Recognition of Affect, Judgment, and Appreciation in Text

Alena Neviarouskaya
University of Tokyo
`lena@mi.ci.i.u-tokyo.ac.jp`

Helmut Prendinger
Nat. Institute of Informatics Tokyo
`helmut@nii.ac.jp`

Mitsuru Ishizuka
University of Tokyo
`ishizuka@i.u-tokyo.ac.jp`

Abstract

The main task we address in our research is classification of text using fine-grained attitude labels. The developed @AM system relies on the compositionality principle and a novel approach based on the rules elaborated for semantically distinct verb classes. The evaluation of our method on 1000 sentences, that describe personal experiences, showed promising results: average accuracy on the fine-grained level (14 labels) was 62%, on the middle level (7 labels) – 71%, and on the top level (3 labels) – 88%.

1 Introduction and Related Work

With rapidly growing online sources aimed at encouraging and stimulating people's discussions concerning personal, public or social issues (news, blogs, discussion forums, etc.), there is a great need in development of a computational tool for the analysis of people's attitudes. According to the Appraisal Theory (Martin and White, 2005), attitude types define the specifics of appraisal being expressed: affect (personal emotional state), judgment (social or ethical appraisal of other's behaviour), and appreciation (evaluation of phenomena).

To analyse contextual sentiment of a phrase or a sentence, rule-based approaches (Nasukawa and Yi, 2003; Moilanen and Pulman, 2007; Subrahmanian and Reforgiato, 2008), a machine-learning method using not only lexical but also syntactic features (Wilson et al., 2005), and a model of integration of machine learning approach with compositional semantics (Choi and Cardie, 2008) were proposed. With the aim to recognize fine-grained emotions from text on the level of distinct sentences, researchers have employed a keyword spotting technique (Chuang and Wu, 2004; Strapparava et al., 2007), a technique calculating emotion scores using Pointwise Mutual Information (PMI) (Kozareva et al., 2007), an approach inspired by common-sense knowledge (Liu et al., 2003), rule-based linguistic approaches (Boucouvalas, 2003; Chaumartin, 2007), machine-learning methods (Alm, 2008; Aman and Szpakowicz, 2008; Strapparava and Mihalcea, 2008), and an ensemble based multi-label classification technique (Bhowmick et al., 2009).

Early attempts to focus on distinct attitude types in the task of attitude analysis were made by Taboada and Grieve (2004), who determined a potential value of adjectives for affect, judgement and appreciation by calculating the PMI with the pronoun-copular pairs '*I was (affect)*', '*He was (judgement)*', and '*It was (appreciation)*', and Whitelaw et al. (2005), who used a machine learning technique (SVM) with fine-grained semantic distinctions in features (attitude type, orientation) in combination with "bag of words" to classify movie reviews. However, the concentration only on adjectives expressing appraisal and their modifiers greatly narrows the potential of the Whitelaw et al. (2005) approach.

In this paper we introduce our system @AM (**AT**titude **A**nalysis **M**odel), which (1) classifies sentences according to the fine-grained attitude labels (nine affect categories (Izard, 1971): 'anger', 'disgust', 'fear', 'guilt', 'interest', 'joy', 'sadness', 'shame', 'surprise'; four polarity labels for judgment and appreciation: 'POS jud', 'NEG jud', 'POS app', 'NEG app'; and 'neutral'); (2) assigns the strength of the attitude; and (3) determines the level of confidence, with which the attitude is expressed. @AM relies on a compositionality principle and a novel approach

based on the rules elaborated for semantically distinct verb classes.

2 Lexicon for Attitide Analysis

We built a lexicon for attitude analysis that includes: (1) attitude-conveying terms; (2) modifiers; (3) "functional" words; and (4) modal operators.

2.1 The Core of Lexicon

As a core of lexicon for attitude analysis, we employ an Affect database and extended version of the SentiFul database developed by Neviarouskaya et al. (2009). The affective features of each emotion-related word are encoded using nine emotion labels ('anger', 'disgust', 'fear', 'guilt', 'interest', 'joy', 'sadness', 'shame', and 'surprise') and corresponding emotion intensities that range from 0.0 to 1.0. The original version of SentiFul database, which contains sentiment-conveying adjectives, adverbs, nouns, and verbs annotated by sentiment polarity, polarity scores and weights, was manually extended using attitude labels. Some examples of annotated attitude-conveying words are listed in Table 1. It is important to note here that some words may express different attitude types (affect, judgment, appreciation) depending on context; such lexical entries were annotated by all possible categories.

