CAO: A Fully Automatic Emoticon Analysis System Based on Theory of Kinesics

Michal Ptaszynski, *Student Member*, *IEEE*, Jacek Maciejewski, Pawel Dybala, Rafal Rzepka, and Kenji Araki

Abstract—This paper presents CAO, a system for affect analysis of emoticons in Japanese online communication. Emoticons are strings of symbols widely used in text-based online communication to convey user emotions. The presented system extracts emoticons from input and determines the specific emotion types they express with a three-step procedure. First, it matches the extracted emoticons to a predetermined raw emoticon database. The database contains over 10,000 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 idea of *kinemes* from the theory of kinesics. The areas are automatically annotated according to their co-occurrence in the database. The annotation is first based on the eye-mouth-eye triplet, and if no such triplet is found, all semantic areas are estimated separately. This provides hints about potential groups of expressed emotions, giving 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.

Index Terms—Affect analysis, text processing, emotion in human-computer interaction, affect sensing and analysis, emoticon.

1 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 and popularization of Internet connections 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 of the interlocutors) 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). Today, the use of emoticons in online conversation contributes to the facilitation of the online communication process in e-mails, BBS, instant messaging applications, or blogs [1], [2], [3]. Obtaining a sufficient level of computation 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).

Emoticons are virtual representations of body language and their main function is similar, namely to convey information about the speaker's emotional state. Therefore, the analysis of emoticons appearing in online communication can be considered as a task for affect analysis, a subfield of AI focusing on classifying users' emotional expressions (e.g., anger, excitement, joy, etc.). There have been several approaches to analyzing emotive information conveyed by emoticons. For example, Tanaka et al. [4] used kernel methods for extraction and classification of emoticons, Yamada et al. [5] used statistics of n-grams, and Kawakami [6] gathered and thoroughly analyzed a database of 31 emoticons. However, all of these methods struggle with numerous problems, which include a lack of ability to precisely extract an emoticon from a sentence, incoherent emotion classification, manual and inconsistent emoticon sample annotation, inability to divide emoticons into semantic areas, small sample base, and therefore high vulnerability to user creativity in generating new emoticons.

This paper presents a system dealing with all 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: In Section 2, we present previous research concerning emoticons and describe some inadequacies of the previous emoticon analysis systems. Section 3 contains definitions and explanations of the nomenclature used in this paper. In Section 4, we explain the procedures applied during the automatic

- *The authors are with the Language Media Laboratory, Division of Media and Network Technologies, Graduate School of Information Science and Technology, Hokkaido University, Kita-14, Nishi-9, Kita-ku, Sapporo 060-0814, Japan. E-mail: {ptaszynski, yaci, paweldybala, kabura, araki}@media.eng.hokudai.ac.jp.*

Manuscript received 24 Feb. 2010; revised 25 June 2010; accepted 6 July 2010; published online 19 July 2010.
Recommended for acceptance by A. Hanjalic.
For information on obtaining reprints of this article, please send e-mail to: taffc@computer.org, and reference IEEECS Log Number TAFFC-2010-02-0014.
Digital Object Identifier no. 10.1109/T-AFFC.2010.3.

generation of the emoticon database and describe the structure and statistics concerning the database. Section 5 contains a description of CAO, the emoticon analysis system built on the database. In Section 6, we describe the evaluation settings for the system and present the results of the evaluation in Section 7. Finally, conclusions, future directions, and planned applications are presented in Section 8.

2 Previous Research

Research on emoticons has developed in three general directions. First, research in the fields of social sciences and communication studies have investigated the effects of emoticons on social interaction. There are several examples worth mentioning here. The research of Ip [7] investigates the impact of emoticons on affect interpretation in Instant Messaging. She concludes that the use of emoticons helps the interlocutors in conveying their emotions during the online conversation. Wolf [8] showed further, in her study on newsgroups, that there are significant differences in the use of emoticons by men and women. Derks et al. [2] investigated the influence of social context on the use of emoticons in Internet communication. Finally, Maness [9] performed linguistic analysis of chat conversations between college students, showing that the use of emoticons is an important means of communication in everyday online conversations. The above research is important in its investigation of the pragmatics of emoticons concerned as expressions of the language used online. However, most of such research focuses on Western-type emoticons.

Two practical applications of emoticon research in the field of Artificial Intelligence are to generate and analyze emoticons in online conversations in order to improve computer-related text-based communication, in fields such as Computer-Mediated Communication (CMC) or Human-Computer Interaction (HCI).

One of the first significant attempts at the first problem, emoticon generation, was by Nakamura et al. [10]. They used a Neural Networks-based algorithm to learn a set of emoticon areas (mouths, faces, etc.) and use them later in a dialogue agent. Unfortunately, the lack of a firm formalization of the semantic areas made the choice of emoticons eventually random and the final performance far from ideal. This was one of the reasons for abandoning the idea of exploiting parts of emoticons as base elements for emoticon-related systems. Since that time most of the research on emoticon generation has focused mostly on preprogrammed emoticons [1], [11], [12]. In our research, we revived the idea of exploiting the emoticon areas, although not in the research on emoticon generation but in emoticon extraction and analysis.

There have been several attempts to analyze emoticons or use them in affect analysis of sentences. For example, Reed [13] showed that the use of preprogrammed emoticons can be useful in sentiment classification. Yang et al. [14] made an attempt to automatically build a lexicon of emotional expressions using preprogrammed emoticons as seeds. However, both of the above researches focus only on preprogrammed Western-type emoticons, which are simple in structure. In our research, we focused on more challenging Eastern-type emoticons (for the description of types of emoticons, see the definition of emoticon in Section 3.2).

