From Words to Emoticons: Deep Emotion Recognition in Text and Its Wider Implications

Rafal Rzepka[1], Mitsuru Takizawa[1], Jordi Vallverdú[2], Michal Ptaszynski[3], Pawel Dybala[4], Kenji Araki[1]
[1] Hokkaido University, Japan
rzepka, araki@ist.hokudai.ac.jp

[2] Autonomous University of Barcelona, Spain
jordi.vallverdu@uab.cat

[3] Kitami Institute of Technology, Japan
ptaszynski@cs.kitami-it.ac.jp

[4] Jagiellonian University, Poland
pawel.dybala@uj.edu.pl

ABSTRACT: *This paper summarizes several lexical methods for more comprehensive affect recognition in text using an example of typed utterances. We introduce a set of algorithms that are capable of recognizing emotions of user's statements in order to achieve more effective and smoother human-machine conversation. Aspects often neglected by existing systems working with Japanese language, e.g. compound sentences, double negation sentences, modifiers as adverbs and emoticons were combined and their higher effectiveness in recognizing affect in more complicated sentences was confirmed through evaluation experiments. The results are introduced together with separate analysis of emoticons' influence on emotional load. We also discuss importance of predicting human emotions not only in the field of human-computer interaction but also its meaning for developing ethical chatbots.*

Keywords: Emotion Recognition, Affective Computing, Lexical Approach, Emoticons

Received: 11 March 2017, Revised 9 April 2017, Accepted 19 April 2017

1. Introduction

Detecting affective states from voice, faces, gestures or text [5, 42] is a task with several goals in mind. They are usually human-centered, meaning that people want their computers to understand what is not obvious for human beings: how a product or service is evaluated by users [49, 4], what sentiments are among electorates [6], how depressed bloggers are [56], how emotions change with time and space (of both real people [55] and literary characters [30]), etc. In our case, we concentrate on processing emotions in order to teach computers to extract knowledge about how human emotions change in case of context. Our goal is to treat affective aspects of humans as a part of commonsense knowledge needed for smooth interaction - not only on a conversational level but also when it comes to decision making and taking actions. One example of such approach is sentiment analysis used by artificial tutors [22], where understanding students' emotions helps educational software to evaluate learner's affective state and apply this understanding to more effective teaching process. After the early success of negative-positive polarity recognizers [49], researchers have been exploring the recognition also of different types of emotions (based on various models), they employ rule-based approaches and machine learning methods, moving from very shallow approaches to ones that deal with concepts [4] which step by step leads to taking context into the consideration. In our opinion, lack of annotated data is one of the biggest obstacles for machine learning approaches to processing the context, therefore we aim at automatic annotation [31] of vast textual resources. To achieve this goal, many different problems must be solved to bring the recognition beyond words or word vectors, for example processing metaphors, analogies, sarcasm, irony, dealing with non-verbal clues such as emoticons and go further merging with clues brought by image and video understanding methods. The richer auto-annotation becomes, the better output of machine learning methods can be expected. This process could be done also by stochastic methods, but by controlling the linguistic and semantic features we can avoid low understandability of the statistical approaches (blackbox problem) which is especially important in our automatic moral judgement research. Any way we approach affect discovery, our inborn, natural understanding of emotions is difficult to mimic with machines for a similar reason that computers have problems with processing common sense knowledge - they are often not expressed directly and only our experience tells us which sentiment is natural in a given situation. To deal with this problem, we have been using WWW resources to retrieve people's reactions to similar happenings and to find clues regarding their descriptions [35], emphasizing possibilities for achieving moral decision capabilities by machines. The relationship between morality and emotions was discussed from the times of Aristotle and has been discussed by philosophers till today [2]. But as we believe that coherent meta-ethics should start from the generalities of emotions to reach an ethics of complexity [50], we need to assure that computers can process such a contextual complexity, in which the emotions of hundreds, thousands or millions of beings are generating an "emotional common sense". Language affects moral decisions [9] and we believe that chatbots may positively influence our behaviors [37]. As Cambria [3] predicts, the next important shift in processing affect, after dealing with concepts, is understanding pragmatics. This paper introduces our attempts to broaden lexical processing of text to open a way to deeper understanding of emotions. One of the possible usages of such capabilities is creating conversational agents for further testing [53], therefore we tested our algorithms on written utterances and confirmed that considering emotions in compound sentences and processing double negatives significantly improve the accuracy. 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.

This paper is organized as follows. in the following section we describe the current state of the field of recognizing affect concentrating on applications to Japanese language. In Section 3 we explain emotion classification of our choice and in Section 4 we present our Affect Recognizer together with its processing of compound sentences (4.1), double negatives (4.2) and emotive focus in compound sentences (4.3). Section 5 describes how the sentences were retrieved and used in our experiments, it also shows the analysis of our experimental results (5.4). In the following section (6), we present our inclusion of adverbs as affect degree modifier: how we chose and clustered them (6.1), how we created the correct data set (6.2) and tested our enhanced algorithm (6.3) showing the results in (6.4). We have also investigated influence of emoticons on the change in the emotive evaluation of a sentence. Section 7 is dedicated to the survey we performed and Section 8 underlines the need of deeper affect processing putting stress on utterances. We conclude the paper in Section 9 summarizing all experimental results and sharing our plans for further development.

