

# Analyzing Sentiment Word Relations with Affect, Judgment, and Appreciation

*Alena NEVIAROUSKAYA Masaki AONO*

TOYOHASHI UNIVERSITY OF TECHNOLOGY, 1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Japan  
alena@kde.cs.tut.ac.jp, aono@kde.cs.tut.ac.jp

## ABSTRACT

In this work, we propose a method for automatic analysis of attitude (affect, judgment, and appreciation) in sentiment words. The first stage of the proposed method is an automatic separation of unambiguous affective and judgmental adjectives from miscellaneous that express appreciation or different attitudes depending on context. In our experiments with machine learning algorithms we employed three feature sets based on Pointwise Mutual Information, word-pattern co-occurrence, and minimal path length. The next stage of the proposed method is to estimate the potentials of miscellaneous adjectives to convey affect, judgment, and appreciation. Based on the sentences automatically collected for each adjective, the algorithm analyses the context of phrases that contain sentiment word by considering morphological tags, high-level concepts, and named entities, and then makes decision about contextual attitude labels. Finally, the appraisal potentials of a word are calculated based on the number of sentences related to each type of attitude.

---

KEYWORDS : Appraisal potentials, Attitude lexicon, Minimal path length, Pointwise Mutual Information, Sentiment lexicon, Word-pattern co-occurrence.

---

## 1 Introduction and related work

*'Attitudinal meanings tend to spread out and colour a phase of discourse as speakers and writers take up a stance oriented to affect, judgment or appreciation.'*

Martin and White (2005: 43)

Rapid growth of online media and sources of different genres (blogs, product or service reviews, social networks etc.) has prompted the emergence and development of a sentiment analysis field aimed at automatic analysis of people's preferences, emotions, and attitudes communicated through written language. A variety of lexical resources has been created to support recognition and interpretation of different kinds of subjective phenomena: subjective (Wilson, Wiebe, & Hoffmann, 2005), polarity (Esuli & Sebastiani, 2006; Hatzivassiloglou & McKeown, 1997; Neviarouskaya, Prendinger, & Ishizuka, 2011), affective (De Albornoz, Plaza, & Gervás, 2012; Strapparava & Valitutti, 2004), and appraisal (Argamon, Bloom, Esuli, & Sebastiani, 2007) lexicons.

The subjectivity lexicon developed by Wilson et al. (2005) is comprised by over 8000 subjectivity clues annotated by type (strongly subjective / weakly subjective) and prior polarity (positive/negative/both/neutral). Hatzivassiloglou and McKeown (1997) created a list of 1336 adjectives manually labeled as either positive or negative. Esuli and Sebastiani (2006) developed a SentiWordNet lexicon based on WordNet (Miller, 1990) synsets comprised from synonymous terms. Three numerical scores characterizing to what degree the terms included in a synset are objective, positive, and negative, were automatically determined based on the proportion of eight ternary classifiers that assigned the corresponding label to the synsets of adjectives, adverbs, nouns, and verbs by quantitatively analysing the glosses associated with them. Neviarouskaya et al. (2011) developed a SentiFul lexicon using the core of sentiment lexicon and automatically expanding it through direct synonymy and antonymy relations, hyponymy relations, and manipulations with morphological structure of words (derivation and compounding). Aimed at introducing the hierarchy of affective domain labels, Strapparava and Valitutti (2004) manually created WordNet-Affect, a lexicon of affective concepts. An affective lexicon SentiSense (De Albornoz et al., 2012) that contains concept-level emotional annotations has been developed semi-automatically by considering semantic relations between synsets in WordNet. The appraisal lexicon (Argamon et al., 2007) developed by applying supervised learning to WordNet glosses contains adjectives and adverbs annotated by attitude type and force.

Methods for extracting and annotating sentiment-related terms include: machine learning approaches examining the conjunction relations between adjectives (Hatzivassiloglou & McKeown, 1997); clustering adjectives according to distributional similarity based on a small amount of annotated seed words (Wiebe, 2000); pattern-bootstrapping algorithms to extract nouns (Riloff, Wiebe, & Wilson, 2003); consideration of web-based mutual information in ranking the subjective adjectives (Baroni & Vegnaduzzo, 2004); bootstrapping algorithm employing a small set of seed subjective terms and an online dictionary, plus filtering the candidates based on a similarity measure (Banea, Mihalcea, & Wiebe, 2008); methods employing WordNet structure relations (Andreevskaia & Bergler, 2006; Kamps & Marx, 2002; Kim & Hovy, 2004; Takamura, Inui, & Okumura, 2005); and sentiment tagging based on morphological structure of words (Ku, Huang, & Chen, 2009; Moilanen & Pulman, 2008; Neviarouskaya et al., 2011). To assign subjectivity labels to word senses, methods relying on distributional similarity (Wiebe & Mihalcea, 2006) and on semi-supervised minimum cut algorithm (Su & Markert, 2009) have been proposed.

