Emotion Prediction System for Japanese Language Considering Compound Sentences, Double Negatives and Adverbs

Rafal Rzepka, Mitsuru Takizawa and Kenji Araki
Graduate School of Information Science and Technology
Hokkaido University, Japan
{rzepka,araki}@ist.hokudai.ac.jp

Abstract

In this paper we introduce an algorithm that is capable of recognizing emotions of user's statements in order to achieve more effective and smoother human-machine conversation. Many studies of the emotion recognition have been actively conducted in order to quantify affect, but it is rather difficult to recognize it from more complicated sentences, often having double negatives. We describe our enhancements of emotion recognizer by combining emotive expressions lexicon, web-mining techniques, processing compound sentences, and adverb weighting for emotiveness degree modification. The effectiveness of the proposed algorithm for recognizing more complicated sentences was confirmed through evaluation experiments which results are also introduced.

1 Introduction

Sentiment analysis has become a firm and widely studied subfield of Natural Language Processing [Cambria *et al.*, 2013; Socher *et al.*, 2013], but even various new methods have been proposed, their performance depends strongly on availability of tools and resources. This often makes them difficult to be applied for less-resourced languages as Japanese, especially when dealing with processing longer sentences in which valence depends on other factors than emotive load of words. For that reason in this research we combined various approaches to sentiment analysis of Japanese language and investigated the influence of both adverbs [Takizawa *et al.*, 2013a] and sentence structure [Takizawa *et al.*, 2013b] on precision in recognizing affect[1]. In this paper we introduce these methods and describe experimental results which showed that the proposed algorithms outperform other systems.

Recognizing affect in Japanese language relies mostly on lexicons and the surface features of the input text. [Endo *et al.*, 2006], in addition to recognizing emotive expressions, use beginning and ending phrases of sentences and retrieve events like "to graduate is X" from "it is so sad to graduate". However, they ignore degrees of emotions and negations. Another method, based on the SO-score [Turney, 2002] was proposed by [Takemoto *et al.*, 2012], who utilize positive and negative phrases to apply polarity to noun phrases. Their method covers degree modification but ignores cases with more than one emotion and negated expressions. [Shi *et al.*, 2008] proposed a method for sentences without clearly emotive expressions by searching example sentences in the Internet and then performing lexicon based affect recognition (for English the latest approach to this problem is introduced in [Ding and Riloff, 2016]). Their conditional forms based approach is an interesting and simple extension for lexicon-based methods but no semantic analysis of the retrieved sentences is performed. Both surface only and surface with Web search methods were combined by [Ptaszynski *et al.*, 2009]. Their approach deal with negations but does not consider multiple emotions and compound sentences. Modifying emotiveness calculation by processing adverbs was proved effective for English [Benamara *et al.*, 2007; Subrahmanian and Reforgiato, 2008]. For Japanese language a method considering adverbs was proposed by [Saeki *et al.*, 2003], however they concentrated on polarization (positive vs. negative), while our method extends to emotion categories. [Matsumoto *et al.*, 2006] proposed more sophisticated method for measuring emotiveness in sentences. They also used emotion categories, considered modalities, modifying degrees of adverbs and negations. However, they based their emotion categories on closed, non open-source electronic dictionary and the emotive expressions (idioms) data set also was not publicly available for comparisons. Although the methods can not be equally compared, our novel approach, based on downloadable tools with emotion dictionary built in, achieves higher accuracy even in the weakest setup.

2 Emotion Categories

Emotions convey complicated information and they are based on human senses and internal mechanisms which are not yet fully understood. They depend on personal preferences and when one person think about "drinking coffee" having positive associations from the smell and taste sensory experiences, another person can react negatively to the very same object or act. This makes sentiment analysis difficult – more precise analysis and multiple detailed output are required.

[1]The second author, the main contributor of this research, is currently with Hitachi, Co.