2. State of The Art

Ways of expressing emotions vary between cultures and it was shown with Japanese as an example of differences in face expressions [24] and language usage [45, 27]. Recognizing affect in Japanese language text relies mostly on lexicons and the surface features of the input text, although [16] has shown that utilizing techniques from machine translation task can lead to improvement. However, their method, as well as older [46] and the latest machine learning approaches [59], require annotated data and can be used if massive auto-annotation we aim at becomes possible. For that reason we concentrate on methods which could be used without costly and time-consuming annotations. To improve quality of this process, we are trying to deal with finer lexical details and use them to achieve higher accuracy of affect recognition. Already a decade ago, in addition to recognizing Japanese emotive expressions, Endo et al. [12] used beginning and ending phrases of sentences and retrieved events like “to graduate is X” from sentences as “it is so sad to graduate”, but their research did not cope with negations and degrees of emotions. Another method, based on the SO-score [49] was proposed by Takemoto et al. [47], who utilized 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. [41] proposed a method for evaluating phrases without clearly emotive expressions by searching example sentences in the Internet and then performing lexicon based affect recognition. Currently researchers try to tackle this task also for English [11]). Shi’s conditional forms based approach is an interesting and simple extension for lexicon-based methods but in our opinion lack of semantic analysis of the retrieved sentences is its biggest flaw. Both surface only and surface with Web search methods were combined by Ptaszynski et al. [32]. Their approach deals with negations but does not consider degrees or compound sentences, which we cover in this research. For that reason our project can be treated as an extension of their work. Modifying emotiveness calculation by processing adverbs was proved effective for English [1, 44]. For Japanese language a method considering adverbs was proposed by [38] , however they concentrated on polarization (positive vs. negative), while our method extends to emotion categories.

Matsumoto et al. [25] proposed more sophisticated method for measuring emotiveness in Japanese 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 compared with the same input, our novel approach, based on open-source tools with emotion dictionary built in, achieves higher accuracy than reported by Matsumoto et al., even in case utilizing the weakest setup.

3. Utilized Emotion Classification

Although natural (as a part of folk psychology) in everyday life, information conveyed by emotions is complicated and hard to process because it is based on human senses and internal mechanisms which are not yet fully understood [10]. Emotional reactions also depend on individual preferences so while one person thinks about “drinking coffee” having positive associations from the smell and taste sensory experiences (memories), 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. 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 [58], but for Japanese language a convenient set of ten categories with expression examples from literature was proposed by [28]. These expressions are part of ML-Ask tool for recognizing affect in sentences [32] and we decided to use it as a basic method worth expanding. All phrases used in this research originate from Nakamura’s dictionary [28] which contains 2,167 lexical examples divided into ten basic emotions characteristic (as the dictionary’s author concludes his work) for Japanese (see Table 1). Nakamura has painstakingly classified expressions from Japanese literature for decades and we turned his endeavor into a helpful base for predicting emotion categories from text.

4. Overview of our Affect Recognizer

As mentioned above, to address problems with shallow sentiment analysis of Japanese language[1], 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). Finally, we examined how the appearance of emoticons modifies the perception of an utterance. For the basic affect analysis we used ML-Ask [32] 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 [34] (see Figure 1).

If an input sentence contains emotive words or features they are labelled accordingly, but if no emotive expression

[1] *Original Japanese words are represented in italic throughout the paper.*

Figure 1. Valence axis according to Russel's theory

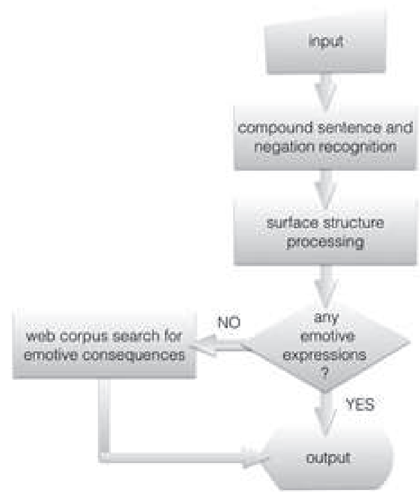

Figure 2. Sentiment analysis regarding compound sentences. Input is a sentence and output is recognized emotions

from Nakamura's dictionary exists in the input,phrases from this sentence become an input for Shi's method for retrieving emotional consequences from the Web [41]. To assure basic context processing, this method prioritizes acts instead of single words (processing "losing money" achieves better accuracy than when "losing" and "money" are treated separately). 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 [31] after sending the exact match query.

4.1 Compound Sentences

Often utterances describe 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.