The goal of our research is to develop a method for automatic analysis of attitude expressed by sentiment words. Such method will support analytical applications relying on recognition of fine-grained context-dependent attitudes conveyed in text. According to the Appraisal Theory (Martin & White, 2005), there are three high-level attitude types: affect (a personal emotional state, feeling, or reaction), judgment (an ethical appraisal of person’s character, behaviour, skills etc.), and appreciation (an aesthetic evaluation of semiotic and natural phenomena, events, objects etc.). We distinguish sentiment-related adjectives expressing unambiguous attitude type (e.g., *happy* conveys affect, *fainthearted* – judgment, and *tasty* – appreciation) and ambiguous attitude type that depends on context (e.g., *useless* expresses affect in the context of *my useless attempts*, judgment in case of *his useless skills*, and appreciation in the phrase *useless information*).

In the first stage of the proposed method, unambiguous affective and judgmental adjectives are automatically separated from miscellaneous adjectives expressing unambiguous appreciation or different attitudes depending on context. The classification is based on a machine learning algorithm employing three feature sets based on Pointwise Mutual Information (PMI), word-pattern co-occurrence, and minimal path length. An early attempt to determine the potentials of an adjective to express affect, judgment or appreciation in evaluative discourse was made by Taboada and Grieve (2004), who calculated the PMI with the pronoun-copular pairs ‘*I was (affect)*’, ‘*He was (judgement)*’, and ‘*It was (appreciation)*’. However, affect-conveying adjectives (e.g., ‘*depressed*’) may equally well occur not only with first person pronouns, but also with third person pronouns, thus describing emotional states experienced by oneself or by other person. Our PMI features are inspired by the approach from (Taboada & Grieve, 2004). However, as distinct from their method, we calculate the strength of the association between attitude-conveying adjectives and patterns, in which they most probably occur (the example patterns for affect and judgment are ‘*feel XX*’ and ‘*XX personality*’, respectively). The next stage of the proposed method is to estimate the potentials of miscellaneous adjectives to convey affect, judgment, and appreciation. Based on the sentences automatically collected for each adjective, the algorithm analyses the context of phrases that contain sentiment word and makes decision about contextual attitude labels. Finally, the appraisal potentials of a word are calculated based on the number of sentences related to each type of attitude.

The remainder of the paper is structured as follows: In Section 2, we describe the method for separation of unambiguous affective and judgmental adjectives from miscellaneous. The algorithm for estimation of the potentials of miscellaneous adjectives to express affect, judgment, and appreciation is detailed in Section 3. In next section, we conclude the paper.

## 2 Method for separation of unambiguous affective and judgmental adjectives from miscellaneous

### 2.1 Data set

For the evaluation of the proposed methodology, we have extracted 1500 attitude-annotated adjectives from the AttitudeFul database (Neviarouskaya, 2011). These adjectives are annotated by at least one of 13 labels: nine for affect (AFF), two for positive and negative judgment (JUD), and two for positive and negative appreciation (APP). As we are interested in separating unambiguous affective (e.g., *joyful*) and judgmental (e.g., *egoistic*) adjectives from miscellaneous (MISC, e.g., *good*) that express appreciation or different attitudes depending on context (for example, *good feeling* expresses positive affect, *good parent* is positive judgment, and *good book* is positive appreciation), we have considered the following top-level labels: AFF, JUD, and MISC (namely, APP and combinations AFF-APP, AFF-JUD, JUD-APP, and AFF-JUD-APP).

The distribution of classes is as follows: AFF – 510 (34.0%), JUD – 414 (27.6%), and MISC – 576 (38.4%) adjectives. The examples are listed in Table 1.

Class	Adjectives
AFF	Euphoric, disheartened, frightened, infuriated, impressed
JUD	Altruistic, brave, diligent, high-principled, tenderhearted, despotic, egoistic, ill-famed, unkind
MISC	APP: comfortable, tasty, poorly-adapted AFF-APP: healthy, devastated AFF-JUD: enthusiastic, jealous JUD-APP: adorable, cheap, nonproductive AFF-JUD-APP: balanced, calm, genuine, unfriendly, worthless

TABLE 1 – Examples of adjectives from the data set.

## 2.2 Feature sets

In our experiments we employed the following feature sets that are further described in details:

1. Pointwise Mutual Information (PMI) based.
2. Word-pattern co-occurrence (WPC) based.
3. Minimal path length (MPL), or proximity, based.

The complete feature set is comprised of 88 features. These features were automatically defined for each adjective from the attitude-annotated data set in order to conduct experiments with cross-validation process.

### 2.2.1 Pointwise Mutual Information (PMI) based feature set

The Pointwise Mutual Information had been used by researchers to calculate the strength of the semantic association between words (Church & Hanks, 1989), to determine the semantic orientation (positive or negative) of words (Turney & Littman, 2002), and to measure the strength of the association between attitude-conveying adjectives and pronoun-copular pairs, such as ‘*I was*’, ‘*he was*’, and ‘*it was*’ (Taboada & Grieve, 2004). In defining PMI features we partially follow the approach from (Taboada & Grieve, 2004). However, as distinct from their method, we calculate the strength of the association between attitude-conveying adjectives and patterns, in which they most probably occur.

The Pointwise Mutual Information is calculated based on the following equation:

$$PMI(word, pattern) = \log_2 \frac{hits(word \text{ in a pattern})}{hits(word) \times hits(pattern)}, \quad (1)$$

where *word* stands for one of the adjectives; *pattern* – one of the patterns for affect or judgment; and *hits* – number of hits in the search engine.

