Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10)

CAO: A Fully Automatic Emoticon Analysis System

Michal Ptaszynski, Jacek Maciejewski, Pawel Dybala, Rafal Rzepka and Kenji Araki
Graduate School of Information Science and Technology, Hokkaido University
Kita-ku, Kita 14 Nishi 9, 060-0814 Sapporo, Japan
{ptaszynski,yaci,paweldybala,kabura,araki}@media.eng.hokudai.ac.jp

Abstract

This paper presents CAO, a system for affect analysis of emoticons. Emoticons are strings of symbols widely used in text-based online communication to convey emotions. It extracts emoticons from input and determines specific emotions they express. Firstly, by matching the extracted emoticons to a raw emoticon database, containing over ten thousand emoticon samples extracted from the Web and annotated automatically. The emoticons for which emotion types could not be determined using only this database, are automatically divided into semantic areas representing "mouths" or "eyes", based on the theory of kinesics. The areas are automatically annotated according to their co-occurrence in the database. The annotation is firstly based on the eye-mouth-eye triplet, and if no such triplet is found, all semantic areas are estimated separately. This provides the system coverage exceeding 3 million possibilities. The evaluation, performed on both training and test sets, confirmed the system's capability to sufficiently detect and extract any emoticon, analyze its semantic structure and estimate the potential emotion types expressed. The system achieved nearly ideal scores, outperforming existing emoticon analysis systems.

Introduction

One of the primary functions of the Internet is to connect people online. The first developed online communication media, such as e-mail or BBS forums, were based on text messages. Although later improvement of Internet connection allowed for phone calls or video conferences, the text-based message did not lose its popularity. However, its sensory limitations in communication channels (no view or sound) prompted users to develop communication strategies compensating for these limitations. One such strategy is the use of emoticons, strings of symbols imitating body language (faces or gestures). The use of emoticons contributes to the facilitation of the online communication process in e-mails, BBS or blogs (Suzuki and Tsuda, 2006; Derks 2007). Obtaining a sufficient computation level for this kind of communication would improve machine understanding of language used online, and contribute to the creation of more natural human-machine interfaces. Therefore, analysis of emoticons is of great importance in such fields as Human-Computer Interaction (HCI), Computational Linguistics (CL) or Artificial Intelligence (AI).

There have been several attempts to analyze Eastern type emoticons, which this paper focuses on. Tanaka et al. (2005) used kernel methods for extraction and classification of emoticons. However, their extraction was incomplete and the classification of emotion types incoherent and eventually set manually. Yamada et al. (2007) used statistics of n-grams. Unfortunately, their method was unable to extract emoticons from sentences. Moreover, they struggled with errors, as some characters were calculated as "eyes", although they represented "mouths", etc. Small coverage of emoticon databases in such research makes them inapplicable in affect analysis of the large numbers of original emoticons appearing on the Internet. All of the previous systems strictly depend on their primary emoticon databases and therefore are highly vulnerable to user creativity in generating new emoticons.

This paper presents CAO, a system dealing with most of those problems. The system extracts emoticons from input and classifies them automatically, taking into consideration semantic areas (representations of mouth, eyes, etc.). It is based on a large database collected from the Internet and improved automatically to coverage exceeding 3 million possibilities. The performance of the system is thoroughly verified with a training set and a test set based on a corpus of 350 million sentences in Japanese. The outline of the paper is as follows. Firstly, general terms related to the research described in this paper are defined. Secondly, database collection methods are explained and structure of the emoticon analysis system build on this database is presented. This is followed by description of experiments and achieved results. Finally, concluding remarks are presented and perspectives for future applications of the system are proposed.

Definitions

Classification of Emotions

We focused on emoticons used in online communication in Japanese. Therefore, we needed to choose the classification of emotions proven to be the most appropriate for the Japanese language. We applied the general definition of emotions as every temporary state of mind, feeling, or affective state evoked by experiencing different sensations (Lewis et al. 2008). As for the classification of emotions, we applied that of Nakamura (1993), who after over 30 years of study in the lexicography of the Japanese language and emotive expressions, distinguishes 10 emotion types as the most appropriate for the Japanese language and culture. These are: ki/yorokobi (joy, delight), do/ikari (anger), ai/aware (sorrow, sadness, gloom), fu/kowagari (fear), chi/haji (shame, shyness), ko/suki (liking, fondness), en/iya (dislike), ko/takaburi (excitement), an/yasuragi (relief) and kyo/odoroki (surprise, amazement). Emoticons in our research are then annotated according to this classification.