POS	Word	Category	Intensity
adjective	*honorable*	POS jud	0.3
	unfriendly	NEG aff (sadness)	0.5
		NEG jud	0.5
		NEG app	0.5
adverb	*gleefully*	POS aff (joy)	0.9
noun	*abnormality*	NEG app	0.25
verb	*frighten*	NEG aff (fear)	0.8
	desire	POS aff (interest)	1.0
		POS aff (joy)	0.5

Table 1. Examples of attitude-conveying words and their annotations.

2.2 Modifiers and Functional Words

We collected 138 modifiers that have an impact on contextual attitude features of related words, phrases, or clauses. They include:

1. Adverbs of degree (e.g., '*significantly*', '*slightly*' etc.) and affirmation (e.g., '*absolutely*', '*seemingly*') that have an influence on the strength of the attitude of related words. Two annotators gave coefficients for intensity degree strengthening or weakening (from 0.0 to 2.0) to each adverb, and the result was averaged (e.g., coeff('*slightly*') = 0.2).

2. Negation words (e.g., '*never*', '*nothing*' etc.) reversing the polarity of related statement.

3. Adverbs of doubt (e.g., '*scarcely*', '*hardly*' etc.) and falseness (e.g., '*wrongly*' etc.) reversing the polarity of related statement.

4. Prepositions (e.g., '*without*', '*despite*' etc.) neutralizing the attitude of related words.

5. Condition operators (e.g., '*if*', '*even though*' etc.) that neutralize the attitude of related words.

We distinguish two types of "functional" words that influence contextual attitude and its strength:

1. *Intensifying* adjectives (e.g., '*rising*', '*rapidly-growing*'), nouns (e.g., '*increase*'), and verbs (e.g., '*to grow*', '*to rocket*') that increase the strength of attitude of related words.

2. *Reversing* adjectives (e.g., '*reduced*'), nouns (e.g., '*termination*'), and verbs (e.g., '*to decrease*', '*to limit*', '*to diminish*'), which reverse the prior polarity of related words.

2.3 Modal Operators

Consideration of the modal operators in the tasks of opinion mining and attitude analysis is very important, as they indicate a degree of person's belief in the truth of the proposition, which is subjective in nature (Hoye, 1997). Modals are distinguished by their *confidence level*. We collected modal operators of two categories: modal verbs (13 verbs) and modal adverbs (61 adverbs). Three human annotators assigned the *confidence level* ranging from 0.0 to 1.0 to each modal verb and adverb; these ratings were averaged (e.g., conf('*vaguely*') = 0.17, conf('*arguably*') = 0.63, conf('*would*') = 0.8, conf('*veritably*') = 1.0).

3 Compositionality Principle

Our algorithm for attitude classification is designed based on the *compositionality principle*, according to which we determine the attitudinal meaning of a sentence by composing the pieces that correspond to lexical units or other linguistic constituent types governed by the rules of *polarity reversal, aggregation (fusion), propagation, domination, neutralization,* and *intensification,* at various grammatical levels.

Polarity reversal means that a phrase or statement containing an attitude-conveying

term/phrase with prior positive polarity becomes negative, and vice versa. The rule of *polarity reversal* is applied in three cases: (1) negation word-modifier in relation with an attitude-conveying statement (e.g., '*never*' & POS('*succeed*') => NEG('*never succeed*')); (2) adverb of doubt in relation with attitude-conveying statement (e.g., '*scarcely*' & POS('*relax*') => NEG('*scarcely relax*')); (3) functional word of *reversing* type in relation with attitude-conveying statement (e.g., adjective '*reduced*' & POS('*enthusiasm*') => NEG('*reduced enthusiasm*')). In the case of judgment and appreciation, the use of the *polarity reversal* rule is straightforward ('POS jud' <=> 'NEG jud', 'POS app' <=> 'NEG app'). However, it is not trivial to find pairs of opposite emotions in the case of a fine-grained classification, except for 'joy' and 'sadness'. Therefore, we assume that (1) the opposite emotion for three positive emotions, i.e. 'interest', 'joy', and 'surprise', is 'sadness' ('POS aff' => 'sadness'); and (2) the opposite emotion for six negative emotions, i.e. 'anger', 'disgust', 'fear', 'guilt', 'sadness', and 'shame', is 'joy' ('NEG aff' => 'joy').

The rules of *aggregation (fusion)* are as follows: (1) if polarities of attitude-conveying terms in adjective-noun, noun-noun, adverb-adjective, adverb-verb phrases have opposite directions, mixed polarity with dominant polarity of a pre-modifier is assigned to the phrase (e.g., POS('*beautiful*') & NEG('*fight*') => POS-neg('*beautiful fight*'); NEG('*shamelessly*') & POS('*celebrate*') => NEG-pos('*shamelessly celebrate*')); otherwise (2) the resulting polarity is based on the equal polarities of terms, and the strength of attitude is measured as a maximum between polarity scores (intensities) of terms (*max*(score1,score2)).