Based on the definitions from the Appraisal theory (Martin & White, 2005), we defined the patterns as indicators of affect and judgment (10 and 20 patterns, respectively). They are given in Table 2.

Affect patterns	Judgment patterns	
feel XX (e.g., <i>feel happy</i> )	XX character	XX is a character
XX emotion	XX personality	XX is a personality
XX is an emotion (e.g., <i>[being] happy is an emotion</i> )	XX trait	XX is a trait
XX as an emotion	XX behavior	XX is a behavior
XX feeling	XX behaviour	XX is a behaviour
XX is a feeling	XX skill	XX is a skill
XX as a feeling	XX skills	admire XX
XX mood	criticise XX	criticize XX
XX is a mood	praise XX	condemn XX
XX as a mood	to sanction XX	to esteem XX

TABLE 2 – Patterns for affect and judgment adjectives.

The schematic representation of the algorithm for PMI calculation is shown in Fig. 1. As a search engine, we selected BING (<http://www.bing.com/>). In our work, each BING query is submitted through BING search API (<http://www.bing.com/toolbox/bingdeveloper/>) using the following structure ensuring retrieval of exact phrases in web documents written in English:

[http://api.search.live.net/xml.aspx?Appid=\[application\\_id\]&sources=web&query=inbody:\["word\\_or\\_phrase"\]language=en](http://api.search.live.net/xml.aspx?Appid=[application_id]&sources=web&query=inbody:[)

The total number of the returned query results (that is the number of hits) is then retrieved from the downloaded XML file.

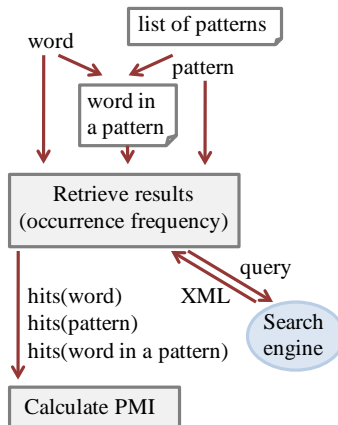


FIGURE 1 – Working flow of the PMI calculation algorithm.

There are four groups of PMI based features employed in our experiments:

1. *PMI*: PMI of an adjective with each affect pattern and each judgment pattern (in total, 30 features).
2. *maxPMI*: maximum PMI with affect patterns and maximum PMI with judgment patterns (2 features).
3. *avgPMI*: average PMI with affect patterns and average PMI with judgment patterns (2 features).
4. *%undefPMI*: percent of "undefined" PMI with affect patterns and percent of "undefined" PMI with judgment patterns (2 features). PMI with a particular pattern is "undefined" in case the search engine returns 0 for number of hits of a word in this pattern (i.e., PMI equals negative infinity).

### 2.2.2 Word-pattern co-occurrence (WPC) based feature set

In addition to PMI based features, we considered the following four co-occurrence based features (*max%rate*):

1. maximum percent rate of *hits(word in a pattern)* to *hits(pattern)* among affect patterns.
2. maximum percent rate of *hits(word in a pattern)* to *hits(pattern)* among judgment patterns.
3. maximum percent rate of *hits(word in a pattern)* to *hits(word)* among affect patterns.
4. maximum percent rate of *hits(word in a pattern)* to *hits(word)* among judgment patterns.

### 2.2.3 Minimal path length (MPL) based feature set

To establish the relatedness of a given adjective with affect or judgment, we decided to employ features based on estimation of proximity between two adjectives through synonymy relation in WordNet (Miller, 1990).

We adopted the following definitions of MPL from (Kamps & Marx, 2002):

*Two words  $w_0$  and  $w_n$  are  $n$ -related if there exists an  $(n+1)$ -long sequence of words  $\langle w_0, w_1, \dots, w_n \rangle$  such that for each  $i$  from 0 to  $n-1$  the two words  $w_i$  and  $w_{i+1}$  are in the same SYNSET.*

*Let MPL be a partial function such that  $MPL(w_i, w_j) = n$  if  $n$  is the smallest number such that  $w_i$  and  $w_j$  are  $n$ -related.*

For the exploration of WordNet relations, we employed Java API for WordNet Searching (JAWS) publicly available at <http://tyle.smu.edu/~tspell/jaws>. Automatically analysing synonymy relations in WordNet, we estimate the shortest synonymy paths from a given adjective to each word from the representative lists of affect and judgment adjectives using Equation (2). These representative lists were created manually and include 25 affect adjectives (e.g., *angry, afraid, happy, downhearted, surprised*, and others) and 20 judgment adjectives (e.g., *clever, well-mannered, cynical, dishonorable*, etc.).

$$MPL(w_i, w_j) = \min (N), \quad (2)$$

where  $w_i$  stands for one of the adjectives;  $w_j$  – one of the adjectives from representative word lists for affect and judgment; and  $N$  – set of path lengths  $\{n_0, n_1, \dots, n_k\}$ , where  $n_k$  is the number of direct-synonymy links in a synonymous sequence  $k$  between words  $w_i$  and  $w_j$ .

To make the task of analysing large synonymous network in WordNet feasible, we established the maximum limit for MPL, as the relatedness between non-direct synonyms disappears quickly when the number of synonymy links grows. Therefore, if  $MPL(w_i, w_j)$  is outside the range from 0 to 4, it is considered to be >4 or undefined (no synonymy path between two words).

