# 論文の内容の要旨

# Compositional Approach for Automatic Recognition of Fine-Grained Affect, Judgment, and Appreciation in Text

## (テキスト中の感情・判断・評価認識のための構成的アプローチ)

ネビアロスカヤ　アレナ

Sharing feelings, pleasant or painful impressions, showing sincere empathy or indifference, exchanging tastes and points of view, advancing moral values, expressing praise or reprehension are indispensable for full-value and effective social interplay between people. With rapidly growing online sources (news, blogs, discussion forums, product or service reviews, social networks etc.) aimed at encouraging and stimulating people's discussions concerning personal, public, or social issues, there is a great need in development of robust computational tools for the analysis of people's preferences and attitudes. Sentiment or subjectivity analysis is nowadays a rapidly developing field with a variety of emerging approaches targeting the recognition of sentiment reflected in written language. Automatic recognition of positive and negative opinions and classification of text using emotion labels have been gaining increased attention of researchers. However, the topic of recognition of fine-grained attitudes expressed in text has been ignored. According to the Appraisal theory proposed by Martin and White (2005), attitude types define the specifics of appraisal being expressed:

(1) *Affect* – personal emotional state or reaction.

(2) *Judgement* – ethical appraisal of person's character, behaviour, skills etc. according to various normative principles.

(3) *Appreciation* – aesthetic evaluation of semiotic and natural phenomena, events, artifacts etc.

The main objectives of our research are:

(1) Fine-grained classification of sentences using attitude types:

*Affect*: nine emotions defined by (Izard 1971): 'Anger', 'Disgust', 'Fear', 'Guilt', 'Interest', 'Joy', 'Sadness', 'Shame', and 'Surprise'.

*Judgment*: positive and negative judgment: 'POS jud' and 'NEG jud'.

*Appreciation*: positive and negative appreciation: 'POS app' and 'NEG app'.

(2) Novel way of deep attitude analysis based on the compositional approach and the semantics of terms.

(3) Analysis of the strength of the attitude and determination of the level of confidence, with which the attitude is expressed, in the interval [0.0, 1.0].

(4) Development of applications driven by attitude-sensing system.

In the thesis, first we describe the developed Affect Analysis Model (AAM) that is based on rule-based linguistic approach for classification of sentences using nine emotion labels or neutral. The proposed algorithm consists of five main stages: (1) symbolic cue analysis; (2) syntactic structure analysis; (3) word-level analysis; (4) phrase-level analysis; and (5) sentence-level analysis. We demonstrate the results of AAM evaluation on two data sets represented by sentences from diary-like blog posts. Averaged accuracy of our system is up to 81.5 percent in fine-grained emotion classification (nine emotion labels and neutral) and up to 89.0 percent in polarity-based classification.

As lexicon-based systems strongly depend on the availability of sentiment-conveying terms in their databases, in order to overcome the problem of lexicon coverage, we introduce original methods for building and expanding sentiment lexicon (SentiFul) represented by sentiment-conveying words that are annotated by sentiment polarity, polarity scores and weights. The main features of the SentiFul are as follows: (1) it is built using not only methods exploring direct synonymy ('*congratulate*'$^{\text{Pos = 0.4}}$ => '*compliment*'$^{\text{Pos = 0.4}}$), antonymy ('*reward*'$^{\text{Pos = 0.2}}$ => '*penalty*'$^{\text{Neg = 0.2}}$), and hyponymy ('*fault*'$^{\text{Neg = 0.6}}$ => '*betise*'$^{\text{Neg = 0.6}}$) relations, but also innovative methods based on derivation and compounding with known lexical units; (2) it is larger than the existing lists of sentiment words; (3) it includes polarity scores, in contrast to most existing sentiment dictionaries that lack assignments of degree or strength of sentiment. The originality and valuable contribution lie in the elaborate patterns/rules for the derivation and compounding processes that have not been considered before. We propose to distinguish the following types of affixes (used to derive new words) depending on the role they play with regard to sentiment features:

(1) *Propagating* affixes preserve sentiment features of the original lexeme and propagate them to newly derived lexical unit. For example: ‘*en-*’ + ‘*rich*’$^{Pos = 0.2}$ => ‘*enrich*’$^{Pos = 0.2}$; ‘*scary*’$^{Neg = 0.9}$ + ‘*-fy*’ => ‘*scarify*’$^{Neg = 0.9}$.

(2) *Reversing* affixes change the orientation of sentiment features of the original lexeme. For example: ‘*harm*’$^{Neg = 0.88}$ + ‘*-less*’ => ‘*harmless*’$^{Pos = 0.88}$; ‘*dis-*’ + ‘*honest*’$^{Pos = 0.1}$ => ‘*dishonest*’$^{Neg = 0.1}$.

(3) *Intensifying* affixes (e.g., ‘*super-*’ in ‘*superhero*’, ‘*over-*’ in ‘*overawe*’) and *Weakening* affixes (e.g., ‘*semi-*’ in ‘*semisweet*’, ‘*mini-*’ in ‘*mini-recession*’) increase/decrease the strength of sentiment features of the original lexeme.