The rule of *propagation* is useful, as proposed in (Nasukawa and Yi, 2003), for the task of the detection of local sentiments for given subjects. "Propagation" verbs propagate the sentiment towards the arguments; "transfer" verbs transmit sentiments among the arguments. The rule of *propagation* is applied when a verb of "propagation" or "transfer" type is used in a phrase/clause and sentiment of an argument that has prior neutral polarity needs to be investigated (e.g., PROP-POS('*to admire*') & '*his behaviour*' => POS('*his behaviour*'); '*Mr. X*' &

TRANS('*supports*') & NEG('*crime business*') => NEG('*Mr. X*')).

The rules of *domination* are as follows: (1) if polarities of a verb (this rule is applied only for certain classes of verbs) and an object in a clause have opposite directions, the polarity of verb is prevailing (e.g., NEG('*to deceive*') & POS('*hopes*') => NEG('*to deceive hopes*')); (2) if compound sentence joints clauses using coordinate connector '*but*', the attitude features of a clause following after the connector are dominant (e.g., 'NEG(*It was hard to climb a mountain all night long*), *but* POS(*a magnificent view rewarded the traveler at the morning*).' => POS(whole sentence)).

The rule of *neutralization* is applied when preposition-modifier or condition operator relate to the attitude-conveying statement (e.g., '*despite*' & NEG('*worries*') => NEUT('*despite worries*')).

The rule of *intensification* means strengthening or weakening of the polarity score (intensity), and is applied when:

1. adverb of degree or affirmation relates to attitude-conveying term (e.g., Pos_score('*happy*') < Pos_score('*extremely happy*'));

2. adjective or adverb is used in a comparative or superlative form (e.g., Neg_score('*sad*') < Neg_score('*sadder*') < Neg_score ('*saddest*')).

Our method is capable of processing sentences of different complexity, including simple, compound, complex (with complement and relative clauses), and complex-compound sentences. We employ Connexor Machinese Syntax parser (http://www.connexor.eu/) that returns lemmas, parts of speech, dependency functions, syntactic function tags, and morphological tags. When handling the parser output, we represent the sentence as a set of primitive clauses. Each clause might include Subject formation, Verb formation and Object formation, each of which may consist of a main element (subject, verb, or object) and its attributives and complements. For the processing of complex or compound sentences, we build a so-called "relation matrix", which contains information about dependences (e.g., coordination, subordination, condition, contingency, etc.) between different clauses in a sentence. While applying the compositionality principle, we consecutively assign attitude fea-

tures to words, phrases, formations, clauses, and finally, to the whole sentence.

4 Consideration of the Semantics of Verbs

All sentences must include a verb, because the verb tells us what action the subject is performing and object is receiving. In order to elaborate rules for attitude analysis based on the semantics of verbs, we investigated VerbNet (Kipper et al., 2007), the largest on-line verb lexicon that is organized into verb classes characterized by syntactic and semantic coherence among members of a class. Based on the thorough analysis of 270 first-level classes of VerbNet and their members, 73 verb classes (1) were found useful for the task of attitude analysis, and (2) were further classified into 22 classes differentiated by the role that members play in attitude analysis and by rules applied to them. Our classification is shown in Table 2.

For each of our verb classes, we developed set of rules that are applied to attitude analysis on the phrase/clause-level. Some verb classes (e.g., "*Psychological state or emotional reaction*", "*Judgment*", "*Bodily state and damage to the body*", "*Preservation*" etc.) include verbs annotated by attitude type, prior polarity orientation, and the strength of attitude. The attitude features of phrases that involve positively or negatively charged verbs from such classes are context-sensitive and are defined by means of rules designed for each of the class.

As an example, we provide short description and rules elaborated for the subclass "*Object-centered (oriented) emotional state*".
Features: subject experiences emotions towards some stimulus; verb prior polarity: positive or negative; context-sensitive.
Verb-Object rules (subject is ignored):
1. "Interior perspective" (subject's inner emotion state or attitude):

S & V+('*admires*') & O+('*his brave heart*') => (fusion, max(V_score,O_score)) => 'POS aff'.

S & V+('*admires*') & O-('*mafia leader*') => (verb valence dominance, V_score) => 'POS aff'.

S & V-('*disdains*') & O+('*his honesty*') => (verb valence dominance, V_score) => 'NEG aff'.