Definition of Emoticon

Emoticons have been used in online communication for over thirty years. Number of them has been developed, depending on the language of use, letter input system or the kind of community they are used in, etc. They can be roughly divided into: a) one-line Western and b) Eastern; and c) Multi-line ASCII art type. We focused on Eastern one-line emoticons[1], in Japan called *kaomoji*. Comparing to the western ones, these are usually unrotated. Some examples are: (^_^) (smiling face) or (ToT) (crying face). They are made of three to over twenty characters written in one line and are sophisticated enough to have a large number of variations expressing different meanings. Since emoticons are considered as representations of body language in text based conversation, their analysis can be based on approach similar to the one from the research on body language. In particular we apply the theory of kinesics to define semantic areas as separate kinemes, and then automatically assign to them emotional affiliations.

Theory of Kinesics

The word *kinesics* refers to all non-verbal behavior related to movement, such as postures, gestures and facial expressions, and functions as a term for body language in current anthropology. It is studied as an important component of nonverbal communication together with paralanguage (e.g. voice modulation) and proxemics (e.g. social distance). The term was coined by Birdwhistell (1952, 1970), who founded the theory of kinesics. The theory assumes that non-verbal behavior is used in everyday communication systematically and can be studied in a similar way to language. A minimal part distinguished in kinesics is a *kineme*, the smallest meaningful set of body movements, e.g. raising eyebrows, etc.

Emoticons from the Viewpoint of Kinesics. One of the current applications of kinesics is in annotation of affect display in psychology to determine which emotion is represented by which body movement or facial expression. Emoticons are representations of body language in online text-based communication. This suggests that the reasoning applied in kinesics is applicable to emoticons as well. Therefore, for the purposes of this research we defined emoticon as a one-line string of symbols containing at least one set of semantic areas, which we classify as: “mouth” [M], “eyes” [E_L], [E_R], “emoticon borders” [B_1], [B_2], and “additional areas” [S_1] - [S_4] placed between the above. Each area can include any number of characters. We also allowed part of the set to be of empty value, which means that the system could analyze an emoticon precisely even if some of the areas are absent. See Table 4 for examples of emoticons and their semantic areas. The analysis of emotive information of emoticons can therefore be based on annotations of the particular semantic areas grouped in an emoticon database.

Database of Emoticons

To create a system for emoticon analysis we first needed a coherent database of emoticons classified according to the emotions they represent. Firstly, a set of raw emoticons was extracted from seven online emoticon dictionaries[2].

Database Naming Unification

The data in each dictionary is divided into numerous categories, such as “greetings”, “hobbies”, “love”, “anger”, etc. However, the number of categories and their nomenclature is not unified. To unify them we used Ptaszynski et al.'s (2009a) affect analysis system. One of the procedures in this system is to classify words according to the emotion type they express, based on Nakamura's emotion classification. Categories with names suggesting emotional content were selected and emoticons from those categories were extracted, giving a total number of 10,137 unique emoticons.

Extraction of Semantic Areas

After gathering the database of emoticons and classifying them according to emotion types, we performed an extraction of all semantic areas appearing in unique emoticons. The extraction was done in agreement with the definition of emoticons and according to the following procedure. Firstly, possible emoticon borders are defined and all unique eye-mouth-eye triplets are extracted together (E_LME_R). From those triplets we extracted mouths (M) and pairs of eyes (E_L,E_R). Finally, having extracted the triplets and defined the emoticon borders (or their absence) we extracted all remaining additional areas (S1-S4).

Emotion Annotation of Semantic Areas

Having divided the emoticons into semantic areas, occurrence frequency of the areas in the emotion type database was calculated for every triplet, eyes, mouth and each of the additional areas. All unique areas were summarized in order of occurrence within the database for each emotion type. Each area's occurrence rate is considered as the probability of which emotion they tend to express.

[1] We did not focus on the other two for several reasons. Western emoticons, characteristic as rotated by 90 degrees, such as :-) (smiling face), or :-((sad face), have a simple structure of usually two to four characters and can be grouped in a list of no more than thirty, which makes them not challenging enough for language processing research. Moreover, we focused on emoticons used in Japanese online communities, where Western emoticons usually do not appear. Therefore, we excluded them from our research, although, a list of them can be simply added to our system in the end. Multi-line ASCII art type emoticons, consist of number of characters written in up to several dozens of lines. When looked at from a distance they make up a certain image, e.g. of a face. However, from the computational point of view, their multi-line structure makes their analysis more an image processing task. However, since some of them include also one-line emoticons, a system for processing of one-line emoticons could help process this type as well.