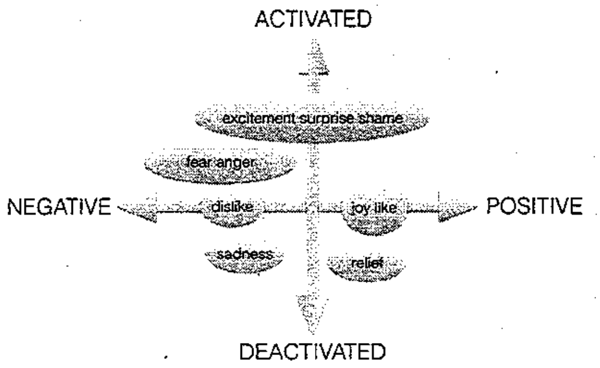

Figure 1: Valence axis according to Russel's theory.

Moreover, our feelings are not easily classified into bipolar categories of positives and negatives which is the mainstream in many applications of sentiment analysis as opinion mining but are insufficient in tasks as dialog processing. Several ways to categorize human emotions were proposed [Whissell, 1989], but for Japanese language a convenient set of ten categories with expression examples from literature was proposed by [Nakamura, 1993]. These expressions are part of ML-Ask tool for recognizing affect in sentences [Ptaszynski *et al.*, 2009] and we decided to use it as a basic method worth expanding. Nakamura's dictionary contains 2,167 lexical examples divided into ten basic emotions characteristic for Japanese: *yorokobi* (joy, delight), *ikari* (anger), *aware* (sorrow, sadness), *kowagari* (fear), *haji* (shame, shyness, bashfulness), *suki* (liking, fondness), *iya* (dislike, detestation), *takaburi* (excitement), *yasuragi* (relief), and *odoroki* (surprise, amazement).

3 System Overview

To address problems with shallow sentiment analysis of Japanese language, we enriched existing algorithms by adding processing of compound sentences and double negations (see Figure 2), then we investigated usefulness of adverbs as the emotion degree modifiers (see Figure 3). For the basic affect analysis we used ML-Ask [Ptaszynski *et al.*, 2009] which covers emotive phrases recognition, lexical clues as exclamation marks or interjections and changes output if the phrases are negated accordingly to Russel's circumplex model of affect [Russell, 1980] (see Figure 1). If no emotive expression from Nakamura's dictionary existed in an input sentence, a set of phrases was retrieved from it and Shi's method for retrieving emotional consequences from the web was used. For acquiring phrase candidates we used trigrams following three rules: a) the first unigram is a particle, b) the last unigram is a verb or adjective, c) exclude cases with symbols and emoticons. Then the Shi's system adds conditionals to the phrase and retrieves 100 snippets from YACIS corpus [Ptaszynski *et al.*, 2012] after sending the exact match query.

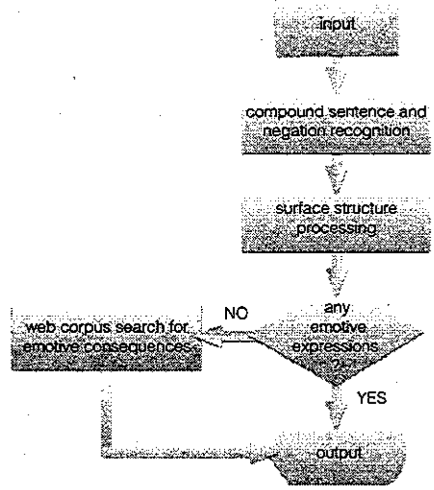

Figure 2: Sentiment analysis regarding compound sentences.

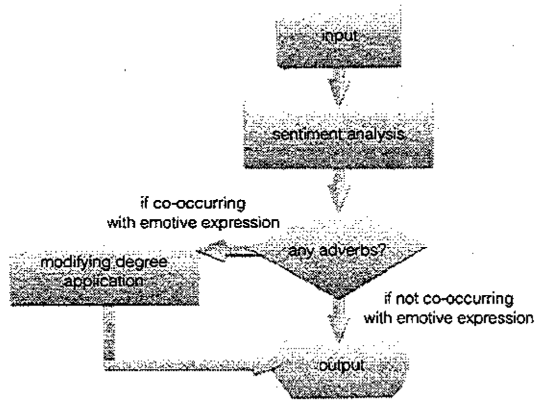

Figure 3: Applying modification degrees to adverbs.