Verb class (verb samples)
1 Psychological state or emotional reaction
1.1 Object-centered (oriented) emotional state (*adore*)
1.2 Subject-driven change in emotional state (trans.) (*charm*, *inspire*, *bother*)
1.3 Subject-driven change in emotional state (intrans.) (*appeal to*, *grate on*)
2 Judgment
2.1 Positive judgment (*bless*, *honor*)
2.2 Negative judgment (*blame*, *punish*)
3 Favorable attitude (*accept*, *allow*, *tolerate*)
4 Adverse (unfavorable) attitude (*discourage*, *forbid*)
5 Favorable or adverse calibratable changes of state (*grow*, *decline*)
6 Verbs of removing
6.1 Verbs of removing with neutral charge (*delete*)
6.2 Verbs of removing with negative charge (*expel*)
6.3 Verbs of removing with positive charge (*evacuate*)
7 Negatively charged change of state (*break*, *crush*)
8 Bodily state and damage to the body (*sicken*, *injure*)
9 Aspectual verbs
9.1 Initiation, continuation of activity, and sustaining (*begin*, *continue*, *maintain*)
9.2 Termination of activity (*quit*, *finish*)
10 Preservation (*defend*, *insure*)
11 Verbs of destruction and killing (*damage*, *poison*)
12 Disappearance (*disappear*, *die*)
13 Limitation and subjugation (*confine*, *restrict*)
14 Assistance (*succor*, *help*)
15 Obtaining (*win*, *earn*)
16 Communication indicator/reinforcement of attitude (*guess*, *complain*, *deny*)
17 Verbs of leaving (*abandon*, *desert*)
18 Changes in social status or condition (*canonize*)
19 Success and failure
19.1 Success (*succeed*, *manage*)
19.2 Failure (*fail*, *flub*)
20 Emotional nonverbal expression (*smile*, *weep*)
21 Social interaction (*marry*, *divorce*)
22 Transmitting verbs (*supply*, *provide*)

Table 2. Verb classes for attitude analysis.

S & V-('*disdains*') & O-('*criminal activities*') => (fusion, max(V_score,O_score)) => 'NEG aff'.
2. "Exterior perspective" (social/ethical judgment):

S & V+('*admires*') & O+('*his brave heart*') => (fusion, max(V_score,O_score)) => 'POS jud'.

S & V+('*admires*') & O-('*mafia leader*') => (verb valence reversal, max(V_score,O_score)) => 'NEG jud'.

S & V-('*disdains*') & O+('*his honesty*') => (verb valence dominance, max(V_score,O_score)) => 'NEG jud'.

S & V-('*disdains*') & O-('*criminal activities*') => (verb valence reversal, max(V_score,O_score)) => 'POS jud'.

3. In case of neutral object => attitude type and prior polarity of verb, verb score (V_score).

Verb-PP (prepositional phrase) rules:

1. In case of negatively charged verb and PP starting with '*from*' => verb dominance:

S & V-('*suffers*') & PP-('*from illness*') => interior: 'NEG aff'; exterior: 'NEG jud'.

S & V-('*suffers*') & PP+ ('*from love*') => interior: 'NEG aff'; exterior: 'NEG jud'.

2. In case of positively charged verb and PP starting with '*in*'/'*for*' => treat PP the same way as object (see above):

S & V+('*believes*') & PP-('*in evil*') => interior: 'POS aff'; exterior: 'NEG jud'.

S & V+('*believes*') & PP+('*in kindness*') => interior: 'POS aff'; exterior: 'POS jud'.

In the majority of rules the strength of attitude is measured as a maximum between attitude scores (for example, the attitude conveyed by '*to suffer from grave illness*' is stronger than that of '*to suffer from slight illness*').

In contrast to the rules of "*Object-centered (oriented) emotional state*" subclass, which ignore attitude features of a subject in a sentence, the rules elaborated for the "*Subject-driven change in emotional state (trans.)*" disregard the attitude features of object, as in sentences involving members of this subclass object experiences emotion, and subject causes the emotional state. For example (due to limitation of space, here and below we provide only some cases):

S('*Classical music*') & V+('*calmed*') & O-('*disobedient child*') => interior: 'POS aff'; exterior: 'POS app'.

S-('*Fatal consequences of GM food intake*') & V-('*frighten*') & O('*me*') => interior: 'NEG aff'; exterior: 'NEG app'.

The Verb-Object rules for the "*Judgment*" subclasses, namely "*Positive judgment*" and "*Negative judgment*", are very close to those defined for the subclass "*Object-centered (oriented) emotional state*". However, Verb-PP rules have some specifics: for both positive and negative judgment verbs, we treat PP starting with '*for*'/'*of*'/'*as*' the same way as object in Verb-Object rules. For example:

S('*He*') & V-('*blamed*') & O+('*innocent person*') => interior: 'NEG jud'; exterior: 'NEG jud'.