[2] Respectively: http://www.facemark.jp/facemark.htm, http://kaomojiya.com/, http://www.kaomoji.com/kao/text/, http://kaomoji-cafe.jp/, http://rsmz.net/kaopara/, http://matsucon.net/material/dic/, http://kaosute.net/jisyo/kanjou.shtml

Database Statistics

The number of unique E_LME_R triplets was 6,185. The number of unique eyes (E_L,E_R) and mouth areas (M) was 1,920 and 1,654, respectively. The number of unique additional areas was respectively S_1=5,169, S_2=2,986, S_3=3,192, S_4=8,837 (Overall 20,184). The distribution of all area types in the database is shown in Table 1.

Database Coverage

In previous research on emoticon classification one of the most popular approaches was the assumption that every emoticon is a separate entity, and therefore is not divided into separate areas or characters. However, this approach is strongly dependent on the number of emoticons in the database and is heavily vulnerable to user creativity in generating new emoticons. For example, emoticon databases in Yamada et al. (2007) had coverage of 693 emoticons, in Tanaka et al. (2005) it was 1075. The approach presented here assumes that emoticons can be analyzed more efficiently when divided into semantic areas. To confirm this, we estimated the coverage of the raw emoticon database (10,137) in comparison to the number of all possible combinations of triplets calculated as $E_L,E_R \times M$. Even excluding the additional areas, the number was equal to 3,175,680[3]. Therefore the coverage of the raw emoticon database, containing a somewhat large number of 10,137 samples, does not exceed 0.32% of the whole coverage of this method. This means that a method based only on a raw emoticon database would lose 99.68% of possible coverage, which in our approach is retained.

CAO - Emoticon Analysis System

The databases of emoticons and their semantic areas described above were applied in CAO, a system for *emotiCon Analysis and decOding of affective information*. It performs three main procedures. Firstly, it detects whether input contains any emoticons. Secondly, if emoticons were detected, the system extracts them from input and divides them into semantic areas. Thirdly, the system estimates the expressed emotions by matching the extracted emoticon in stages until it finds a match in the databases of: 1) raw emoticons, 2) E_LME_R triplets and additional areas S_1,…,S_4, and 3) Separately for the eyes E_L,E_R, mouth M and the additional areas.

Emoticon Detection in Input

The first procedure after obtaining input is responsible for detecting the presence of emoticons. The presence of an emoticon is determined when at least three symbols usually used in emoticons appear in a row. A set of 455 symbols was statistically selected as symbols appearing most frequently in emoticons.

Emoticon Extraction from Input

The extraction of emoticons from input is done in stages, looking for a match with: 1) the raw emoticon database; in

1.Input; (e.g.: ・゜・(ノД`;)・゜・)
2.Determine emotion types according to raw emoticon database; (・゜・(ノД`;)・゜・ : sorrow/sadness(3), excitement(2))
3.If no match,determine emotion types for E_LME_R triplet;
(ノД`:excitement(14),anger(2),sorrow(1),fear(1),joy(1),fondness(1))
4. If no emotion types for triplet found, find emotion types for separate semantic areas E_LE_R and M;
(ノ`:sorrow(3),shame(3),joy(2),fondness(2),fear(1),excitement(1),anger(1))
(Д:sorrow(53),excitement(52),anger(42),surprise(37), joy(28), fondness(25),dislike(22),fear(12),shame(9))
5. Determine emotion types for additional areas;
(・゜・:..., ;:..., ・゜・:...)
6. Proceed to next emoticon in the character string;
7.If no more emoticons, summarize scores;

Figure 1. The flow of the procedure for affect analysis of emoticon.

case of no match, 2) any ELMER triplet from the triplet database. If a triplet is found the system uses all databases of additional areas and emoticon borders and matches them to the regular expression: S1B1S2ELMERS3B2S4; 3) if the triplet match was not found, the system looks for: 3a) any triplet match from all 3 million ELMER combinations; or as a last resort 3b) a match for any of all areas separately.

Although the extraction procedure could function also as a detection procedure, it is more time consuming for the use of a regular expression. The differences in processing time are not noticeable when the number of consecutive inputs is small. However, we plan to use CAO to annotate large corpora including several million entries. With this code improvement the system skips sentences with no potential emoticons, which shortens the processing time.