3.1 Compound Sentences

There are many cases when there are descriptions of more than one sentiment, for example "I like soccer, but I also like basketball", and they can be expressed in multiple ways as the continuative form of a verb. For this research purpose we heuristically chose nine most frequent connectors of clauses in Japanese compound sentences (*kedo, keredo, demo, to, node, ga, shi, te* and *noni*) which cover meanings of "but", "and", "then", "still", etc.

3.2 Double Negatives

An input sentence can contain double negatives as in "I am not saying I don't like basketball" and this also can be expressed in various ways, however such expressions are to some extent similar [Lin, 2005]. The authors investigated their usage and chose seven popular forms and added them to the system. For example, if a phrase is followed by *-nai towa kagiranai* ("it's not always true that NEG..."), *-nai towa ienai* ("it can't be said that NEG...") or *naku-wa nai* ("it's not that... NEG"), its emotive value is inverted according to the Russel's theory and "I can't say I don't love curry" is labelled as "like" not "dislike" as in case of ML-Ask system which deals only with single negations.

3.3 Emotive Focus in Compound Sentences

In sentences as "I like soccer, but I don't like watching it" we have two contradictory sentiments to different aspects of one element. From a survey (described in 4.2) we concluded that, as in the soccer example, the more important message is included in the final, second clause. Therefore we set the emotive focus on the final clauses and our system assigns their emotion category as one describing sentiment of the whole sentence.

4 Confirming Effectiveness of Double Negatives and Compound Sentences Processing

4.1 Choosing and Clustering Adverbs

One of the aspects of more precise sentiment analysis is the degree (strength of emotion) measurement. For further experiments we investigated emotion modification capability of adverbs. As the base set we took a list of 93 modifying and mood changing phrases from degree study of [Kawabata, 2011] and used Google[2] search engine to remove rare ones. Those with less than 1,000 hits cooccurring with Nakamura's phrases were deleted, leaving 86.

Next we performed a survey using six adjectives accompanied by the adverbs – three positive (*suki* "like", *ureshii* "happy" and *tanoshii* "fun" and three negative ones (*kirai* "dislike", *kanashii* "sad" and *tsurai* "painful"). Ten students (7 males and 3 females in their 20's) have then evaluated the modification degree of an emotion on the -5 to 5 scale (-5 as negative, 0 as neutral and 5 as positive). We also asked subjects if the pair is natural and if less than seven subjects did not agree, pair was omitted as unnatural. The results of

[2]http://google.co.jp

Table 1: Adverbs clustered into modifying "strong" degree class number 1.

Japanese	English	Japanese	English
kyokutan-ni	extremely	*ketachigai-ni*	incomparably
ketahazure-ni	extraordinarily	*mōretsu-ni*	fiercely
monosugoku	terribly	*samajiku*	tremendously
zetsudai-ni	enormously	*totetsumonaku*	extravagantly
mottomo	most	*osoroshiku*	horribly
ichiban	best	*kyokudo-ni*	extremally
saikō-ni	maximally	*kiwamete*	really
tondemonaku	outrageously	*tokubetsu-ni*	especially
berabō-ni	awfully	*batsugun-ni*	outstandingly
hanahadashiku	seriously	*hijō-ni*	very
danchigai-ni	distinctly	*subarashiku*	wonderfully
shigoku	extremely	*tohōmonaku*	ridiculously
kyōretsu-ni	intensly	*omoikkiri*	as much as
mechakucha	a lot		

Table 2: Second "strong" degree cluster (class number 2).