S('*They*') & V-('*punished*') & O('*him*') & PP-('*for his misdeed*') => interior: 'NEG jud'; exterior: 'POS jud'.

Verbs from classes "*Favorable attitude*" and "*Adverse (unfavorable) attitude*" have prior neutral polarity and positive or negative reinforcement, correspondingly, that means that they only impact on the polarity and strength of non-neutral phrase (object in a sentence written in active voice, or subject in a sentence written in passive voice, or PP in case of some verbs). The rules are:

1. If verb belongs to the "*Favorable attitude*" class and the polarity of phrase is not neutral, then the attitude score of the phrase is intensified (symbol '^' means intensification):

S('*They*') & [V pos. reinforcement]('*elected*') & O+('*fair judge*') => 'POS app'; O_score^.

S('*They*') & [V pos. reinforcement]('*elected*') & O-('*corrupt candidate*') => 'NEG app'; O_score^.

2. If verb belongs to the "*Adverse (unfavorable) attitude*" class and the polarity of phrase is not neutral, then the polarity of phrase is reversed and score is intensified:

S('*They*') & [V neg. reinforcement]('*prevented*') & O-('*the spread of disease*') => 'POS app'; O_score^.

S+('*His achievements*') & [V neg. reinforcement]('*were overstated*') => 'NEG app'; S_score^.

Below are examples of processing the sentences with verbs from "*Verbs of removing*" class.

"*Verbs of removing with neutral charge*":

S('*The tape-recorder*') & [V neutral rem.]('*automatically ejects*') & O-neutral('*the tape*') => neutral.

S('*The safety invention*') & [V neutral rem.]('*ejected*') & O('*the pilot*') & PP-('*from burning plane*') => 'POS app'; PP_score^.

"*Verbs of removing with negative charge*":

S('*Manager*') & [V neg. rem.]('*fired*') & O-('*careless employee*') & PP('*from the company*') => 'POS app'; *max*(V_score,O_score).

"*Verbs of removing with positive charge*":

S('*They*') & [V pos. rem.]('*evacuated*') & O('*children*') & PP-('*from dangerous place*') => 'POS app'; *max*(V_score,PP_score).

Along with modal verbs and modal adverbs, members of the "*Communication indicator/reinforcement of attitude*" verb class also in-

dicate the confidence level or degree of certainty concerning given opinion. Features are: subject (communicator) expresses statement with/without attitude; statement is PP starting with '*of*', '*on*', '*against*', '*about*', '*concerning*', '*regarding*', '*that*', '*how*' etc.; ground: positive or negative; reinforcement: positive or negative. The rules are:

1. If the polarity of expressed statement is neutral, then the attitude is neutral:

S('*Professor*') & [V pos. ground, pos. reinforcement, confidence:0.83]('*dwelled*') & PP-neutral('*on a question*') => neutral.

2. If the polarity of expressed statement is not neutral and the reinforcement is positive, then the score of the statement (PP) is intensified:

S('*Jane*') & [V neg. ground, pos. reinforcement, confidence:0.8]('*is complaining*') & PP-('*of a headache again*') => 'NEG app'; PP_score^; confidence:0.8.

3. If the polarity of expressed statement is not neutral and reinforcement is negative, then the polarity of the statement (PP) is reversed and score is intensified:

S('*Max*') & [V neg. ground, neg. reinforcement, confidence:0.2]('*doubt*') & PP-{'*that*' S+('*his good fortune*') & [V termination]('*will ever end*')} => 'POS app'; PP_score^; confidence:0.2.

In the last example, to measure the sentiment of PP, we apply rule for the verb '*end*' from the "*Termination of activity*" class, which reverses the non-neutral polarity of subject (in intransitive use of verb) or object (in transitive use of verb). For example, the polarity of both sentences '*My whole enthusiasm and excitement disappear like a bubble touching a hot needle*' and '*They discontinued helping children*' is negative.

5 Decision on Attitude Label

The decision on the most appropriate final label for the clause, in case @AM annotates it using different attitude types according to the words with multiple annotations (e.g., see word '*unfriendly*' in Table 1) or based on the availability of the words conveying different attitude types, is made based on the analysis of:

1) morphological tags of nominal heads and their premodifiers in the clause (e.g., first person pronoun, third person pronoun, demonstrative pronoun, nominative or genitive noun, etc.);

2) the sequence of hypernymic semantic relations of a particular noun in WordNet (Miller, 1990), which allows to determine its conceptual domain (e.g., "person, human being", "artifact", "event", etc.);

3) the annotations from the Stanford Named Entity Recognizer (Finkel et al. 2005) that labels PERSON, ORGANIZATION, and LOCATION entities.