4.2 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 [23]. 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

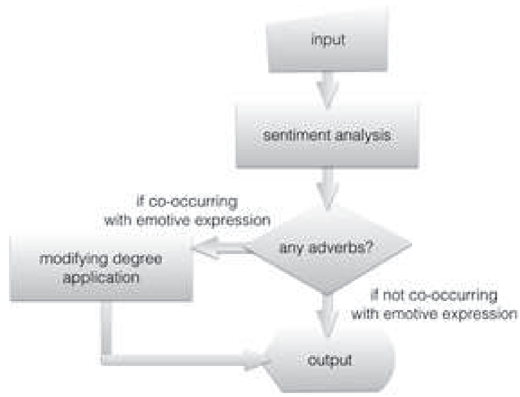

Figure 3. Applying modification degrees to adverbs

Category	Translation
yorokobi	joy, delight, enjoyment
ikari	anger, rage, fury
aware	gloom, sorrow, sadness
kowagari	fear, terror, dread
haji	shame, shyness, bashfulness
suki	liking, fondness, loving
iya	dislike, detestation, unwillingness
takaburi	excitement, agitation
yasuragi	relief, peace
odoroki	surprise, amazement, astonishment

Table 1. Ten categories of emotions by Nakamura

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.

4.3 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 the one describing sentiment of the whole sentence.

5. Confirming Effectiveness of Double Negatives and Compound Sentences Processing

5.1 Extracting Sentences for Experiments

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 [28]. If there were more than two types of emotions, they were asked to choose two, if there was no 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

	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

Table 2. Averages determining the degree classes of adverbs

	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 3. Result comparison of the proposed methods (compound sentences and negations) and the baseline

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.

5.2 Experiments with Compound Sentences and Negation Processing

We used 50 sentences described in previous section as input. In the first experiment we set the surface method [32] and the Web method [41] 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% and "joy" 25%) 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.

5.3 Results Analysis

As shown in Table 3, 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% of cases. The errors were caused by not sufficient coverage of Nakamura's lexicon (4% of all sentences), incorrect emotion conversion in double negations (12%) and improper emotive consequences found in search snippets (4 %).

As shown in Table 4, Precision of focusing on the final emotion was ideal, although recall was lower than expected (62.9%). Subjects labelled emotive focus in 35 sentences out of 50 (70%), however the proposed method was not able to find 11 of them (22%). Regarding compound sentences, not processing them caused failure in 4% of cases. More problematic were sentences which had focused clauses but it was not possible to automatically categorize emotions (18%); 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% recall and 64.3% precision, which means that it agreed with 3/4 of human evaluators, and in remaining 25% of sentences categorization was wrong or performed on parts not labelled during the survey as emotive (8% 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.

6. Using Adverbs as Sentiment Degree Modifiers

6.1 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

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

Table 4. Recall and precision for compound sentence focus and double negation ambiguity

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 5. Adverbs clustered into modifying "strong" degree class number 1

and mood changing phrases from degree study of [19] 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.

6.2 Correct Set Preparation

We have extracted 500 sentences from Twitter for every adverb we have clustered and left only sentences containing emotive phrases using furthest neighbor method. 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).

6.3 Adverb Processing 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).

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

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 6. Second “strong” degree cluster (class number 2)

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 7. The “medium” degree adverbs (classes 3 and 6)

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 (see Tables 5, 6, 7 and 8). 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 9. 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 10.

6.4 Results Analysis

The emotion categorization task achieved 82.3% accuracy and most errors were caused by expressions missing from Nakamura’s lexicon (examples shown in Table 11). Actually words like *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 12 and 13). 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 14 and 15).

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 8. The "weak" degree (classes 4 and 5)

	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 (agreement with human evaluators)

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

7. Influence of Emoticons

To see what are differences between adverbs and emoticons in influencing emotional load, we have conducted additional study [36]. In order to investigate this matter, we designed a set of sentences for questionnaires that were later answered by Japanese native speakers. Previously it was confirmed that both emoticons [17, 43] and adverbs [1, 48] are important in sentiment analysis but our scope was to see what role they play in context of intensifying affect. To create credible questionnaires we first retrieved a set of Japanese sentences that included emotive expressions from Nakamura's set. The source was previously used YACIS [31] corpus. We limited emoticons to the ones used in work of Urabe et al. which do not use any additional characters outside the base face borders set by brackets (e.g. these which symbolize hand gestures, movement or items like cigarettes). To assure that emotion expressed by a sentence is compatible with the emoticon we limited sentence-emoticon pairs to these representing the same emotion category which helped avoiding sarcasm sometimes conveyed by opposite affect emoticon. To make experiment less burdensome for evaluators and to keep semantical consistency we added usual adverbs if they were missing in the retrieved sentences. The last step in preparing data for our questionnaires was to show the sentences to Japanese native speakers in order to check if they are syntactically and semantically correct. This process brought us 22 sentences (see Table 16) that included at least one emotive word, one adverb and one emoticon. Using those selected base sentences we prepared four sets of questionnaires with randomized sets of versions a) with both adverb and emoticon omitted; b) containing adverb only; c) containing emoticon only; and d) containing both adverb and emoticon. An example set of sentences is given below (last sentence is the unstripped original).