The feature set based on MPL contains two groups of features:

1. *MPL*: MPL between an adjective and each representative affect or judgment adjective (in total, 45 features).
2. *minMPL*: minimal MPL among MPLs between an adjective and affect adjectives and minimal MPL among MPLs between an adjective and judgment adjectives (in total, 2 features).

### 2.3 Classification algorithms

With the aim to find the best performing machine learning algorithm classifying attitude adjectives into AFF, JUD, and MISC classes, we conducted a series of experiments with the following algorithms from WEKA software (Hall, Frank, Holmes, Pfahringer, Reutemann, & Witten, 2009):

1. J48 (Decision Trees).
2. Naive Bayes (Bayesian classifier).
3. SMO (Sequential Minimal Optimization algorithm for training a support vector classifier).

As a baseline, we considered rule-based classifier ZeroR that classifies data using the most frequent label.

### 2.4 Evaluation results

We performed 10-fold cross-validations on our data set in order to get reasonable estimate of the expected accuracy on unseen adjectives.

First, we evaluated the effectiveness of distinct groups of features. The results (percents of correctly classified instances) are given in Table 3.

Groups of features	Accuracy rate (%)			
	ZeroR	J48	Naive Bayes	SMO
<i>%undefPMI</i>	38.40	<b>46.56*</b>	44.22*	45.14*
<i>maxPMI</i>		<b>49.91*</b>	47.99*	48.42*
<i>max%rate</i>		<b>52.30***</b>	36.01	38.80
<i>avgPMI</i>		51.67*	52.39*	<b>53.55*</b>
<i>minMPL</i>		<b>54.40*</b>	54.25*	54.25*
<i>MPL</i>		<b>55.06**</b>	44.13*	53.51**
<i>PMI</i>		47.68*	54.17**	<b>55.08**</b>
Best results are given in bold. * Significantly higher than the baseline. ** Significantly higher than the baseline and one of the other methods. *** Significantly higher than the baseline and two other methods.				

TABLE 3 – Classification results using distinct groups of features.

Paired t-tests with significance level of 0.05 showed that all ML algorithms (J48, Naive Bayes, and SMO) employing distinct groups of features outperformed the baseline method with statistically significant difference in accuracy rate, with the exceptional cases of Naive Bayes and SMO using *max%rate* features. As seen from the obtained results, algorithms based on the decision trees (J48) and support vectors (SMO) overall resulted in higher accuracy than Naive Bayes classifier. *PMI* and *MPL* features proved to be more effective than other features, when employed independently in SMO and J48 algorithms, respectively.

In our next experiment, to analyse the importance of different groups of features, first we evaluated the performance of the classification algorithms with *PMI* features only, then we cumulatively added other features to the algorithms. The results in terms of accuracy rate at each step of this experiment are given in Table 4 for each classification algorithm.

Features	Accuracy rate (%)			
	ZeroR	J48	Naive Bayes	SMO
<i>PMI</i>	38.40	47.68*	54.17**	<b>55.08**</b>
<i>PMI + maxPMI</i>		50.54*	54.29**	<b>55.40**</b>
<i>PMI + maxPMI + avgPMI</i>		51.17*	55.16**	<b>56.85**</b>
<i>PMI + maxPMI + avgPMI + %undefPMI</i>		50.50*	54.37**	<b>57.61***</b>
<i>PMI + maxPMI + avgPMI + %undefPMI + max%rate</i>		52.74*	50.79*	<b>57.77***</b>
<i>PMI + maxPMI + avgPMI + %undefPMI + max%rate + MPL</i>		57.64*	54.78*	<b>61.88***</b>
<i>PMI + maxPMI + avgPMI + %undefPMI + max%rate + MPL + minMPL</i>		58.47*	57.15*	<b>61.81***</b>
Best results are given in bold.				
* Significantly higher than the baseline.				
** Significantly higher than the baseline and one of the other methods.				
*** Significantly higher than the baseline and two other methods.				

TABLE 4 – Classification results based on features cumulatively added to the algorithms.

The evaluation revealed that the support vector classifier SMO significantly outperformed other methods at each step of the experiment, with only statistically insignificant difference in case of comparison to Naive Bayes at first three steps. As was expected, the obtained results indicate that the classification algorithm benefits from consideration of all groups of features. The analysis of results from the best-performing algorithm (SMO) shows that adding *PMI* based features, such as *maxPMI*, *avgPMI*, and *%undefPMI*, to *PMI* features allows obtaining 2.53% gain in accuracy. Insignificant improvement is observed after inclusion of WPC based features (namely, *max%rate*), and this is not surprising, as these features proved to be ineffective when independently employed in SMO (i.e., there is almost no improvement over the baseline, as seen in Table 3). Statistically significant gain in accuracy is obtained after inclusion of *MPL* based features (namely, *MPL* and *minMPL*). It is important to note, however, that the performance of SMO classifier does not benefit from *minMPL* features, in contrast to J48 and Naive Bayes classifiers.

The detailed accuracy of SMO with full set of features by class (AFF, JUD, and MISC) in terms of precision, recall, and F-measure is given in Table 5.