The schematic illustration of our derivation and scoring algorithm is shown in Figure 1.

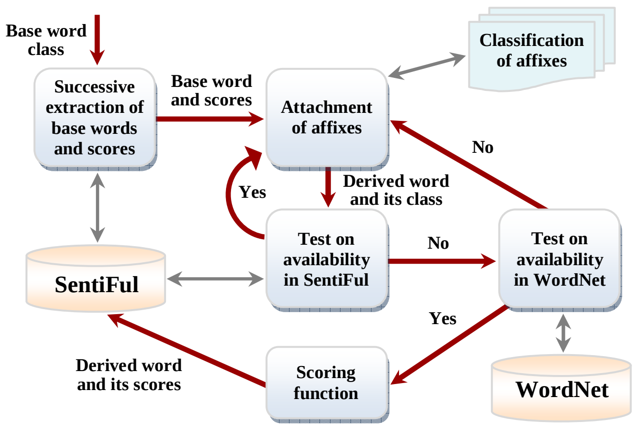


**Figure 1** The algorithm of derivation and scoring of the new words

Besides derivation, we considered important process of finding new words such as compounding, which is a highly productive process, especially in the case of nouns and adjectives. We elaborated the algorithm for automatic extraction of new sentiment-related compounds from WordNet (Miller 1990) using words from SentiFul as seeds for sentiment-carrying base components and applying the patterns of compound formations (for example, ‘*ill*’$^{Neg = 0.467}$ + ‘*famed*’$^{Pos = 0.475}$ => ‘*ill-famed*’$^{Neg = 0.467}$; ‘*pain*’$^{Neg = 0.8}$ + ‘*killer*’$^{Neg = 0.35}$ => ‘*pain-killer*’$^{Pos = 0.575}$; ‘*risk*’$^{Neg = 0.567}$ + ‘*free*’[valence shifter] => ‘*risk-free*’$^{Pos = 0.567}$). We assume that if a compound contains at least one base component that conveys sentiment features, we can predict the valence of this compound. The evaluations of the proposed methods showed that they achieved high accuracy in assigning dominant polarity labels and polarity scores to the words. The method based on compounding performed with the highest accuracy in assigning dominant positive or negative labels, followed by the methods considering hyponymy relations, derivation process, synonymy relations, and antonymy relations (this method yielded noisy results).

In this thesis, we introduce novel compositional linguistic approach for attitude recognition in text. We built a lexicon for fine-grained attitude analysis (AttitudeFul) that includes attitude-conveying terms (e.g., '*honorable*' [POS jud: 0.3], '*unfriendly*' [Sadness: 0.5; NEG jud: 0.5; NEG app: 0.5]), extensive sets of modifiers, contextual valence shifters, and modal operators, which contribute to robust analysis of contextual attitude and its strength. The architecture of the developed Attitude Analysis Model (@AM) is presented in Figure 2.

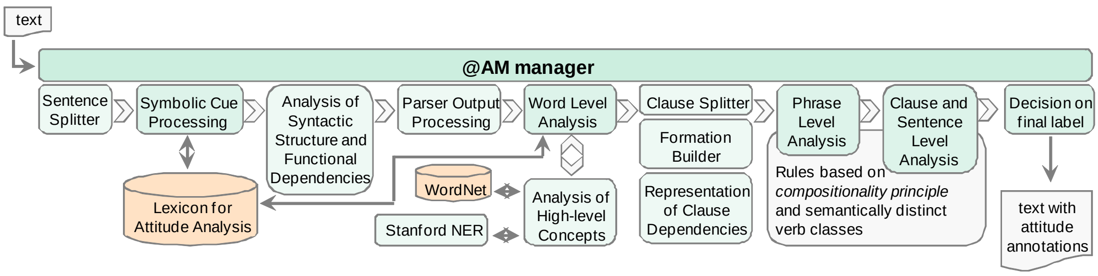


**Figure 2** Architecture of @AM

During the 'Symbolic Cue Processing' stage, the system analyses the occurrences of emoticons, abbreviations and acronyms, interjections, 'question mark' and 'exclamation mark', repeated punctuation, and capital letters. The analysis of syntactic structure and functional dependencies of a sentence is performed by the Connexor Machinese Syntax. On the 'Word Level Analysis' stage, the system checks the availability of the sentence tokens in the AttitudeFul database and gets their annotations depending on the category. In case of an attitude-conveying word, its attitude features are represented as a vector of attitude strengths (intensities): a = [POS jud, NEG jud, POS app, NEG app, Anger, Disgust, Fear, Guilt, Interest, Joy, Sadness, Shame, Surprise]. For example: a('*high-spirited*') = [0.7 (POS jud),0,0,0,0,0,0,0,0,0.7 (Joy),0,0,0]. There are several categories of modifiers registered in the AttitudeFul database: adverbs of degree, adverbs of affirmation, negation words, adverbs of doubt, adverbs of falseness, prepositions, and condition operators.