- *Itoshii sugata desu-ne* ("It's such a beloved appearance")
- *Mechakucha itoshii sugata desu-ne* ("It's such an amazingly beloved appearance")
- *Itoshii sugata desu-ne* (ˆ_ˆ) ("It's such a beloved appearance (ˆ_ˆ")

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

Table 10. Accuracy of automatic emotion degree estimation limited to sentences with correct emotion category

Processed sentence and its translation	**Human**	**System**
Keta-hazure-ni omoroi eiga yatta (That movie was extraordinarily good)	Joy	NoEmo
Jitsu-ni shiawase-na shōgakkōjidai 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)	Sad	Excite

Table 11. Examples of incorrect emotion category estimations

- *Mechakucha itoshii sugata desu-ne* (ˆ_ˆ) ("It's such an amazingly beloved appearance (ˆ_ˆ)")

The purpose of separating similar sentences and randomization was to make sure that the participants would not be influenced by their previous answers. The purpose of separating similar sentences and randomization was to make sure that the participants would not be influenced by their previous answers.

7.1 Emoticon Influence Survey

Twenty one subjects replied to our survey (one person was asked to fill in 4-5 sets). There were 17 males and 4 females, from 21 to 43 years old (average age 25.2), mostly with background in computer science (16) and few in humanities (5). We asked every subject about their emoticon usage and only two subjects answered they never use them. Six subjects said they always use emoticons, other six that they use them occasionally, seven admitted to use them rarely.

The task was to mark emotional load of every sentence by choosing an emotion degree on 1-5 scale where 1 is no load and 5 is a strong load.

7.2 Survey Results and Analysis

As expected, sentences with only emotive words scored lower than those containing also adverbs and emoticons. What was more surprising, the highest scores were not so much lower for sentences with only emoticons (36.4%) and both emoticons and adverbs (40.9%) - see Figure 4.

There were five cases (22.7%) where emotive phrase and adverb sentences conveyed the most distinct emotions, which could mean the emoticons have a stronger influence on reader's affective perception on a sentence than adverbs. However, a closer look at these examples does not allow to head to any certain conclusions. Below we discuss various trends we have observed.

Sentences created from "*Unto tanoshinde kudasai* (ˆoˆ)" ("Please have lots of fun (ˆoˆ)") were evaluated as expected, meaning the highest score (joy: 4.8, like: 4.6) for the original with adverb and emoticon, lower evaluation for emotive phrase only (joy: 3.4, like: 2.8) and for emoticon only (joy: 3.8, like: 3.8). However, its most simplified version with emotive phrase only scored second high (joy: 3.7, like: 4.0). "*Tanoshinde kudasai*" ("Please have fun") as a dry request-like expression may be concealing emotional associations which cannot be properly conveyed by pure text, therefore a reader might be assuming involvement of other emotions. Average sums of scores for all types of sentences did not differ significantly, but in 27.3% cases the simplest sentences had the highest average totals, showing that subjects had quite different opinions and were assigning scores to emotions which were evaluated as non-existent

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

Figure 12. Examples of correct recognition for "weak" (W), "medium" (M) and "strong" (S) categorization of adverbs

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

Figure 13. Examples of incorrect recognition in "weak" (W),"medium" (M) and "strong" (S) categorization of adverbs

in sentences with emoticons and adverbs. For instance, sentence "*Kyūkutsu deshita-ga, kawaikatta desu*" (≧ ω ≦)) ("It looked a bit tight but it was cute (≧ ω ≦)") had average total of scores 20.6 (sums of all points given by subjects to a sentence), while its most simplified version's average was 21.7, because subjects have chosen all types of emotions except sadness. In case of the semantically richer original, only half categories were chosen (joy, likeness, relief, excitement and surprise, no negative interpretations). Subjects seem to be guessing what the utterer could feel in the semantically the poorest versions and probably for that reason there were also scores for anger, sadness and shame. The same trend was seen in sentences as "*Kyō-wa shinpai-sugiru hi desu-ne*" ("Today is the day I am worried"), or *Hikōki-ni noru toki-wa tanoshii desu-ne* ("It's fun when you take an airplane"). Scores for sadness and dislike may suggest that some subjects suspect traces of irony or they project their own emotional associations while evaluating the sentences.

In other example, shame / shyness in "*Watashi-mo jitsu-wa hazukashigarimono desu-kara*!" ("Because actually I'm a shy guy, too!") scored higher than the version with adverb *hijō-ni* ("very") and the version with emoticon (*^_^*). We think we can witness two phenomena in such cases. When it comes to using an adverb to underline one's own weakness, it might be perceived as an intensifier of author's modesty, not of the shyness itself. Smiling emoticon, on the other hand, seems to weaken the statement, making it more funny and lighter than just a straight confession. There is a possibility that there is a relation between the strength of passive utterances and a their smaller number of adverbs and emoticons.