Class	Detailed accuracy of SMO		
	Precision	Recall	F-measure
AFF	0.748	0.594	0.662
JUD	0.594	0.551	0.571
MISC	0.558	0.689	0.617

TABLE 5 – Detailed accuracy of SMO with full set of features.

The classifier achieved the highest level of precision in classifying adjectives related to AFF (0.748), while it was least precise in case of MISC (0.558) adjectives. F-measures indicate that it is easier for SMO algorithm to classify AFF adjectives than MISC and JUD adjectives.

The confusion matrix (Table 6) shows that AFF and JUD adjectives were predominantly incorrectly predicted as MISC adjectives, while MISC adjectives were mostly confused with JUD ones. This is due to the fact that the MISC class in the data set includes adjectives that are annotated by multiple labels (AFF-APP, AFF-JUD, JUD-APP, AFF-JUD-APP) and may express affect or judgment depending on the context. Interesting observation is that AFF and JUD adjectives were rarely confused: only 10% of AFF adjectives were incorrectly labeled as JUD, while about 6.8% of JUD adjectives were confused with AFF ones), thus demonstrating that PMI and MPL based features proposed in our work are good enough in characterizing these categories of adjectives.

Class	AFF	JUD	MISC
AFF	303	51	156
JUD	28	228	158
MISC	74	105	397

TABLE 6 – Confusion matrix.

### 3 Estimation of appraisal potential

The next stage of the proposed method is to estimate the potentials of MISC adjectives to express affect, judgment, and appreciation. The schematic representation of the algorithm for appraisal potential estimation is shown in Fig. 2.

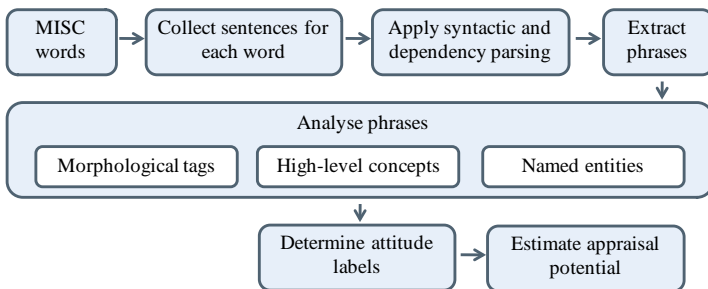


FIGURE 2 – Working flow of the algorithm for appraisal potential estimation.

The algorithm starts with the collection of sentences for each MISC word from online ABBYY Lingvo.Pro dictionary (<http://lingvopro.abbyyonline.com/en>). This dictionary allows access to unique online storage of sentences taken from real texts of different genres and language styles (classic and modern literature, web sites, technical publications, and legal documents) with the purpose to demonstrate typical use of a word. To restrict the number of example sentences extracted for each MISC adjective, the upper limit was set to 75 sentences.

Given 576 MISC adjectives, the algorithm collected 16217 sentences. About 78% of all MISC adjectives were productive, resulting in at least one example sentence. The average number of sentences per productive word is about 36. The percent distribution of productive words is as follows: low-productive adjectives (from 1 to 25 sentences) – 51.1%, including *truthful*, *inhumane*; medium-productive adjectives (from 26 to 50 sentences) – 11.3%, including *gorgeous*, *irrational*; and highly productive adjectives (from 51 to 75 sentences) – 37.6%, including *successful*, *difficult* etc. The analysis of non-productive adjectives (for example, *glamorous*, *ill-proportioned*, *uninspiring*) that did not yield any example sentence revealed that about 57% of them are hyphenated compound adjectives (for comparison, such adjectives occur only in 13% of productive ones). To collect example sentences for MISC adjectives that turned out non-productive in online ABBYY Lingvo.Pro dictionary, the algorithm may employ other online sources (for example, news, forums, blogs etc.); however, this is out of scope of this work.

Then, Connexor Machine Syntax parser (Connexor Oy, <http://www.connexor.eu/technology/machine/machinesyntax/>) is applied to each sentence in order to get lemmas, syntactic relations, dependencies, syntactic and morphological information.

Using the parser output, the method then extracts phrases that include the corresponding adjective. Some examples of sentences that contain MISC adjective *beautiful* are demonstrated in Table 7.

Sentence	Phrase	Annotations	Attitude label
Thus all my <i>beautiful</i> feelings ended in smoke.*	my <i>beautiful</i> feelings	my [PRON PERS GEN SG1] <i>beautiful</i> feelings [N NOM PL] [FEELING]	AFF
She helped him to get well, and he fell madly in love with the <i>beautiful</i> young Indian and married her.**	<i>beautiful</i> young Indian	<i>beautiful</i> young Indian [N NOM SG] [PERSON]	JUD
'He apologizes for any inconvenience and hopes you will enjoy your stay in his <i>beautiful</i> city,' said Inigo.***	his <i>beautiful</i> city	his [PRON PERS GEN SG3] <i>beautiful</i> city [N NOM SG] [LOCATION]	APP
* Youth. Tolstoy, Leo. ** The Fire From Within. Castaneda, Carlos. *** Fifth Elephant. Pratchett, Terry.			

TABLE 7 – Analysis of sentences that contain MISC adjective *beautiful*.

Three types of annotations are considered in the stage of phrase analysis (example annotations are given in Table 7):

1. morphological tags.
2. high-level concepts.
3. named entities.