After the word level annotations are taken from the database, the system turns to the analysis of high-level concepts, which will play the key role in the decision on final attitude label of a sentence. A high-level concept of each noun in the sentence is determined based on:

(1) Analysis of the sequence of hypernymic semantic relations of a particular noun in WordNet (Miller 1999). For example: ‘*student*’ => *PERSON*; ‘*miracle*’ => *EVENT*; ‘*decoration*’ => *ARTIFACT*.

(2) Annotations from the Stanford Named Entity Recognizer (Stanford NER) (Finkel et al. 2005): *PERSON*, *ORGANIZATION*, and *LOCATION*.

Using the data from the ‘Clause Splitter’, the ‘Formation Builder’ module represents each clause as a set of formations: Subject formation (SF), Verb formation (VF) and Object formation (OF), each of which may consist of a main element (subject, verb, or object) and its attributives and complements. The ‘Representation of Clause Dependencies’ module is responsible for building a so-called ‘relation matrix’, which contains information about the dependencies between different clauses in a compound, complex, or complex-compound sentences.

Words in a sentence are interrelated and, hence, each of them can influence the overall meaning and attitudinal bias of a statement. 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.

In order to elaborate rules for the 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. For example, @AM classifies sentence ‘*They prevented* [verb of adverse attitude] *the spread of disease*' as positive appreciation, and '*My whole enthusiasm and excitement disappear* [verb of disappearance] *like a bubble touching a hot needle*' – as conveying negative emotion ('Sadness').

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 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; (2) high-level concepts of nouns based on WordNet; and (3) high-level concepts of named entities based on the annotations from the Stanford NER. For example, @AM outputs different attitude labels for the following sentences containing only one attitude-conveying word '*unfriendly*' (a('*unfriendly*') = [0,0.5 (NEG jud),0,0.5 (NEG app),0,0,0,0,0,0,0.5 (Sadness),0,0]): '*I feel highly unfriendly attitude towards me*', '*The salesperson was really unfriendly*', and '*Plastic bags are environment unfriendly*':

(1) *I* [NomFPP] *feel highly* [modifier: adverb of degree: 1.7] *unfriendly* [NEG aff (Sadness): 0.5; NEG jud: 0.5; NEG app: 0.5] *attitude* [WN: *COGNITION*] *towards me* [AccFPP] =>
=> 'NEG aff' ('Sadness'): 0.85.

(2) *The salesperson* [WN: *PERSON*] *was really* [modifier: adverb of degree: 1.55] *unfriendly* [NEG aff (Sadness): 0.5; NEG jud: 0.5; NEG app: 0.5] =>
=> 'NEG jud': 0.78.

(3) *Plastic bags* [WN: *ARTIFACT*] *are environment* [WN: *STATE*] *unfriendly* [NEG aff (Sadness): 0.5; NEG jud: 0.5; NEG app: 0.5] =>
=> 'NEG app': 0.5.

There are several aspects that distinguish our Attitude Analysis Model from other systems. First, our method classifies individual sentences using fine-grained attitude labels (nine for different affective states, two for positive and negative judgment, and two for positive and negative appreciation), as against other methods that mainly focus on two sentiment categories (positive and negative) or six basic emotions. Next, our Attitude Analysis Model is based on the analysis of syntactic and dependency relations between words in a sentence; the *compositionality principle*; a novel linguistic approach based on the rules elaborated for semantically distinct verb classes; and a method considering the hierarchy of concepts. As distinct from the state-of-the-art approaches, the proposed compositional linguistic approach for automatic recognition of fine-grained affect, judgment, and appreciation in text (1) is domain-independent; (2) extensively deals with the semantics of terms, which allows accurate and robust automatic analysis of attitude type, and broadens the coverage of sentences with complex contextual attitude; (3) processes sentences of different complexity, including simple, compound, complex (with complement and relative clauses), and complex-compound sentences; (4) handles not only correctly written text, but also informal messages written in an abbreviated or expressive manner; and (5) encodes the strength of the attitude and the level of confidence, with which the attitude is expressed, through numerical values in the interval [0.0, 1.0]. The performance of our Attitude Analysis Model was evaluated on data sets represented by sentences from different domains. @AM achieved high level of accuracy on sentences from personal stories about life experiences, fairy tales, and news headlines, outperforming other methods on several measures. In fine-grained attitude classification (14 labels) our system achieved averaged accuracy of 62.1 percent, and in coarse-grained classification (3 labels) – 87.9 percent.

Using Affect Analysis Model and Attitude Analysis Model, we have developed several applications: AffectIM (Instant Messaging application integrated with AAM), EmoHeart (application of AAM in 3D world Second Life), iFeel_IM! (innovative real-time communication system with rich emotional and haptic channels), and web-based @AM interface. We believe that the output of our systems can contribute to the robustness of the following society-beneficial and analytical applications: public opinion mining, deep understanding of a market and trends in consumers' subjective feedback, attitude-based recommendation system, economic and political forecasting, affect-sensitive and empathic dialogue agent, emotionally expressive storytelling, integration into online communication media and social networks.