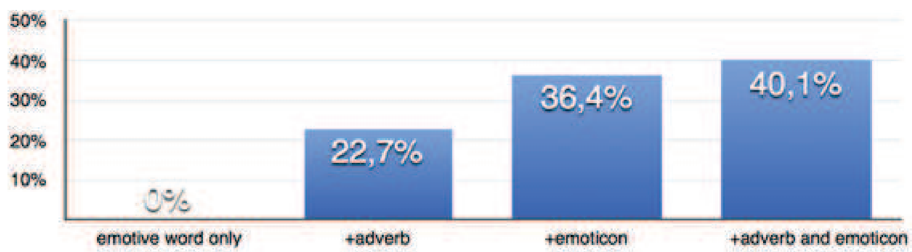

Figure 4. Rate differences between the highest scores depending on a sentence variation

Processed sentence and its translation	Human	Polarized	Overall
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

Figure 14. Examples of correct recognitions for polarized and overall non-polarized averages

Processed sentence and its translation	Human	Polarized	Overall
Jitsu-ni shiawase-na shōgakkōjidai 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

Figure 15. Examples of incorrect recognitions for polarized and overall non-polarized averages

"*Shōshō tanoshimi-ni shite ita-no da-ga, sore-hodo-no kangeki-wa nakatta*" ("I was looking for it a bit but eventually I wasn't moved that much") is an example where adverb *shōshō* ("a bit") was influential enough to made this version score significantly higher in sadness category than the richest version (4.3 vs 3.0). Our first intuition was that also in this case the emoticon has softened the interpretation but when we confirmed average scores, it appeared that emoticon version scored lower (3.2), and the lowest when enriched also by adverb (3.0) as mentioned above. Emotive word only sentence scored slightly higher (3.4) leaving the adverb only version to be felt as distinctly sadder. The same trend is visible in evaluation of sentence "*Hannya-no Kawashima-san-ni ore-wa totemo tekii-o mochimashita*" ("I feel such a hostility toward Mr. $\gg$ Hannya $\ll$Kawashima") and *Nodo-ga kakudan-ni raku-ni narimashita* ("My throat is distinctly better now"). In the first case, emoticon that caused a big drop in anger evaluation (from 4.4 to 2.2) was ($'\varepsilon'$). As a less common one it is possible that users felt puzzled by it, especially with the epsilon symbol between the eyes. Another potential explanation is that because more than half of cases where adverbs were the most influential were conveying negative emotions, the lack of emoticons could be intensifying these emotions (when added they naturally tend to weaken the negative load). Another example of interesting adverb - emoticon relation is contained in a sentence "*Zutto akogareteita-kara kiai jūbun* (*^^*)" ("Because I admired him for ages, I have enough of spirit (*^^*)"). It seems that omitting zutto ("for ages") and adding the emoticon increases the emotional strength of the sentence.

8. Wider Implications of Deeper Affect Discovery in Utterances

The overwhelming and undisputed emotional nature of human communication requires an updated analysis through the new spaces created by social media, basically because emotions shape the way we use symbolic contents [18]. For example, the data we extracted for this research comes come from Twitter, and this as well as similar spaces follows special rules of emotional contagion and collective emotions [8, 13, 20]. Thus, detecting emotions in social media is one of the most important research areas in contemporary studies [57], and in some cases, the emotional analysis can have very useful consequences, like a fast alert detection thanks to confirmed social communication [39], or even political monitoring [33]. In a close future, humans not will only increase their electronic communications with other humans, but also with artificial cognitive systems like chatbots, artificial social agents or other AI systems [21, 52]. The reinforcement of emotional trustability is a key aspect of all human as well as human-machine interactions [14, 54], especially interesting for specific users like elderly [7]. Similarity network or hierarchical clustering, among other methods can be useful for the initial understanding of current human emotional expressions and the following dynamics, like emotion entrainment [15, 60]. Some possible different cultural uses of 'universal' emotional tools like emojis, are for example under debate [26]. These research attempts can help electronic developers avoid frequent problems as well as attract possible users to natural and comfortable environments. Anyhow, we must always to have in mind that users are not always coherent regarding what they think and do [51], especially when possible important emotive actions are involved into the communicative processes. This is, at the end, not only a necessity for linguistic analysis but for correct engineering, practical, and ethical implementations of electronic communication systems.

9. Overall 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%) 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% 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. When it comes to emoticon influence investigation, the used 88 utterances were semantically similar variants of the basic 22 originals and this number is not sufficient to draw indubitable conclusions. We see the results as useful not only for language analysis but also generation and we are planning to utilize the findings in our ongoing project for a dialog system with affective processing. We found out that ignoring emoticons may carry important consequences of incomplete emotion measurements and losing this semiotic information might be more costly than utilizing only adverbs weighting. Adverbs intensifying function differ from context to context so clustering methods for setting strengths as used by [48] must be learned on as big data sets as possible. Except increasing accuracy of the affect recognizer and testing different emotion categorizations as Plutchik's Wheel [29] and Geneva Emotion Wheel [40], we intend to test it in the chatbot environment with actual users participating conversation and to investigate how our idea can be extended to constantly changing contexts and their emotional fluctuations. When sufficient accuracy is achieved, we plan to test our combined algorithm as an automatic supervisor in a machine learning task and to investigate how it compares with human supervision.

References

[1] Benamara, Farah., Cesarano, Carmine., Picariello, Antonio., Reforgiato, Diego Recupero., Subrahmanian, Venkatramana, S. (2007). Sentiment analysis: Adjectives and adverbs are better than adjectives alone. *In* Proceedings of the International Conference on Weblogs and Social Media (ICWSM).