Japanese	English	Japanese	English
zutto	much	*yake-ni*	awfully
ōi-ni	greatly	*zuibun*	quite
danzen	absolutely	*moro-ni*	completely
sōtō	considerably	*toku-ni*	especially
hitokiwa	exceptionally	*taihen*	really
totemo	very	*ichijirushiku*	significantly
sugoku	very	*sukoburu*	extremely
meppō	exorbitantly	*eraku*	greatly
hidoku	quite	*taisō*	greatly
kanari	quite	*unto*	severely
yatara	excessively	*yoppodo*	great deal
daibu	quite	*toriwake*	particularly

Table 3: The "medium" degree adverbs (classes 3 and 6).

Japanese	English	Japanese	English
muyake-ni	recklessly	*kekkō*	quite
tada	only	*jitsu-ni*	truly
makoto-ni	very	*koto-ni*	especially
yokei	too much	*iya-ni*	terribly
kotonohoka	unusually	*ii kagen*	pretty
jūbun	enough	*nakanaka*	quite
wari-to	comparatively	*sokosoko*	just
māmā	kind of	*kokoro-mochi*	a little
aru teido	to some extent	*hikakuteki*	comparatively
masumasu	more and more	*wariai*	proportionally
isasaka	a bit	*ikuraka*	a bit
ikubunka	rather	*yaya*	slightly

Table 4: The "weak" degree (classes 4 and 5).

Japanese	English	Japanese	English
amari	not too	*anmari*	not too
tada	only	*jitsu-ni*	truly
tashō	a bit	*sukoshi*	a little
chotto	some	*choppiri*	a little bit
wazuka-ni	slightly	*jakkan*	somewhat
shōshō	slightly		

Table 5: Averages determining the degree classes of adverbs.

	1	2	3	4	5	6
Positive	4.7	3.9	3.0	1.9	1.5	2.1
Negative	-4.7	-3.9	-3.3	-0.9	-1.5	-2.3

the survey were used for clustering by the farthest neighbor method and six classes were obtained. We labelled them into three groups: strong, medium and weak (see Tables 1, 2, 3 and 4) and the borderline averages are shown in Table 5.

4.2 Extracting Sentences for Experimenting

To confirm newly added features to the state of the art methods, we performed experiments to investigate effectiveness of compound sentences processing and double negatives. For collecting input sentences we used Japanese Twitter and YACIS blog corpus from which we retrieved 50 sentences by querying with 9 usual compound sentences connectors and negative suffixes. Because in this research we do not deal with emoticons yet, we limited the input sentences to those without symbolic representations of faces. Then, to acquire correct data, we asked 8 students (7 males and 1 female in their 20's) to annotate emotiveness of these sentences by categorizing them into ten categories proposed by [Nakamura, 1993]. If there were more than two types of emotions, they were asked to choose two, if there was not distinct emotion, they labelled the sentence as not emotive. We also instructed the subjects to mark emotive ambiguities. If two different emotions were recognized in separate clauses, they had to choose which one, in their opinion, indicates sentiment of the whole sentence and mark a focus label on it. If two or more subjects agreed on an emotion category, we treated such sentence as a correct one and usable for the evaluation experiment. Because out of 35 sentences with different emotions clauses 32 had the focus label put on the second clause, we decided to treat emotions in the final part of a sentence as decisive.

4.3 Experiment

We used 50 sentences described above as input. In the first experiment we set the surface method [Ptaszynski *et al.*, 2009] and the web method [Shi *et al.*, 2008] combination as the baseline and compared its performance with the proposed version, i.e. processing compound sentences and negations. Because the baseline method outputs more than one emotion if categories share more than half of all discovered types (e.g. "like" 45.0% and "joy" 25.0%) the subjects were also allowed to choose more than one emotion. For a system's output to be treated as correct at least one of the categories had to agree with the correct data.

Table 6: Result comparison of the proposed methods (compound sentences and negations) and the baseline.