Morphological tags of our particular interest that are taken from the output of Connexor Machine Syntax parser are related to pronouns and nouns. They include N (noun), PRON (pronoun), PERS (personal), NOM (nominative), GEN (genitive), ACC (accusative), SG1/PL1 (singular/plural, first person), SG3/PL3 (singular/plural, third person), <Refl> (reflexive), <Rel> (relative), <Interr> (interrogative), and WH (wh-pronoun).

In addition to morphological tags, high-level concepts of nouns are automatically extracted from WordNet based on the analysis of bottom-up sequence of hypernymic semantic relations. The hierarchy of high-level concepts used in our approach is given in Table 8. For example, *musician* is related to high-level concept *PERSON*, *virtuosity* – to *SKILL*, and *contest* – to *EVENT*.

<i>ENTITY</i>	
<i>1. ABSTRACTION</i> <i>ATTRIBUTE</i> <i>PERSONALITY</i> <i>SHAPE</i> <i>SKILLFULNESS</i> <i>TRAIT</i> <i>SELF-POSSESSION</i> <i>COMMUNICATION</i> <i>FEELING</i> <i>GROUP</i> <i>ETHNIC GROUP</i> <i>PEOPLE</i> <i>SOCIAL GROUP</i> <i>PSYCHOLOGICAL FEATURE</i> <i>COGNITION</i> <i>ATTITUDE</i> <i>BELIEF, incl. OPINION, JUDGMENT</i> <i>MIND</i> <i>SKILL</i> <i>MOTIVATION, incl. ETHICAL MOTIVE</i> <i>QUANTITY</i> <i>RELATION</i> <i>TIME</i>	<i>2. ACTIVITY</i> <i>3. BODY</i> <i>4. EVENT</i> <i>5. FOOD</i> <i>6. LOCATION</i> <i>7. OBJECT</i> <i>ARTIFACT</i> <i>NATURAL OBJECT</i> <i>8. ORGANISM</i> <i>ANIMAL</i> <i>HUMAN</i> <i>PERSON</i> <i>MAN</i> <i>RELATIVE</i> <i>9. PLANT</i> <i>10. POSSESSION</i> <i>11. PROCESS</i> <i>NATURAL PHENOMENON</i> <i>12. STATE</i> <i>13. SUBSTANCE</i>

TABLE 8 – The hierarchy of high-level concepts.

For further annotations the algorithm employs Stanford Named Entity Recognizer (Finkel, Grenager, & Manning, 2005) to detect named entities related to *PERSON*, *ORGANIZATION*, and *LOCATION*.

Next stage is to determine attitude label for the MISC adjective depending on phrase context. The algorithm (1) analyses the morphological tags, high-level concepts, and named entities in the phrase, (2) applies rules depending on these features, and (3) makes decision about attitude label. For example, *beautiful* expresses affect in the context of *my beautiful feelings*, judgment in case of *beautiful young Indian*, and appreciation in the phrase *his beautiful city*.

The attitude label rules were developed in accordance with the definitions of affect, judgment, and appreciation given in the Appraisal Theory by (Martin & White, 2005).

- Affect is a personal emotional state, feeling, or reaction to behaviour, process, or phenomena.
- Judgment is an ethical appraisal of person's character, behaviour, skills etc. according to various normative principles.
- Appreciation is an aesthetic evaluation of semiotic and natural phenomena, events, objects etc.

The features related to AFF, JUD and APP are listed below (note that some features are common for both AFF and JUD).

- AFF: nominal head of a phrase, or subject (where adjective functions as a subject complement), or object (where adjective functions as an object complement) is
  - nominative first person pronoun (*I, we*), second person pronoun (*you*), or third person pronoun (*he, she*);
  - accusative first person pronoun (*me, us*), second person pronoun (*you*), or third person pronoun (*him, them*);
  - reflexive first person pronoun (*myself, ourselves*), second person pronoun (*yourself*), or third person pronoun (*herself, himself*);
  - relative wh-pronoun (*who, whoever, whom, whomever*);
  - named entity (nominative) labelled as *PERSON*;
  - one of high-level concepts: *FEELING, PERSON, MAN, HUMAN, RELATIVE, PEOPLE, ETHNIC GROUP*, or *SOCIAL GROUP*;
  - high-level concept *ACTIVITY* pre-modified by genitive first person pronoun (for example, *my useless attempts*).

Examples of sentences, where MISC adjectives (underlined> are related to affect, include:

*It was a beneficent pause, relaxed, and filled with {peaceful satisfaction [N NOM SG] [FEELING]} in respect of work already accomplished.* (The Magic Mountain. Mann, Thomas).

*Again was all {my [PRON PERS GEN SG1] arduous labor [N NOM SG] [ACTIVITY]} gone for naught.* (The Warlord of Mars. Burroughs, Edgar Rice).