[2] Callahan, Sidney. (1988). The role of emotion in ethical decision making. Hastings Center Report, 18 (3) 9-14.

[3] Cambria, Erik. (2016). Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems*, 31 (2) 102- 107.

[4] Cambria, Erik., Fu, Jie., Bisio, Federica Poria, Soujanya. (2005). Affectivespace 2: Enabling affective intuition for concept-level sentiment analysis. *In*: AAAI, p. 508-514.

[5] Cambria, Erik., Schuller, Bjorn., Yunqing, Xia., Havasi, Catherine New avenues in opinion mining and sentiment analysis. Intelligent Systems, IEEE, 28 (2)15-21, March 2013.

[6] Andrea Ceron, Luigi Curini, Stefano Iacus, M. (2015). Using sentiment analysis to monitor electoral campaigns method matters-evidence from the United States and Italy. Social Science Computer Review, 33 (1) 3-20.

[7] Chattaraman, Veena., Kwon, Wi-Suk., Gilbert, Juan E. (2012). Virtual agents in retail web sites: Benefits of simulated social interaction for older users. *Computers in Human Behavior*, 28 (6) 2055-2066.

[8] Chmiel, Anna., Sienkiewicz, Julian., Thelwall, Mike., Paltoglou, Georgios., Buckley, Kevan., Kappas, Arvid., Janusz, A. (2011). Holyst. Collective emotions online and their influence on community life. PloS one, 6 (7) e 22207.

[9] Albert Costa, Alice Foucart, Sayuri Hayakawa, Melina Aparici, Jose Apesteguia, Joy Heafner, Boaz Keysar. (2014). Your morals depend on language. *PloS one*, 9 (4) e94842.

[10] Antonio, Damasio, R. (1999). The feeling of what happens: Body and emotion in the making of consciousness. Houghton Mifflin Harcourt.

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

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

[13] Ferrara, Emilio., Yang, Zeyao. (2015). Measuring emotional contagion in social media. *PloS one*, 10 (11) e0142390.

[14] Frances, S. (2015). Grodzinsky, KeithW. Miller, Marty J.Wolf. Developing automated deceptions and the impact on trust. *Philosophy & Technology*, 28 (1) 91-105.

[15] He, Saike., Zheng, Xiaolong., Zeng, Daniel Luo, Chuan Zhang, Zhu(2016). . Exploring entrainment patterns of human emotion in social media. *PloS one*, 11 (3) e0150630.

[16] Hiroshi, Kanayama., Tetsuya, Nasukawa., Hideo, Watanabe. (2004). Deeper sentiment analysis using machine translation technology. *In*: Proceedings of the 20th International Conference on Computational Linguistics, p. 494. Association for Computational Linguistics, 2004.

[17] Hogenboom, Alexander., Bal, Daniella., Frasincar, Flavius., Bal, Malissa de Jong, Franciska., Kaymak, Uzay (2013). Exploiting emoticons in sentiment analysis. *In*: Proceedings of the 28th Annual ACM Symposium on Applied Computing, SAC ‘13, p. 703-710, New York, NY, USA, 2013. ACM.

[18] Jimenez-Ortega, Laura., Martin-Loeches, Manuel., Casado, Pilar., Sel, Alejandra., Fondevila, Sabela., De Tejada, Pilar Herreros., Schacht, Annekathrin., Sommer, Werner. (2012). How the emotional content of discourse affects language comprehension. *PloS one*, 7 (3) e33718.

[19] Kawabata, Motoko. (2011). Study regarding perspectives on classifying expressions of degree (in Japanese). *In*: Technical report of Aichi Institute of Technology, vol. 47, p. 115-124.

[20] Kramer, Adam D.I., Guillory, Jamie E., Hancock, Jeffrey T.(2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111 (24) 8788-8790.

[21] Kugurakova, Vlada., Talanov, Maxim., Manakhov, Nadir., Ivanov, Denis. (2015). Anthropomorphic artificial social agent with simulated emotions and its implementation. *Procedia Computer Science*, 71. 112-118, 2015.

[22] Li, Xin Guang., Li, Su., Mei, Jiang., Li, Rui., Zhang., Sheng, Bin.(2014). Study of english pronunciation quality evaluation system with tone and emotion analysis capabilities. *In* :Applied Mechanics and Materials, volume 475, p. 318-323. Trans Tech Publ, 2014.

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

[24] Matsumoto, David., LeRoux, Jeff., Wilson-Cohn, Carinda., Raroque, Jake., Kooken, Kristie., Ekman, Paul., Yrizarry, Nathan., Loewinger, Sherry., Uchida, Hideko., Yee, Albert et al. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman’s Japanese and Caucasian Brief Affect Recognition Test (JACBART). Journal of Nonverbal Behavior, 24 (3) 179-209, 2000.

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

[26] Miller, Hannah Thebault-Spieker, Jacob., Chang, Shuo., Johnson, Isaac., Terveen, Loren., Hecht, Brent. “Blissfully happy” or “ready to fight”: Varying interpretations of emoji. ICWSM’16, 2016.