	Correct sentences	Incorrect sentences	Accuracy [%]
Baseline	24	26	48.0
Compound sent.	30	20	60.0
Double negatives	26	24	52.0
Both	39	11	78.0

Table 7: Recall and precision for compound sentence focus and double negation ambiguity.

	Recall[%]	Precision[%]
Compound sentence focus	62.9	100.0
Double negation ambiguity	75.0	64.3

4.4 Results Analysis

As shown in Table 6, even when using compound sentences processing and considering double negatives separately, the performance was better than the baseline's. Furthermore, the results showed that combining both methods clearly improved the accuracy (almost 30 points).

One of the reasons for such increase is that clauses were separately processed and phrase queries for the web corpus differed for each clause, which improved the output in 10.0% of cases. The errors were caused by not sufficient coverage of Nakamura's lexicon (4.0% of errors), incorrect emotion conversion in double negations (12.0%) and improper emotive consequences found in search snippets (4.0%).

As shown in Table 7, precision of focusing on the final emotion was perfect, although recall was lower than expected (62.9%). Subjects labelled emotive focus in 35 sentences out of 50 (70.0%), however the proposed method was not able to find 11 of them (22.0%). Regarding compound sentences, not processing them caused failure in 4.0% of cases. More problematic were sentences which had focused clauses but it was not possible to automatically categorize emotions (18.0%); our system was set not to focus on clauses without emotions which caused such errors. Taking into account the emotional ambiguities in double negatives, system achieved 75.0% recall and 64.3% precision, which means that it agreed with 3/4 of human evaluators, and in remaining 25.0% of sentences categorization was wrong or performed on parts not labelled during the survey as emotive (8.0% of such cases). However, although the tests were on a small scale and some obvious problems are remaining, the results showed that processing compound sentences and double negatives can contribute to improvement of existing sentiment analysis methods.

5 Confirming Effectiveness of Using Adverbs as Sentiment Degree Modifiers

We have extracted 500 sentences from Twitter for every adverb we have clustered and left only sentences containing

Table 8: Accuracy of automatic emotion degree estimation (agreement with human evaluators).

	Correct sentences	Incorrect sentences	Accuracy [%]
Weak-Medium-Strong	68	18	79.1
Separated Polarities	51	35	59.3
Overall Average	57	29	66.3

Table 9: Accuracy of automatic emotion degree estimation limited to sentences with correct emotion category.

	Correct sentences	Incorrect sentences	Accuracy [%]
Weak-Medium-Strong	58	13	80.3
Separated Polarities	47	24	66.2
Overall Average	57	29	66.3

emotive phrases. After half-randomly choosing a sentence for every adverb (authors first manually chose candidates by e.g. eliminating sentences with emoticons, then performed the random choice), 86 sentences for experiments were evaluated by 8 students (6 males and 2 females in their twenties) who labelled the sentences with 10 emotion categories, their degrees on scale from -5 to 5, and polarity (positive / negative marks). First, we measured an average modifying degree which was 2.8. Sentences which significantly differed from this level had exclamation marks, ellipsis like "..." or interjections like "whooow" (18.6% cases). Ellipsis were found in 8.1% of sentences and they indicated lower emotiveness. On the other hand exclamations existed in 15.1% of sentences and they had higher emotiveness as well as 2.3% of sentences with a symbol "♪". In 11.6% of sentences symbols and interjections as "aghhhrrr" at the end of a sentence were found and their existence probably influenced the evaluators. However, sentence ending particles like *yo* ("you know") or *ne* ("isn't it?") had a relatively low influence on the degree (0.06 average vs. 0.52 average of other endings).

5.1 Experiment

We tested our algorithm for considering degree modification by adverbs by inputting the 86 sentences described above and automatic emotion category estimation achieved 82.3% accuracy (71 correct and 15 incorrect judgements). Because we wanted to investigate how precise is the degree estimation, we set three types of comparing results with human-created correct set: a) three levels estimation ("weak" - "medium" and "strong"), b) agreement level considering polarities separately and c) overall agreement level without separating polarities. The borderline values for a) were set by averages in clustering process. Two remaining types were based on human subjects evaluation and if the difference between system's output and human average estimation was not larger than 0.5 (within 1 point) we treated the output as correct.