Affect Analysis Procedure

In the affect analysis procedure, CAO estimates which emotion types an emoticon potentially could expresses. This is done by matching the recognized emoticon to the emotions annotated on database elements and checking their occurrence statistics. The system first checks which emotion types were annotated on raw emoticons. If no emotion was found, it looks for a match with emotion annotations with E_LME_R triplets. In case of no match, the semantic area databases for eyes E_LE_R and mouth M are considered separately and the matching emotion types are extracted. Finally, emotion type annotations for additional areas are determined. The flow of this procedure is shown with an example in Figure 1.

Output Calculation

After extracting emotion annotations for semantic areas,the final emotion ranking output is calculated. In the process of evaluation we calculated the score in three different ways to specify the most efficient method of result calculation.

Occurrence. Number of occurrences of an element (emoticon/triplet/semantic area). The higher occurrence hit-rate of an element in the emotion type database, the higher it scored. For more elements, the final score was calculated as the sum of all occurrence scores for all emotion types. The scores were placed in descending order of the final sums of their occurrences.

[3] However, including the additional areas gives an overall number of possible combinations equal to at least $1.382613544823877 \times 10^{21}$

Table 1. Distribution of all types of unique areas for which occurrence statistics were calculated across all emotion types in the database.

Area type	E_LME_R	S_1	S_2	E_LE_R	M	S_3	S_4
joy, delight	1298	1469	653	349	336	671	2449
anger	741	525	321	188	239	330	1014
sorrow	702	350	303	291	170	358	730
fear	124	72	67	52	62	74	133
shame	315	169	121	110	85	123	343
fondness	1079	1092	802	305	239	805	1633
dislike	527	337	209	161	179	201	562
excitement	670	700	268	243	164	324	1049
relief	81	50	11	38	26	27	64
surprise	648	405	231	183	154	279	860
Overall	**6185**	**5169**	**2986**	**1920**	**1654**	**3192**	**8837**

Frequency. Calculated as the occurrence number of a matched element (emoticon or semantic area) divided by the number of all elements in the particular emotion type database. The higher the frequency rate for a matched element in the emotion type database, the higher it scored. The final score for an emotion type was calculated as the sum of all frequency scores of the matched elements for an emotion type. The scores were placed in descending order of the final sums of their frequencies.

Unique Frequency. Calculated similarly to frequency. The difference is that the denominator (division basis) is not the number of all elements in the particular emotion type database, but the number of all unique elements.

Two-dimensional Model of Affect

According to Solomon (1993), people sometimes misinterpret specific emotion types, but rarely their valence. One might, for example, confuse anger and irritation, but it is unlikely they would confuse admiration with detestation. Therefore, we checked whether the general features of the extracted emotion types were in agreement. By "general features", we mean those proposed by Russell (1980) in his theory of a two-dimensional model of affect, where he argues that all emotions can be described in a space of two dimensions: valence and activation. An example of positive-activated emotion would be elation, positive-deactivated would be relief; negative-activated and negative-deactivated emotions would be indignation and depression, respectively. Nakamura's emotion types were mapped onto Russell's model and their affiliation to the spaces was determined as in Ptaszynski et al. (2009b). These groups are then used to estimate whether the emotions extracted by CAO belong to the same feature group. Figure 2 shows the details of the emotion type mapping.

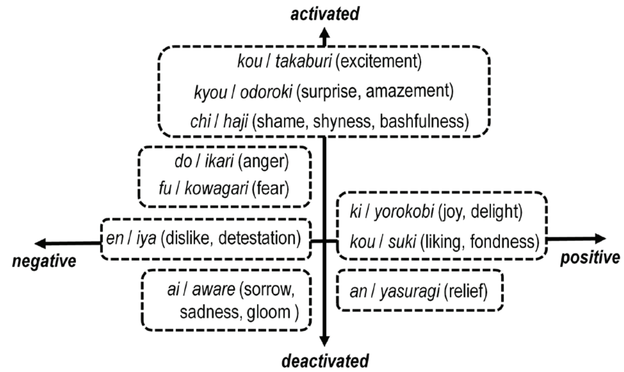

Figure 2: Grouping Nakamura's classification of emotions on Russell's two-dimensional space.

Evaluation of CAO

To fully verify the system's performance we carried out an exhaustive evaluation using a training set and a test set. The evaluated areas were: emoticon detection in a sentence, emoticon extraction from input, division of emoticons into semantic areas, and emotion classification.

