous adjectives. *Proceedings of SemEval 2010.*

Yuen R., Chan T., Lai T., Kwong O., T'sou B. 2004. Morpheme-based derivation of bipolar semantic orientation of Chinese words. *Proceedings of COLING'04.*