For b) we used separate averages for positive and negative annotations, while for c) we calculated an overall average for both polarities combined. The results of the experiment are shown in Table 8. We have also checked how the degree estimation algorithms works if emotion category is correctly recognized. The results of this test run on data with 16 excluded incorrect categorizations are shown in Table 9.

Table 10: Examples of incorrect emotion category estimations.

	Hum	Sys
Keta-hazure-ni omoroi eiga yatta (That movie was extraordinarily goo)	Joy	NoEmo
Jitsu-ni shiawase-na shōgakkō jidai datta-wa (They were really happy primary school times)	Relief	Joy
Kyō-wa hisabisa-ni hitome-mo ki-ni sezu-ni omoikkiri naita (Today, for the first time after ages, I could cry without caring about being seen by somebody)	Sad	Excite

Table 11: Examples of correct recognition for "weak" (W), "medium" (M) and "strong" (S) categorization of adverbs.

	Human	Proposed (W-M-S)
Konkai keta-chigai-ni omoshiroi desu-yo-ne (This time it's incomparably funny, isn't it?)	S	S
Ima-to natte-wa mō anmari tanoshiku-nai keredo (But now, it's not so pleasant anymore)	W	W
Mada choppiri itai-kedo-na... (It still hurts a bit...)	W	W

5.2 Results Analysis

The emotion categorization task achieved 82.3% accuracy and most errors was caused by expressions missing from Nakamura's lexicon (examples shown in Table 10). Actually words as *omoshiroi* ("interesting") and *suki* ("like") existed in the lexicon but in their basic forms; in the input the former was written in a colloquial manner (*omoroi*) and the former in syllables (*kana*) instead of Chinese characters (*kanji*). The degree estimation itself has achieved 79.1% of accuracy when simple weak-medium-strong categories were used (see Tables 11 and 12). The most problematic was the "medium" class which adverbs were found in 13 incorrectly estimated sentences (15.1%), which suggest that further categorization might be needed. The results for more strict evaluation by comparing averages showed 59.3% of accuracy and 66.2% for correct emotion recognitions (see Tables 13 and 14).

Main reason for incorrect recognitions were due to ambiguous emotive expressions as *hazukashii* which can be interpreted as strongly negative "shameful" or more positive

Table 12: Examples of incorrect recognition in "weak" (W), "medium" (M) and "strong" (S) categorization of adverbs.

	Human	Proposed (W-M-S)
Jitsu-ni shiawase-na shōgakkō jidai datta-wa (That was really happy primary school time)	S	M
Kuro-basu kanren-wa mō majime-ni iretakunee-tte hodo wari-to okotteru-zo (They are quite angry and even say they don't wanna use anything related to "Kuroko's Basketball")	S	M
Sakura-mo, kōyo-mo nai-kara yaya monotarinai... (I'm a bit disappointed that there is no cherry blossom, no red leafs	W	M

Table 13: Examples of correct recognitions for polarized (POL) and overall non-polarized averages (OVR).

	HUM	POL	OVR
Yūrei nanka-yori ikiteiru ningen-no kyōki-no-ga yoppodo kowai wa (I'm much more scared by living people's madness than by ghosts and stuff)	-4.25	-4.11	4.11
Meganekko-ga tokubetsu-ni suki!! (I especially like girls in glasses!!)	4.50	4.25	4.25
O-demukae makoto-ni ureshii kagiri degozaimasu (I am truly glad you came to meet me)	3.88	3.33	3.33

Table 14: Examples of incorrect recognitions for polarized (POL) and overall non-polarized averages (OVR).