[27] Junko Minato, Kazuyuki Matsumoto, Fuji Ren, and Shingo Kuroiwa. Corpus-based analysis of japanese-english

emotional expressions. *In*: 2007 International Conference on Natural Language Processing and Knowledge Engineering, pages 413-418. IEEE, 2007.

[28] Akira Nakamura. (1993). Kanjo hyogen jiten [Dictionary of Emotive Expressions]. Tokyodo Publishing (1993).

[29] Plutchik, Robert. (1982). A psychoevolutionary theory of emotions. Social Science Information / Sur les sciences sociales, (1982).

[30] Ptaszynski, Michal., Dokoshi, Hiroaki., Oyama, Satoshi., Rzepka, Rafal., Kurihara, Masahito., Araki, Kenji., Momouchi, Yoshio. (2103). Affect analysis in context of characters in narratives. *Expert Systems with Applications*, 40 (1) 168- 176.

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

[32] Ptaszynski, Michal Dybala, Pawel Shi, Wenhan ., Rzepka, Rafal Araki, Kenji. (2009). 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, p. 1469-1474, Pasadena, July.

[33] Rheault, Ludovic., Beelen, Kaspar., Cochrane, Christopher., Hirst, Graeme. (2106). Measuring emotion in parliamentary debates with automated textual analysis. PloS one, 11 (12):e0168843.

[34] Russell, James A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39 (6) 1161-1178.

[35] Rzepka, Rafal., Araki, Kenji. (2005). What statistics could do for ethics? - The idea of common sense processing based safety valve. In Papers from AAAI Fall Symposium on Machine Ethics, FS-05-06, p. 85-87.

[36] Rzepka, Rafal., Jagla, Urszula., Dybala, Pawel., Araki, Kenji. (2016). Influence of emoticons and adverbs on affective perception of Japanese texts. *In*: Proceedings the 30th Annual Conference of the Japanese Society for Artificial Intelligence, 1K4-3H4-OS-17b-3.

[37] Rzepka, Rafal., Mazur, Micha.,l Clapp, Austin., Araki, Kenji. (2016). Global brain that makes you think twice. *In*: 2016 AAAI Spring Symposium Series.

[38] Saeki, Mika., Tokuhisa, Masato., Murakami, Jin'ichi., Ikehara, Satoru. (2003). Judgment of affective expressions by adverb and adjective (in Japanese). *In*: Forum on Information Technology, p. 117-118.

[39] Sasahara, Kazutoshi. (2014). Quantifying collective mood by emoticon networks. *In*: Proceedings of the 2014 *ACM conference on Web science*, p. 253-254. ACM.

[40] Scherer, Klaus R. (2005). What are emotions? And how can they be measured? *Social science information*, 44 (4) 695-729.

[41] Shi, Wenhan., Rzepka, Rafal., Araki, Kenji. (2008). Emotive information discovery from user textual input using causal associations from the Internet. *In*: Proceedings of the Forum on Information Technology (FIT-08), p.267 - 268.

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

[43] Solakidis, Georgios S., Vavliakis, Konstantinos N., Mitkas, Pericles A. (2014). Multilingual sentiment analysis using emoticons and keywords. *In*: Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)- 02, p. 102-109. IEEE Computer Society.

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

[45] Szemerey, Marton.(2012). Linguistic representation of emotions in Japanese and Hungarian: Quantity and abstractness. *Acta Linguistica Asiatica*, 2 (1) 61-72.

[46] Takamura, Hiroya., Inui, Takashi., Okumura, Manabu. (2005). Extracting semantic orientations of words using spin model. *In*: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, p. 133 -140. Association for Computational Linguistics.

[47] Takemoto, Sho., Tokuhisa, Masato., Murata, Masaki. (2012). Sentiment prediction using SO-score and patterns lexicon (in Japanese). *In*: Proceedings of the Twelve Annual Meeting of The Association for Natural Language Processing.

[48] Takizawa, Mitsuru., Rzepka, Rafal., Araki, Kenji. (2013). Evaluating polarity considering influence of adverbs (in Japanese). *In*: Technical Reports of Kotoba Kogaku, volume SIG-LSE-B302-3, p. 13-18.

[49] Turney, Peter D. (2002). 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, p. 417 - 424. Association for Computational Linguistics.

[50] Vallverdu, Jordi. (2013). An Ethic of Emotion. Amazon Digital Services LLC, ASIN: B00EFM7KMU.

[51] Vallverdu, Jordi. (2016). Ambient stupidity. *In*: Trends in Ambient Intelligent Systems, p.173-186. Springer.

[52] Vallverdu, Jordi., Casacuberta, David. (2015). Ethical and technical aspects of emotions to create empathy in medical machines. *In*: Machine Medical Ethics, p. 341-362. Springer.

[53] Jordi Vallverdu, Huma Shah, David Casacuberta. (2010). Chatterbox challenge as a test-bed for synthetic emotions. Creating Synthetic Emotions through Technological and Robotic Advancements, p. 118-144.

[54] Villata, Serena., Boella, Guido., Gabbay, Dov M., Van Der Torre, Leendert (2013). A socio-cognitive model of trust using argumentation theory. *Int. J. Approx. Reasoning*, 5 (4) 541-559, June.