Training Set Evaluation

The training set for the evaluation included all 10,137 unique emoticons from the emoticon database. However, to avoid perfect matching with the database (and therefore scoring 100% accuracy) we made the system skip the first step, matching input to the raw emoticon database, and continue with further procedures (matching triplets and separate semantic areas). The system's score was calculated as follows. If the system annotated an emoticon taken from a specific emotion type database with the name of the database as the highest one on the list of all annotated emotions, it counted as 1 point. Therefore, if the system annotated 5 emotion types on an emoticon taken from the "joy" database and the "joy" annotation appeared as the first one on the list of 5, the system's score was 1 point (5/5). However, if "joy" appeared on the second place, the score was 0.8 point (4/5), and so on.

Test Set Evaluation

In the test set evaluation we used Yacis Blog Corpus. It is an unannotated corpus consisting of 354,288,529 Japanese sentences. Average sentence length is 28.17 Japanese characters. Yacis Corpus was assembled using data obtained automatically from the pages of Ameba Blog (www.ameblo.co.jp), one of the largest Japanese blogging services. It consists of 12,938,606 downloaded and parsed web pages written by 60,658 unique bloggers. There are 6,421,577 pages containing 50,560,024 comments (7.873 comments per page that contains at least one comment). We used this corpus as it has been shown before that communication on blogs is rich in emoticons.

Experiment Settings

From Yacis Blog Corpus we randomly extracted 1000 middle-sized sentences as the test set. 418 of those sentences included emoticons. We calculated CAO's performance in detecting emoticons in sentences (with Cohen's agreement coefficient, kappa), and emoticon extraction, including division of emoticons into semantic areas (with balanced F-score). In the evaluation of the emotion estimation procedure 42 people annotated emotions on separate emoticons appearing in the sentences. This was used to verify system performance in specifying emotion types of particular emoticons. Additionally, the annotators annotated emotions on the whole sentences with emoticons (however, the emoticon samples were different). We did so in order to check how much of the emotive information encapsulated in a sentence is conveyed only with emoticons. Since meaning of written/typed sentences is perceived primarily on the basis of lexical information, we expected these results to be lower than those from only emoticon evaluation. The system's results were calculated considering human annotations as a gold standard. Moreover, we checked the results of annotations for specific emotion types and groups of emotions belonging to the same groups from Russell's two-dimensional affect space.

Comparing CAO with Other Systems

We also compared CAO to other emoticon analysis systems where possible. Tanaka et al.'s (2005) system was capable of simple emoticon extraction from sentences. Therefore emoticon extraction was compared to their system. Emotion estimation of emoticons was compared to the system developed by Yamada et al. (2007), as their approach is similar to ours in the method of exploiting the statistical occurrence of parts of emoticons. The two methods are described briefly below.

Kernel Method for Emoticon Extraction. The system for extraction and analysis of emoticons with kernel methods was proposed by Tanaka et al. (2005). They used popular tools for processing sentences in Japanese, a POS tagger *ChaSen* (Matsumoto et al. 2000) and a chunker *yamcha* (Kudo and Matsumoto, 2001). After chunking sentences they separated parts of speech from unrecognized elements, which they defined naively as emoticons. However, their method was significant as it was the first evaluated attempt to extract emoticons from lexical input. Unfortunately, the method was unable to extract emoticons from input other than a chunkable sentence (therefore, if a user was typing in a heightened emotional state and made a typo, the method would not activate). It was also not able to extract emoticons not containing both traditional emoticon borders (parenthesis). This made their method vulnerable to user creativity, although in a closed test on a set of prepared sentences their result was somewhat high (Precision=85.5%, Recall=86.7%, balanced F-score = 86%).

Their classification of emoticons into emotion types however, was not ideal. The set of six emotion types was determined manually and the classification process was based on a small sample set. Therefore as the system for comparison of emotion type classification we used a later one developed by Yamada et al. (2007).

N-gram Method for Emoticon Affect Estimation. Yamada and colleagues used statistics of n-grams to determine emotion types of emoticons. Although their method was not able to detect or extract emoticons from input, their set of emotion types was not set by the researchers, but borrowed from a classification appearing on BBS Web sites with emoticon dictionaries. Although not ideal, such classification was less subjective than their predecessors. To classify emoticons they used statistics of all characters occurring in emoticons without differentiating them into semantic areas. This caused errors, as some characters were calculated as "eyes", although they represented "mouths", etc. However, the accuracy of their method still achieved somewhat high scores of about 76% to 83%. For comparison with CAO we rebuilt this system and improved it with our emotion type classification (without this improvement their system would always score 0% for the lacking emotion types) and emoticon extraction from input, which capability the system did not posses. Moreover, we also used our database of raw emoticon samples, which improved the coverage of their system's database from 693 to 10,137. We used this system in evaluation of CAO to verify the performance of our system in comparison with other methods in the fairest way possible. We also used three versions of Yamada's system, based on unigrams, bigrams and trigrams.