	HUM	POL	OVR
Jitsu-ni shiawase-na shōgakkō jidai datta-wa (That was really happy primary school time)	3.88	3.30	3.30
Chinami-ni watashi-wa kafunshōnanode o-hana-wa amari suki-ja arimasen (By the way I am allergic to pollen so I don't like flowers too much)	-2.00	-0.90	0.90
Mada choppiri itai-kedo-na... (It still hurts a bit...)	-1.38	-1.20	1.20

"shy". Such context dependencies were found in 17.4% of all sentences. When polarization is not possible, the overall average is calculated and big discrepancies happen which is visible in the results. Achieved accuracy was 66.3% and it increased to 73.2% for correct emotion category recognitions. One of the reasons for the incorrect recognitions are the sentence endings both lexical as "ya know" (5.8%) and symbolic as "!", "..." or "♪" (7.0%). Therefore we utilized ML-Ask's capability for recognizing such features and added weights to every match. According to the findings from human evaluation, we experimentally set higher weight (0.20) to symbolic endings and lower (0.10) to lexical endings as particles (if the sum with added weights was higher than 5 or lower than -5, the algorithm treated the output as maximum values of 5 and -5). This addition caused 4.7 points increase in accuracy for separated polarities and 4.6 for overall averages.

6 Conclusions and Future Work

In this paper we introduced additional features to the combination of lexicon-based and web-based method for sentiment analysis. Firstly, we confirmed that considering emotions in compound sentences and processing double negatives improve the accuracy (78.0%) in a significant way (almost 30 points of increase when compared to the baseline). Then we reported the results of our experiments with using emotion degree modification capability of adverbs for more precise emotion estimation. We were able to confirm that the adverbs help to achieve high accuracy of sentiment analysis (82.3%). We also investigated how close the clustered adverbs-based degree estimations are to human evaluators, and the proposed system reached 79.1% agreement for three "weak", "medium" and "strong" degree types, showing that further clustering for "medium" type might be needed. In the second stage of this investigation we compared how similar our proposed method was to human subjects in estimating emotiveness of sentences with adverbs. When averages of negative and positive polarity were concerned, the proposed system achieved accuracy of 59.3% and if not concerned 66.2%. After adding weights to emotive sentence endings, the algorithm achieved 64.0% and 70.9% for both measuring methods respectively. The experiments revealed that for better agreement with human subjects we need more adequate, not ambiguous phrases in lexicons to avoid incorrect recognitions like in sentences containing word "envy" which polarity depends on situation. Refining the lexicon is one of our current works, as well as investigating emoticons' influence on affect. We also plan to conduct bigger scale experiments and test other clustering methods.

References

[Benamara *et al.*, 2007] Farah Benamara, Carmine Cesarano, Antonio Picariello, Diego Reforgiato Recupero, and Venkatramana S Subrahmanian. Sentiment analysis: Adjectives and adverbs are better than adjectives alone. In *ICWSM*. Citeseer, 2007.

[Cambria *et al.*, 2013] E. Cambria, B. Schuller, Yunqing Xia, and C. Havasi. New avenues in opinion mining and sentiment analysis. *Intelligent Systems, IEEE*, 28(2):15–21, March 2013.

[Ding and Riloff, 2016] Haibo Ding and Ellen Riloff. Acquiring knowledge of affective events from blogs using label propagation. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*, 2016.

[Endo *et al.*, 2006] Daisuke Endo, Manami Saito, and Kazuhide Yamamoto. Extracting expressions evoking emotions using dependency structure (in japanese). In *Proceedings of The Twelve Annual Meeting of The Association for Natural Language Processing*, 2006.

[Kawabata, 2011] Motoko Kawabata. Study regarding perspectives on classifying expressions of degree (in japanese). In *Technical report of Aichi Institute of Technology*, volume 47, pages 115–124, 2011.

[Lin, 2005] Lechang Lin. A cognitive study of double negatives : Focus on the correlation between form and meaning. *Studia humanitatis*, 3:27–39, mar 2005.