[55] Bao-Khanh, H. (2013). Vo and Nigel Collier. Twitter emotion analysis in earthquake situations. *International Journal of Computational Linguistics Research*, 4 (1) 159-173.

[56] Wang, Xinyu., Zhang, Chunhong., Ji, Yang., Sun, Li., Wu, Leijia., Bao, Zhana (2013). A depression detection model based on sentiment analysis in micro-blog social network. *In*: Pacific-Asia Conference on Knowledge Discovery and Data Mining, p. 201-213. Springer.

[57] Wang Yichen., Pal, Aditya. (2015). Detecting emotions in social media: (2015). A constrained optimization approach. *In*: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015), p. 996-1002.

[58] CWhissell, Cynthia M. (1989). The Dictionary of Affect in Language. Emotion: Theory, Research and Experience: 4, The Measurement of Emotions. New York: Academic.

[59] Zhang, Peinan., Komachi. Mamoru (2015). Japanese sentiment classification with stacked denoising autoencoder using distributed word representation. *In*: Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation, PACLIC 29, Shanghai, China, October 30 - November 1.

[60] Zollo, Fabiana Kralj Petra., Novak, Michela Del Vicario., Bessi, Alessandro., Mozetic, Igor., Scala, Antonio., Caldarelli, Guido., Quattrociocchi, Walter. (2015). Emotional dynamics in the age of misinformation. *PloS one*, 10 (9) e0138740.

Sentence	Emoticon	Top Emotions
Iikagen kutsurogi-sugita-n-de sorosoro kaerimasu I've relaxed way too much, so I'm gonna go back home soon	(^^)	RELIEF / JOY
Mechakucha itoshii sugata desu-ne It's such an amazingly beloved appearance	(^_^)	LIKE / JOY
Unto tanoshinde kudasai Please have lots of fun	(^o^)	JOY / LIKE
Kyūkutsu deshita-ga, kawaikatta desu It looked a bit tight but it was cute	(≧ ω ≦)	LIKE / JOY
Zutto akogareteita-kara kiai jūbun Because I admired him for ages, I have enough of spirit	(*^^*)	JOY / LIKE
Dakara yokei shiawase-ni narunda-ne And that's why you become too happy	(*^∇^*)	JOY / RELIEF
Hannya-no Kawashima-san-ni ore-wa totemo tekii-o mochimashita I feel such a hostility toward Mr. "Hannya" Kawashima	(′ε′)	ANGER / DISLIKE
Nodo-ga kakudan-ni raku-ni narimashita My throat is distinctly better now	(^o^)	JOY / LIKE
Shōshō tanoshimi-ni shite ita-no da-ga, sorehodo-no kangeki-wa nakatta I was looking for it a bit but there was no deeper emotions	(^_^)	SADNESS / DISLIKE
Aida-ni shōchū hasandara choppiri fuwafuwa shite orimasu-ga Cause I had some shochu in between, I'm floating in the air	(●′∀′●)	JOY / LIKE
Kotoba-to kokoro-to karada-no okonai-ga icchi suru-to sugoku kokochi ii It is very pleasant when words and mind and actions of the body match	(^_^)	RELIEF / JOY
Kyō-wa saikō-ni shinpai-sugiru hi desu-ne Today is the day I am most worried	(‘ ;ω ; ‘)	SADNESS / DISLIKE
Kotoshi-wa kyonen-yori-mo yaru koto-ga muzukashiku nattete, wari-to tanoshimemashita-yo This year it was getting more difficult to do than last year, so I enjoyed it	(∧0∧)	JOY / LIKE
Saigo-de chotto sabishii-keredo tobikiri tanoshii jikan-ni shiyō I feel lonely a bit now at the end, but let's have lots of fun	(^_^)	SAD / JOY
Watashi-wa kekkō kō kan motemashita-yo I was quite nicely impressed	(^0^)	LIKE / JOY
Watashi-mo hijō -ni hazukashigari-na mono desukara! Actually I am very shy guy, too!	(*^_^*)	EXCITE / SHAME
Jūbun atsui-no-ga tsutawaru riakushon deshita-yo It was a reaction showing that he/she was quite heated	(^0^)	JOY / LIKE
Nebusoku-no karada-ni-wa nakanaka ureshii o-shokuji desu It is a very delightful meal to my body deprived of sleep	(^^)	JOY / RELIEF
Nagashima-ga yo-no naka-de mottomo osoreteiru no-wa konchū -to watashi-no gekido desu Insects and my rage are what Nagashima is the most afraid of in the world	(′;∀;′)	ANGER / EXCITE
Hikō ki-ni noru toki-wa toriwake tanoshii desu-ne It's especially fun when you take an airplane	(*^∇^*)	JOY / LIKE
Yachū-wa, tashō raku desu-ka-ne Is it a little easier in the middle of the night?	(*^∇^*)	JOY / LIKE
Odotteite danzen tanoshii desu It's such a great fun to dance	(^_^)	JOY / LIKE

Figure 16. Twenty two sentences used in our study with two top emotional category annotations (versions with adverbs and emoticons)