Results and Discussion

Training Set Evaluation

Emoticon Extraction from Input. CAO extracted and divided into semantic areas a total number of 14,570 emoticons from the database of the original 10,137. The larger number of extracted emoticons on the output was caused by the fact that many emoticons contain more than one emoticon set. The results for emoticon extraction and division into semantic areas were as follows. CAO was able to extract and divide all of the emoticons, therefore the Recall rate was 100%. As for the Precision, 14,497 out of 14,570 were extracted and divided correctly, which gives the rate of 99.5%, with the balanced F-score equal to 99.75%. This score clearly outperforms Tanaka's system.

Affect Analysis of Emoticons. Firstly, we checked how many of the extracted emoticons were annotated at all. The only emoticons for which the system could not find any emotions were the 73 errors that appeared in the extraction evaluation. This means that the emotion annotation procedure was activated for all of the correctly extracted emoticons. Secondly, we calculated the accuracy in annotation of the particular emotion types on the extracted emoticons. All of the three methods for result calculation (Occurrence, Frequency and Unique Frequency) scored high, from over 80% to over 85%. The highest overall score in the training set evaluation was achieved by, in order: Occurrence (85.2%), Unique Frequency (81.8%) and Frequency (80.4%). Comparison with the other emoticon analysis system showed, that even after the improvements we made, the best score it achieved (80.2%) still did not exceed our worst score. For details see Table 2.

Table 2: Training set evaluation results for emotion estimation of emoticons for each emotion type with all five score calculations in comparison to another system for training set.

Emotion type	Yamada et al (2007) 1-gram	2-gram	3-gram	CAO: Occurrence	Frequency	Unique Frequency
anger	0.702	0.815	**0.877**	**0.811**	0.771	0.767
dislike	0.661	0.809	**0.919**	0.631	**0.800**	0.719
excitement	0.700	0.789	**0.846**	0.786	0.769	**0.797**
fear	**0.564**	0.409	0.397	0.451	**0.936**	0.858
fondness	**0.452**	0.436	0.448	**0.915**	0.778	0.783
joy	0.623	0.792	**0.873**	**0.944**	0.802	0.860
relief	**1.000**	0.999	**1.000**	0.600	**0.990**	0.985
shame	0.921	0.949	**0.976**	0.706	**0.922**	0.910
sorrow	0.720	0.861	**0.920**	**0.814**	0.809	0.791
surprise	0.805	0.904	**0.940**	0.862	0.866	**0.874**
All approx.	0.675	0.751	**0.802**	**0.852**	0.804	0.818

Test Set Evaluation

Emoticon Detection in Input. The system correctly detected the presence or absence of emoticons in 976 out of 1000 sentences (97.6%). In 24 cases (2.4% of all sentences) the system failed to detect that an emoticon appeared in the sentence. However, it achieved an ideal score in detecting the absence of emoticons. This means that there are no errors in the detecting procedure itself, but the database does not cover all possibilities of human creativity. The strength of the Cohen's coefficient of agreement with human annotators was considered to be very good (kappa=0.95).

Emoticon Extraction from Input. From 418 sentences containing emoticons CAO extracted 394 (Recall=94.3%). All of them were correctly extracted and divided into semantic areas (Precision=100%), which gave an overall extraction score of over 97.1% of balanced F-score. With such results the system clearly outperformed Tanaka et al.'s (2005) system in emoticon extraction and presented ideal performance in emoticon division into semantic areas, which capability was not present in the compared system.

Affect Analysis of Separate Emoticons. The highest score was achieved by, in order: Unique Frequency (93.5% for specific emotion types and 97.4% for groups of emotions mapped on Russell's two-dimensional affect space), Frequency (93.4% and 97.1%) and Occurrence (89.1% and 96.7%). The compared system by Yamada et al. (2007), despite the numerous improvements, did not score well, achieving its best score (for trigrams) below worst score of CAO (Occurrence/Types). The scores are shown in the top part of Table 3. The best score was achieved by Unique Frequency, which in training set evaluation achieved the second highest score. This method of score calculation will be therefore used as default score calculation in the system. However, to confirm this, we also checked the results of evaluation of affect analysis of sentences.