[Matsumoto *et al.*, 2006] Kazuyuki Matsumoto, Ren Fuji, and Shingo Kuroiwa. *Computational Intelligence: International Conference on Intelligent Computing, ICIC 2006, Kunming, China, August 16-19, 2006. Proceedings, Part II*, chapter Emotion Estimation System Based on Emotion Occurrence Sentence Pattern, pages 902–911. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[Nakamura, 1993] Akira Nakamura. *Kanjo hyogen jiten [Dictionary of Emotive Expressions]*. Tokyodo Publishing, 1993.

[Ptaszynski *et al.*, 2009] Michal Ptaszynski, Pawel Dybala, Wenhan Shi, Rafal Rzepka, and Kenji Araki. Towards context aware emotional intelligence in machines: Computing contextual appropriateness of affective states. In *Proceedings of Twenty-first International Joint Conference on Artificial Intelligence (IJCAI-09*, pages 1469–1474, Pasadena, July 2009.

[Ptaszynski *et al.*, 2010] Michal Ptaszynski, Pawel Dybala, Wenhan Shi, Rafal Rzepka, and Kenji Araki. Contextual affect analysis: A system for verification of emotion appropriateness supported with contextual valence shifters. *International Journal of Biometrics*, 2(2):134–154, 2010.

[Ptaszynski *et al.*, 2012] Michal Ptaszynski, Pawel Dybala, Rafal Rzepka, Kenji Araki, and Yoshio Momouchi. Yacis: A five-billion-word corpus of japanese blogs fully annotated with syntactic and affective information. In *Proceedings of The AISB/IACAP World Congress*, pages 40–49, 2012.

[Russell, 1980] J.A. Russell. A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161–1178, 1980.

[Saeki *et al.*, 2003] Mika Saeki, Masato Tokuhisa, Jin'ichi Murakami, and Satoru Ikehara. Judgment of affective expressions by adverb and adjective (in japanese). In *Forum on Information Technology*, pages 117–118, 2003.

[Shi *et al.*, 2008] Wenhan Shi, Rafal Rzepka, and Kenji Araki. Emotive information discovery from user textual input using causal associations from the internet. In *Proceedings of the Forum on Information Technology (FIT-08)*, pages 267–268, 2008.

[Socher *et al.*, 2013] Richard Socher, Alex Perelygin, Jean Y Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*, volume 1631, page 1642. Citeseer, 2013.

[Subrahmanian and Reforgiato, 2008] Venkatramana S Subrahmanian and Diego Reforgiato. Ava: Adjective-verb-adverb combinations for sentiment analysis. *Intelligent Systems, IEEE*, 23(4):43–50, 2008.

[Takemoto *et al.*, 2012] S. Takemoto, M. Tokuhisa, and S. Murata. So-score to pataan jisho-o mochiita jousho suitei (sentiment prediction using so-score and patterns lexicon). In *Proceedings of The Twelve Annual Meeting of The Association for Natural Language Processing*, 2012.

[Takizawa *et al.*, 2013a] Mitsuru Takizawa, Rafal Rzepka, and Kenji Araki. Emotion recognition considering modifyuing degree of adverbs (in japanese). In *Technical Report of Language Sense Engineering JSAI SIG-LSE-B302-3*, pages 13–18, 2013.

[Takizawa *et al.*, 2013b] Mitsuru Takizawa, Rafal Rzepka, and Kenji Araki. Improvement of emotion recognition system considering the negative form and compound sentence improvement of emotion recognition system considering the negative form and compound sentence. In *Proceedings of the 29th Fuzzy Systems Symposium*, pages 774–777, 2013.

[Turney, 2002] Peter D Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 417–424. Association for Computational Linguistics, 2002.

[Whissell, 1989] C. M. Whissell. *The Dictionary of Affect in Language. Emotion: Theory, Research and Experience: vol. 4, The Measurement of Emotions*. New York: Academic, 1989.