For ex., '*I feel highly unfriendly attitude towards me*' conveys emotion ('NEG aff': 'sadness'), while '*The shop assistant's behavior was really unfriendly*' and '*Plastic bags are environment unfriendly*' express judgment ('NEG jud') and appreciation ('NEG app'), correspondingly.

6 Evaluation

For the experiments, we used our own data set, as, to the best of our knowledge, there is no publicly available data set of sentences annotated by the fine-grained labels proposed in our work. In order to evaluate the performance of our algorithm, we created the data set of sentences extracted from personal stories about life experiences that were anonymously published on the *Experience Project* website (www.experienceproject.com), where people share personal experiences, thoughts, opinions, feelings, passions, and confessions through the network of personal stories. With over 4 million experiences accumulated (as of February 2010), *Experience Project* is a perfect source for researchers interested in studying different types of attitude expressed through text.

6.1 Data Set Description

For our experiment we extracted 1000 sentences[1] from various stories grouped by topics within 13 different categories, such as "Arts and entertainment", "Current events", "Education", "Family and friends", "Health and wellness", "Relationships and romance" and others, on the *Experience Project* website. Sentences were collected from 358 distinct topic groups, such as "I still remember September 11", "I am intelligent but airheaded", "I think bullfighting is cruel", "I quit smoking", "I am a fashion victim", "I was adopted" and others.

---

[1] This annotated data set is freely available upon request.

TOP	POS				NEG								neutral
MID	POS aff			POS jud	POS app	NEG aff						NEG jud	NEG app	neutral
ALL	interest	joy	surprise	POS jud	POS app	anger	disgust	fear	guilt	sadness	shame	NEG jud	NEG app	neutral

Figure 1. Hierarchy of attitude labels.

We considered three hierarchical levels of attitude labels in our experiment (see Figure 1). Three independent annotators labeled the sentences with one of 14 categories from the ALL level and a corresponding score (the strength or intensity value). These annotations were further interpreted using labels from the MID and the TOP levels. Fleiss' Kappa coefficient was used as a measure of reliability of human raters' annotations. The agreement coefficient on 1000 sentences was 0.53 on ALL level, 0.57 on MID level, and 0.73 on TOP level.

Only those sentences, on which at least two out of three human raters completely agreed, were included in the gold standards for our experiment. Three gold standards were created according to the hierarchy of attitude labels. Fleiss' Kappa coefficients are 0.62, 0.63, and 0.74 on ALL, MID, and TOP levels, correspondingly. Table 3 shows the distributions of labels in the gold standards.

ALL level		MID level	
Label	**Number**	**Label**	**Number**
anger	45	POS aff	233
disgust	21	NEG aff	332
fear	54	POS jud	66
guilt	22	NEG jud	78
interest	84	POS app	100
joy	95	NEG app	29
sadness	133	neutral	87
shame	18	**total**	**925**
surprise	36		
POS jud	66	**TOP level**	
NEG jud	78	**Label**	**Number**
POS app	100	POS	437
NEG app	29	NEG	473
neutral	87	neutral	87
total	**868**	**total**	**997**

Table 3. Label distributions in gold standards.

6.2 Results

The results of a simple method selecting the attitude label with the maximum intensity from the annotations of sentence tokens found in the database were considered as the baseline. After processing each sentence from the data set by the baseline method and our @AM system, we measured averaged accuracy, precision, recall, and F-score for each label in ALL, MID, and TOP levels. The results are shown in Table 4.

As seen from the obtained results, our algorithm performed with high accuracy significantly surpassing the baselines in all levels of attitude hierarchy, thus demonstrating the contribution of the sentence parsing and our hand-crafted rules to the reliable recognition of attitude from text. Two-tailed t-tests with significance level of 0.05 showed that the differences in accuracy between the baseline method and our @AM system are statistically significant ($p<0.001$) in fine-grained as well as coarse-grained classifications.

In the case of fine-grained attitude recognition (ALL level), the highest precision was obtained for 'shame' (0.923) and 'NEG jud' (0.889), while the highest recall was received for 'sadness' (0.917) and 'joy' (0.905) emotions at the cost of low precision (0.528 and 0.439, correspondingly). The algorithm performed with the worst results in recognition of 'NEG app' and 'neutral'.