Affect Analysis of Emoticons in Sentences. The highest score was achieved by, in order: Unique Frequency (80.2% for specific emotion types and 94.6% for groups of emotions mapped on Russell's two-dimensional affect space), Frequency (80% and 94%) and Occurrence (75.5% and 90.8%). It is the same score order, although the evaluation was not of estimating separate emoticons, but the whole sentences. This proves that Unique Frequency is the most efficient method of output calculation for CAO. The compared system scored lower, achieving only one score (for bigrams) higher than our worst score (Occurrence/Types). The scores are shown in the bottom part of Table 3. Two examples of the results have been presented in Table 4.

The score for specific emotion type determination was, as we expected, not ideal (from 75.5% to 80.2%). This confirms that, using only emoticons, it is difficult to perform affect analysis of whole sentences with a result close to ideal. The reason for this is that the emotive information conveyed in sentences consists also of other lexical and contextual information. However, it can be still performed at a reasonable level (80.2%) and the results for two-dimensional affect space were close to ideal (nearly 95%). This means that the emotion types for which human annotators and the system did not agree still had the same general features (valence and activation). This also confirms the statement that people sometimes misinterpret (or use interchangeably) the specific emotion types of which general features are the same (in the test data people annotated, e.g., "fondness" on sentences with emoticons expressing "joy", etc., but never, e.g., "joy" on "fear").

Conclusions and Future Work

In this paper we presented a prototype system for automatic affect analysis of Eastern style emoticons, CAO. The system was created using a database of emoticons containing over ten thousand of unique emoticons collected from the Internet. These emoticons were automatically distributed into emotion type databases with the use of a previously developed affect analysis system. Finally, the emo-

Table 3: Results of the CAO system in Affect Analysis of emoticons in test set. The results summarize three ways of score calculation, specific emotions types and 2-dimensional affect space. The CAO system showed in comparison to another system.

Emotion Estimation on Separate Emoticons								
Yamada et al. (2007)			**CAO**					
			Occurrence		**Frequency**		**Unique Frequency**	
1-gram	**2-gram**	**3-gram**	**Types**	**2D space**	**Types**	**2D space**	**Types**	**2D space**
0.721347	0.865117	**0.877049**	0.891472	0.966778	0.934319	0.971044	**0.935364**	**0.973925**
Emotion Estimation on Sentences								
Yamada et al. (2007)			**CAO**					
			Occurrence		**Frequency**		**Unique Frequency**	
1-gram	**2-gram**	**3-gram**	**Types**	**2D space**	**Types**	**2D space**	**Types**	**2D space**
0.685714	**0.797659**	0.714819	0.755171	0.908911	0.800896	0.940582	**0.802012**	**0.946291**

ticons were divided into semantic areas, such as mouths or eyes and their emotion affiliations were calculated based on occurrence statistics. The division of emoticons into semantic areas was based on Birdwhistell's (1970) idea of kinemes as minimal meaningful elements in body language. The database of emoticon kinemes applied in CAO has coverage of over three million combinations. The number is sufficient enough to cover most emoticon possibilities used by users in online communication in Japanese.

The evaluation on both the training set and the test set showed that the system outperforms previous methods, achieving results close to ideal, and has other capabilities not present in its predecessors: detecting emoticons in input with very strong agreement coefficient (kappa = 0.95); and extracting emoticons from input and dividing them into semantic areas, with balanced F-score of over 97%. The system estimated emotions of separate emoticons with an accuracy of 93.5% for the specific emotion types and 97.3% for groups of emotions belonging to the same two dimensional affect space. Also in affect analysis of whole sentences CAO annotated the expressed emotions with a high accuracy of over 80% for specific emotion types and nearly 95% for two dimensional affect space, which means CAO can be used in this task, or as a support for other systems.

We plan to apply CAO to a number of tasks. Beginning with contribution to computer-mediated communication, we plan to make CAO a support tool for e-mail reader software. Although emoticons are used widely in online communication, there is still a wide spectrum of users (often elderly), who do not understand the emoticon expressions. Such users, when reading a message including emoticons, often get confused which causes future misunderstandings. CAO could help such users interpret the emoticons appearing in e-mails. As processing time in CAO is very short (processing of both training and test sets took no more than a few seconds), this application could be also extended to instant messaging services to help interlocutors understand each other in the text based communication. As a support system for Affect and Sentiment Analysis systems, such as (Ptaszynski et al. 2009a), CAO could also contribute to preserving online security (Ptaszynski et al. 2010), which has been an urgent problem for several years. Finally, we plan to thoroughly annotate large corpora of online communication, like Yacis Blog Corpus, to contribute to linguistic research on emotions in language.