- JUD: head of a phrase, or subject (where adjective functions as a subject complement), or object (where adjective functions as an object complement) is
  - nominative first person pronoun, second person pronoun, or third person pronoun;
  - accusative first person pronoun, second person pronoun, or third person pronoun;
  - reflexive first person pronoun, second person pronoun, or third person pronoun;
  - relative wh-pronoun;
  - named entity (nominative) labelled as *PERSON* or *ORGANIZATION*;
  - one of high-level concepts: *ATTITUDE, BELIEF, MIND, MOTIVATION,*

*PERSONALITY, SELF-POSSESSION, SKILL, SKILLFULNESS, TRAIT, PERSON, MAN, HUMAN, RELATIVE, PEOPLE, ETHNIC GROUP, SOCIAL GROUP;*

- high-level concept *ACTIVITY*
  - (1) pre-modified by genitive second person pronoun (*your*), genitive third person pronoun (*his*), genitive wh-pronoun (*whose*), genitive named entity labelled as *PERSON* (for example, *John's*) or *ORGANIZATION*, or genitive noun related to one of high-level concepts: *PERSON* (for example, *doctor's*), *MAN, HUMAN, RELATIVE, PEOPLE, ETHNIC GROUP, SOCIAL GROUP*, or
  - (2) post-modified by phrase beginning with *of*, where prepositional complement is represented by one of named entities or high-level concepts mentioned above.

For instance, *His acting was perfect* and *Doctor's assistance was productive* convey inscribed JUD and invoked APP, as a person is explicitly mentioned in both sentences.

Examples of sentences, where MISC adjectives (underlined) are related to judgment, include:

*She has {fantastic organizational skills [N NOM PL] [SKILL]} that have been a tremendous help in managing all the information that comes into and goes out of this office. (Upgrading and Repairing Laptops. Mueller, Scott).*

*{Russia's [N GEN SG] [LOCATION] exalted view [N NOM SG] [ATTITUDE] of itself} was rarely shared by the outside world. (Diplomacy. Kissinger, Henry).*

- APP: head of a phrase, or subject (where adjective functions as a subject complement), or object (where adjective functions as an object complement) is
  - named entity labelled as *LOCATION*;
  - one of high-level concepts: *ABSTRACTION, ANIMAL, ARTIFACT, ATTRIBUTE, BODY, COGNITION, COMMUNICATION, ENTITY, EVENT, FOOD, GROUP, LOCATION, NATURAL OBJECT, NATURAL PHENOMENON, OBJECT, ORGANISM, PLANT, POSSESSION, PROCESS, PSYCHOLOGICAL FEATURE, QUANTITY, RELATION, SHAPE, STATE, SUBSTANCE, TIME*;
  - high-level concept *ACTIVITY* used without explicit mention of a person (for example, *successful filtration* is a natural process (APP); the sentence *It was responsible innings* conveys inscribed APP and invoked JUD, as the person is not mentioned explicitly).

Examples of sentences, where MISC adjectives (underlined) are related to appreciation, include:

*The Advisory Committee found {the presentation [N NOM SG] [ACTIVITY] lengthy and cumbersome}, particularly in the addendum to the report. (United Nations 2010).*

*He seemed to be sitting in {a very uncomfortable pram [N NOM SG] [ARTIFACT]}, with some strange insects buzzing around him. (Reaper Man. Pratchett, Terry).*

After all collected sentences were labeled by attitude types, the appraisal potentials of productive MISC adjectives were estimated. The potentials of a word to express affect, judgment, and appreciation were calculated based on the number of sentences related to each type of attitude using Equations (3)-(5).

$$Affect\ Potential\ (word) = \frac{N_{aff}(word)}{N_{aff}(word) + N_{jud}(word) + N_{app}(word)} \quad (3)$$

$$\text{Judgment Potential (word)} = \frac{N_{jud}(\text{word})}{N_{aff}(\text{word}) + N_{jud}(\text{word}) + N_{app}(\text{word})} \quad (4)$$

$$\text{Appreciation Potential (word)} = \frac{N_{app}(\text{word})}{N_{aff}(\text{word}) + N_{jud}(\text{word}) + N_{app}(\text{word})}, \quad (5)$$

where *word* stands for an adjective;  $N_{aff}$ ,  $N_{jud}$ , and  $N_{app}$  – number of sentences, where *word* conveys affect, judgment, and appreciation, correspondingly.

The examples of appraisal potentials calculated for adjectives are given in Table 9.

Adjective	Affect Potential	Judgment Potential	Appreciation Potential
appealing	0.15	0.22	0.63
awkward	0.29	0.31	0.40
bashful	0.38	0.44	0.18
consummate	0.25	0.58	0.17
excellent	0.19	0.22	0.59
genuine	0.32	0.16	0.52
jealous	0.46	0.44	0.10
loving	0.41	0.31	0.28
tasty	0.0	0.0	1.0
unsuitable	0.0	0.05	0.95
upbeat	0.5	0.33	0.17

TABLE 9 – Appraisal potentials.

## Conclusions

In this paper, we proposed a novel method for analysing sentiment word relations with three attitude types, namely affect, judgment, and appreciation. We emphasized the importance of recognition of context-dependent attitudes conveyed by adjectives of ambiguous attitude type. With the aim to find the best performing machine learning algorithm classifying attitude adjectives into affect, judgment, and miscellaneous classes, we created a dataset (1500 attitude-annotated adjectives) and conducted a series of experiments with the following algorithms: Decision Trees, Naive Bayes, and Support Vector classifier. In our experiments we employed three feature sets comprising of 88 features. The evaluation revealed that the classification algorithms benefited from consideration of all groups of features, and the Support Vector classifier significantly outperformed other algorithms (with about 62% accuracy). The classifier achieved the highest level of precision in classifying adjectives related to affect (0.748), while it was least precise in case of miscellaneous (0.558) adjectives. The appraisal potentials of miscellaneous adjectives to convey affect, judgment, and appreciation were estimated based on a novel algorithm analysing contextual attitudes expressed by each word in a set of sentences.