The analysis of a confusion matrix for the ALL level revealed the following top confusions of our system: (1) 'anger', 'fear', 'guilt', 'shame', 'NEG jud', 'NEG app' and 'neutral' were predominantly incorrectly predicted as 'sadness' (for ex., @AM resulted in 'sadness' for the sentence '*I know we have several months left before the election, but I am already sick and tired of seeing the ads on TV*', while human annotations were 'anger'/'anger'/'disgust'); (2) 'interest', 'POS jud' and 'POS app' were mostly confused with 'joy' by our algorithm (e.g., @AM classified the sentence '*It's one of those life changing artifacts that we must have in order to have happier, healthier lives*' as 'joy'(-ful), while human annotations were 'POS app'/'POS app'/'interest').

Our system achieved high precision for all categories on the MID level (Table 4), with the exception of 'NEG app' and 'neutral', although

Level	Label	Baseline method				@AM			
		Accuracy	Precision	Recall	F-score	Accuracy	Precision	Recall	F-score
ALL	anger		0.742	0.511	0.605		0.818	0.600	0.692
	disgust		0.600	0.857	0.706		0.818	0.857	0.837
	fear		0.727	0.741	0.734		0.768	0.796	0.782
	guilt		0.667	0.364	0.471		0.833	0.455	0.588
	interest		0.380	0.357	0.368		0.772	0.524	0.624
	joy		0.266	0.579	0.364		0.439	0.905	0.591
	sadness	0.437	0.454	0.632	0.528	0.621	0.528	0.917	0.670
	shame		0.818	0.500	0.621		0.923	0.667	0.774
	surprise		0.625	0.694	0.658		0.750	0.833	0.789
	POS jud		0.429	0.227	0.297		0.824	0.424	0.560
	NEG jud		0.524	0.141	0.222		0.889	0.410	0.561
	POS app		0.349	0.150	0.210		0.755	0.400	0.523
	NEG app		0.250	0.138	0.178		0.529	0.310	0.391
	neutral		0.408	0.483	0.442		0.559	0.437	0.490
MID	POS aff		0.464	0.695	0.557		0.668	0.888	0.762
	NEG aff		0.692	0.711	0.701		0.765	0.910	0.831
	POS jud		0.405	0.227	0.291		0.800	0.424	0.554
	NEG jud	0.524	0.458	0.141	0.216	0.709	0.842	0.410	0.552
	POS app		0.333	0.150	0.207		0.741	0.400	0.519
	NEG app		0.222	0.138	0.170		0.474	0.310	0.375
	neutral		0.378	0.483	0.424		0.514	0.437	0.472
TOP	POS		0.745	0.796	0.770		0.918	0.920	0.919
	NEG	0.732	0.831	0.719	0.771	0.879	0.912	0.922	0.917
	neutral		0.347	0.483	0.404		0.469	0.437	0.452

Table 4. Results of the evaluation of performance of the baseline method and @AM system.

high recall was obtained only in the case of categories related to affect ('POS aff', 'NEG aff'). These results indicate that affect sensing is easier than recognition of judgment or appreciation from text. TOP level results (Table 4) show that our algorithm classifies sentences that convey positive or negative sentiment with high accuracy (92% and 91%, correspondingly). On the other hand, 'neutral' sentences still pose a challenge.

The analysis of errors revealed that system requires common sense or additional context to deal with sentences like '*All through my life I've felt like I'm second fiddle*' (gold standard: 'sadness'; @AM: 'neutral') or '*For me every minute on my horse is alike an hour in heaven!*' (gold standard: 'joy'; @AM: 'neutral').

We also evaluated the system performance with regard to attitude intensity estimation. The percentage of attitude-conveying sentences (not considering neutral ones), on which the result of our system conformed to the fine-grained gold standard (ALL level), according to the measured distance between intensities given by human raters (averaged values) and those obtained by our system is shown in Table 5. As seen from the table, our system achieved satisfactory results in

estimation of the strength of attitude expressed through text.

Range of intensity difference	Percent of sentences, %
[0.0 – 0.2]	55.5
(0.2 – 0.4]	29.5
(0.4 – 0.6]	12.2
(0.6 – 0.8]	2.6
(0.8 – 1.0]	0.2

Table 5. Results on intensity.

7 Conclusions

In this paper we introduced @AM, which is so far, to the best of our knowledge, the only system classifying sentences using fine-grained attitude types, and extensively dealing with the semantics of verbs in attitude analysis. Our composition approach broadens the coverage of sentences with complex contextual attitude. The evaluation results indicate that @AM achieved reliable results in the task of textual attitude analysis. The limitations include dependency on lexicon and on accuracy of the parser. The primary objective for the future research is to develop a method for the extraction of reasons behind the expressed attitude.