Acknowledgements

This research is partially supported by a Research Grant from the Nissan Science Foundation and The GCOE Program from Japan's Ministry of Education, Culture, Sports, Science and Technology.

References

Birdwhistell, R. L. 1952. *Introduction to kinesics: an annotation system for analysis of body motion and gesture*, University of Kentucky Press.

Birdwhistell, R. L. 1970. *Kinesics and Context*, University of Pennsylvania Press, Philadelphia.

Derks, D., Bos, A.E.R., von Grumbkow, J. 2007. Emoticons and social interaction on the Internet: the importance of social context, *Computers in Human Behavior*, **23**, pp. 842-849.

Kudo, T. and Matsumoto, Y. 2001. Chunking with support vector machines, In *Proc. of NAACL-01*, pp. 192-199.

Lewis, M., Haviland-Jones, J. M., Feldman Barrett, L. (eds.). 2008. *Handbook of emotions*, Guilford Press.

Matsumoto, Y., Kitauchi, A., Yamashita, T., Hirano, Y., Matsuda, H., Takaoka, K. and Asahara, M. 2000. Japanese Morphological Analysis System ChaSen version 2.2.1.

Nakamura, A. 1993. *Kanjo hyogen jiten [Dictionary of Emotive Expressions]*(in Japanese), Tokyodo Publ., Tokyo.

Ptaszynski, M., Dybala, P., Matsuba, T., Masui, F., Rzepka, R. and Araki, K. 2010. Machine Learning and Affect Analysis Against Cyber-Bullying, In *Proc. of AISB'10*, *LaCATODA Symp.*, pp. 7-16.

Ptaszynski, M., Dybala, P., Rzepka, R. and Araki, K. 2009a. Affecting Corpora: Experiments with Automatic Affect Annotation System - A Case Study of the 2channel Forum, In *Proceedings PACLING-09*, pp. 223-228.

Ptaszynski, M., Dybala, P., Shi, W., Rzepka, R. and Araki, K. 2009b. Towards Context Aware Emotional Intelligence in Machines: Computing Contextual Appropriateness of Affective States, In *Proceedings of IJCAI-09*, pp. 1469-1474.

Russell, J. A. 1980. A circumplex model of affect, *J. of Personality and Social Psychology*, **39**(6), pp. 1161-1178.

Solomon, R. C. 1993. *The Passions: Emotions and the Meaning of Life*, Hackett Publishing.

Suzuki, N. and Tsuda, K. 2006. Automatic emoticon generation method for web community, *WBC2006*, pp. 331-334.

Tanaka, Y., Takamura, H., Okumura, M. 2005. Extraction and Classification of Facemarks with Kernel Methods, *IUI'05*, pp. 28-34.

Table 4: Examples of analysis performed by CAO. Presented abilities include: emoticon extraction, division into semantic areas, and emotion estimation in comparison with human annotations of separate emoticons and sentences. Emotion estimation (only the highest scores) given for Unique Frequency.

Example 1: *Itsumo, "Mac, ne-----" tte shibui kao sareru n desu. Windows to kurabete meccha katami ga semai desu* (ノ Д`):･° +:･。

Translation: People would pull a wry face on me saying "Oh, you're using a Mac…?" . It makes me feel so down when compared to Windows (ノ Д`):･° +:･。

S_1	**B_1**	**S_2**	**E_LME_R**	**S_3**	**B_2**	**S_4**
N/A	(	N/A	ノ Д`	N/A	)	:･° +:･。

CAO	**Human Annotation**	
	emoticon	**sentence**
sadness / sorrow (0.00698324)	sadness / sorrow	sadness / sorrow, dislike
...		

Example 2: *2000 bon anda wo tassei shita ato ni iroiro to sainan tsuzuita node nandaka o-ki no doku* ・・・(°. °)

Translation: All these sudden troubles, after scoring 2000 of safe hits. Unbelievable pity ・・・(°. °)

S_1	**B_1**	**S_2**	**E_L**	**M**	**E_R**	**S_3**	**B_2**	**S_4**
・・・	(	N/A	°	.	°	N/A	)	N/A

CAO	**Human Annotation**	
	emoticon	**sentence**
surprise (0.4215457)	surprise	surprise , dislike
...		