## Acknowledgments

This work was supported by a Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (JSPS) through the program for JSPS Postdoctoral Fellowship for Foreign Researchers.

## References

- Andreevskaia, A. and Bergler, S. (2006). Mining WordNet for fuzzy sentiment: sentiment tag extraction from WordNet glosses. In *Eleventh Conference of the European Chapter of the Association for Computational Linguistics*, pages 209–216.
- Argamon, S., Bloom, K., Esuli, A., and Sebastiani, F. (2007). Automatically determining attitude type and force for sentiment analysis. In *Third Language and Technology Conference*.
- Banea, C., Mihalcea, R., and Wiebe, J. (2008). A bootstrapping method for building subjectivity lexicons for languages with scarce resources. In *International Conference on Language Resources and Evaluations (LREC 2008)*.
- Baroni, M. and Vegnaduzzo, S. (2004). Identifying subjective adjectives through web-based mutual information. In *Seventh German Conference on Natural Language Processing*.
- Church, K. W. and Hanks, P. (1989). Word association norms, mutual information and lexicography. In *27th Annual Conference of the Association of Computational Linguistics*, pages 76–83, New Brunswick, NJ: Association for Computational Linguistics.
- De Albornoz, J. C., Plaza, L., and Gervás, P. (2012). SentiSense: An easily scalable concept-based affective lexicon for sentiment analysis. In *Eighth International Conference on Language Resources and Evaluation (LREC 2012)*.
- Esuli, A. and Sebastiani, F. (2006). SentiWordNet: a publicly available lexical resource for opinion mining. In *Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, pages 417–422, Genoa, Italy.
- Finkel, J. R., Grenager, T., and Manning, C. (2005). Incorporating non-local information into information extraction systems by Gibbs sampling. In *43rd Annual Meeting of the Association of Computational Linguistics*, pages 363–370.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA data mining software: An update. *SIGKDD Explorations*, 11(1).
- Hatzivassiloglou, V. and McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. In *35th Annual Meeting of the Association for Computational Linguistics and the 8th Conference of the European Chapter of the ACL*, pages 174–181.
- Kamps, J. and Marx, M. (2002). Words with attitude. In *Belgian-Dutch Conference on Artificial Intelligence*, pages 449–450.
- Kim, S.-M. and Hovy, E. (2004). Determining the sentiment of opinions. In *International Conference on Computational Linguistics (COLING 2004)*, pages 1367–1373.
- Ku, L.-W., Huang, T.-H., and Chen, H.-H. (2009). Using morphological and syntactic structures for Chinese opinion analysis. In *International Conference on Empirical Methods in Natural Language Processing*, pages 1260–1269.
- Martin, J. R. and White, P. R. R. (2005). *The Language of Evaluation: Appraisal in English*. Palgrave, London, UK.
- Miller, G. A. (1990). WordNet: An on-line lexical database. *International Journal of Lexicography*, Special Issue, 3(4):235–312.
- Moilanen, K. and Pulman, S. (2008). The good, the bad, and the unknown: morphosyllabic

- sentiment tagging of unseen words. In *Proceedings of the ACL-08: HLT*, pages 109–112.
- Neviarouskaya, A. (2011). *Compositional Approach for Automatic Recognition of Fine-Grained Affect, Judgment, and Appreciation in Text*. PhD Dissertation, Graduate School of Information Science and Technology, University of Tokyo.
- Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2011). SentiFul: A lexicon for sentiment analysis. *IEEE Transactions on Affective Computing*, 2(1):22–36.
- Riloff, E., Wiebe, J., and Wilson, T. (2003). Learning subjective nouns using extraction pattern bootstrapping. In *Conference on Natural Language Learning*, pages 25–32.
- Strapparava, C. and Valitutti, A. (2004). WordNet-Affect: an affective extension of WordNet. In *International Conference on Language Resources and Evaluation*, pages 1083–1086.
- Su, F. and Markert, K. (2009). Subjectivity recognition on word senses via semi-supervised mincuts. In *North American Association of Computational Linguistics (NAACL 2009)*.
- Taboada, M. and Grieve, J. (2004). Analyzing appraisal automatically. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text*, pages 158–161.
- Takamura, H., Inui, T., and Okumura, M. (2005). Extracting semantic orientation of words using spin model. In *43rd Annual Meeting of the Association of Computational Linguistics*, pages 133–140.
- Turney, P. and Littman, M. (2002). Unsupervised learning of semantic orientation from a hundred-billion-word corpus. *Technical Report ERC-1094 (NRC 44929)*, National Research Council of Canada.
- Wiebe, J. (2000). Learning subjective adjectives from corpora. In *17th Conference of the AAI*.
- Wiebe, J. and Mihalcea, R. (2006). Word sense and subjectivity. In *Joint Conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics (COLING-ACL 2006)*.
- Wilson, T., Wiebe, J., and Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 347–354, Vancouver, Canada: Association for Computational Linguistics.