References

Alm, Cecilia O. 2008. *Affect in Text and Speech*. PhD Dissertation. University of Illinois at Urbana-Champaign.

Aman, Saima, and Stan Szpakowicz. 2008. Using Roget's Thesaurus for Fine-Grained Emotion Recognition. *Proceedings of the Third International Joint Conference on Natural Language Processing*, Hyderabad, India, pp. 296-302.

Bhowmick, Plaban K., Anupam Basu, and Pabitra Mitra. 2009. Reader Perspective Emotion Analysis in Text through Ensemble based Multi-Label Classification Framework. *Computer and Information Science*, 2 (4): 64-74.

Boucouvalas, Anthony C. 2003. Real Time Text-to-Emotion Engine for Expressive Internet Communications. *Being There: Concepts, Effects and Measurement of User Presence in Synthetic Environments*, Ios Press, pp. 306-318.

Chaumartin, Francois-Regis. 2007. UPAR7: A Knowledge-based System for Headline Sentiment Tagging. *Proceedings of the SemEval-2007 International Workshop*, pp. 422-425.

Choi, Yejin, and Claire Cardie. 2008. Learning with Compositional Semantics as Structural Inference for Subsentential Sentiment Analysis. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 793-801.

Chuang, Ze-Jing, and Chung-Hsien Wu. 2004. Multimodal Emotion Recognition from Speech and Text. *Computational Linguistic and Chinese Language Processing*, 9(2): 45-62.

Finkel, Jenny R., Trond Grenager, and Christopher Manning. 2005. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. *Proceedings of the 43nd Annual Meeting of the ACL*, pp. 363-370.

Hoye, Leo. 1997. *Adverbs and Modality in English*. New York: Addison Wesley Longman Inc.

Izard, Carroll E. 1971. *The Face of Emotion*. New York: Appleton-Century-Crofts.

Kipper, Karin, Anna Korhonen, Neville Ryant, and Martha Palmer. 2007. A Large-scale Classification of English Verbs. *Language Resources and Evaluation*, 42 (1): 21-40.

Kozareva, Zornitsa, Borja Navarro, Sonia Vazquez, and Andres Montoyo, A. 2007. UA-ZBSA: A Headline Emotion Classification through Web Information. *Proceedings of the SemEval-2007 International Workshop*, pp. 334-337.

Liu, Hugo, Henry Lieberman, and Ted Selker. 2003. A Model of Textual Affect Sensing Using Real-World Knowledge. *Proceedings of IUI-2003*, pp. 125-132.

Martin, James R., and Peter R.R. White. 2005. *The Language of Evaluation: Appraisal in English*. Palgrave, London, UK.

Miller, George A. 1990. WordNet: An On-line Lexical Database. *International Journal of Lexicography*, Special Issue, 3 (4): 235-312.

Moilanen, Karo, and Stephen Pulman. 2007. Sentiment Composition. *Proceedings of the Recent Advances in Natural Language Processing International Conference*, pp. 378-382.

Nasukawa, Tetsuya, and Jeonghee Yi. 2003. Sentiment Analysis: Capturing Favorability using Natural Language Processing. *Proceedings of the 2nd International Conference on Knowledge Capture*, pp. 70-77.

Neviarouskaya, Alena, Helmut Prendinger, and Mitsuru Ishizuka. 2009. SentiFul: Generating a Reliable Lexicon for Sentiment Analysis. *Proceedings of the International Conference on Affective Computing and Intelligent Interaction*, IEEE, Amsterdam, Netherlands, pp. 363-368.

Strapparava, Carlo, and Rada Mihalcea. 2008. Learning to Identify Emotions in Text. *Proceedings of the 2008 ACM Symposium on Applied Computing*, Fortaleza, Brazil, pp. 1556-1560.

Strapparava, Carlo, Alessandro Valitutti, and Oliviero Stock. 2007. Dances with Words. *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1719-1724.

Subrahmanian, V.S., and Diego Reforgiato. 2008. AVA: Adjective-Verb-Adverb Combinations for Sentiment Analysis. *Intelligent Systems*, IEEE, 23 (4): 43-50.

Taboada, Maite, and Jack Grieve. 2004. Analyzing Appraisal Automatically. *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text*, pp.158-161.

Whitelaw, Casey, Navendu Garg, and Shlomo Argamon. 2005. Using Appraisal Groups for Sentiment Analysis. *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, CIKM, Bremen, Germany, pp. 625-631.

Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing Contextual Polarity in Phrase-level Sentiment Analysis. *Proceedings of HLT-EMNLP-2005*, ACL, pp. 347-354.