There have been three significant attempts to analyze Eastern emoticons. Tanaka et al. [4] used kernel methods for extraction and classification of emoticons. However, their extraction was incomplete and the classification of emotions incoherent and eventually set manually. Yamada et al. [5] used statistics of n-grams. Unfortunately, their method was unable to extract emoticons from sentences. Moreover, as they based their method on simple occurrence statistics of all characters in emoticons, they struggled with errors, as some characters were calculated as "eyes," although they represented "mouths," etc. Finally, Kawakami [6] gathered and thoroughly analyzed a database of 31 emoticons. Unfortunately, his analysis was done manually. Moreover, the small number of samples made his research 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.

In our research, we dealt with all of the above problems. Our system is capable of extraction of emoticons from input and fully automatic affect analysis based on a coherent emotion classification. It also takes into consideration semantic areas (representations of mouth, eyes, etc.). The system is based on a large emoticon database collected from the Internet and enlarged automatically, providing coverage of over 3 million possibilities. The system is thoroughly evaluated with a training set (the database) and a test set (a corpus of over 350 million sentences in Japanese). We summarize all of the previous research in comparison to our system in Table 2.

3 Definitions

3.1 Classification of Emotions

We focused on emoticons used in online communication in Japanese. Therefore, for the classification of emotions, we needed to choose the one 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 [15]. As for the classification of emotions, we applied that of Nakamura [16], who, after over 30 years of thorough 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 (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.

3.2 Definition of Emoticon

Emoticons have been used in online communication for many years and their numbers have developed depending on the language of use, letter input system, the kind of community they are used in, etc. However, they can be roughly divided into three types: 1) Western one-line type, 2) Eastern one-line type, and 3) multiline ASCII art type.

TABLE 1
Examples of Emoticon Division into Sets of Semantic Areas:
[M]—Mouth, $[E_L]$, $[E_R]$—Eyes, $[B_1]$, $[B_2]$—Emoticon Borders, $[S_1]$-$[S_4]$—Additional Areas

No. of sets	Emoticon	S_1	B_1	S_2	E_L	M	E_R	S_3	B_2	S_4	...
1	ヾ(。･ω･)ﾉ	ヾ	(	。	･	ω	･	N/A	)	ﾉ	
1	(——;)	N/A	(	N/A	—	N/A	—	;	)	N/A	
		SET 01									SET 02
2	(^^)人(^^)	N/A	(	N/A	^	N/A	^	N/A	)	人	(^^)
2	☆- (●≧ ▿)人(▿ ≦ ●)- ☆	☆-	(	●	≧	▿	N/A	N/A	)	人	(▿ ≦ ●)- ☆
					SET 01		SET 02		SET 03		SET 04
4	(∇ ´○)皿´★)ω´☆)∀ ´●)				(∇ ´○)		皿´★)		ω´☆)		∀ ´●)

Western emoticons exhibit characteristics as being rotated by 90 degrees, such as ":-)" (smiling face), or ":-D" (laughing face). They are the simplest of the three as they are usually made of two to four characters and are of a relatively small number. Therefore, we excluded them from our research as not being challenging enough to be a part of language processing. Moreover, our research focuses on the use of emoticons by Japanese users, and this type of emoticon is rarely used in Japanese online communities. However, as the Western-type emoticons can be gathered in a list of about 50, such a list could be simply added to our system at the end in a subprocedure.

Multiline ASCII art-type emoticons, on the other hand, consist of a number of characters written in several, or even up to several dozens of lines, which, when looked at from a distance, make up a picture, often representing a face or several faces. Their multiline structure leads their analysis to be considered more as a task for image processing than language processing, as this would be the only way for the computer to obtain an impression of the emoticon from a point of view similar to a user looking at the computer screen. Because of the above, we do not include multiline ASCII art emoticons in our research.

Finally, Eastern emoticons, in contrast to the Western ones are usually unrotated and present faces, gestures, or postures from a point of view easily comprehensible to the reader. Some examples are: "(^o^)" (laughing face), "(^_^)" (smiling face), and "(ToT)" (crying face). They arose in Japan, where they were called *kaomoji*, in the 1980s and since then have been developed in a number of online communities. They are made up of three to over 20 characters written in one line and consist of a representation of at least one face or posture, up to a number of different face-marks. In the research described in this paper, we focused mainly on this type of emoticon, as they have a large variation of appearance and are sophisticated enough to express different meanings. See Table 1 for some examples of this type of emoticon.

Emoticons defined as above can be considered as representations of body language in text-based conversation, where the communication channel is limited to the transmission of letters and punctuation marks. Therefore, we based our approach on the analysis of emoticons on assumptions similar to those from research on body

TABLE 2
Previous Research on Emoticon Analysis in Comparison to Our System

Research (approach) / Capability	Tanaka et al (2005) (kernel methods)	Yamada et al (2007) (n-grams)	Kawakami (2008) (database)	CAO (theory of kinesics)
1. Detection whether input equals emoticon	×	×	×	○
2. Detection of emoticon in sentence input	○ (included in **3.**)	×	×	○
3. Extraction of emoticon from any other string of characters	○	×	×	○
4. Division into semantic areas	×	×	×	○
5. Database coverage	1,075	693	31	10,137 expanded automatically to over 3 million
6. Classification of emotion types	6 types (Subjective)	7 types (BBS-based; Subjective)	6 types (Subjective)	10 types (Language/ Culture Based)
7. Emotion estimation of separate emoticons	○ (included in **8.**)	○	○	○
8. Affect Analysis of sentences with emoticons	○	×	×	○

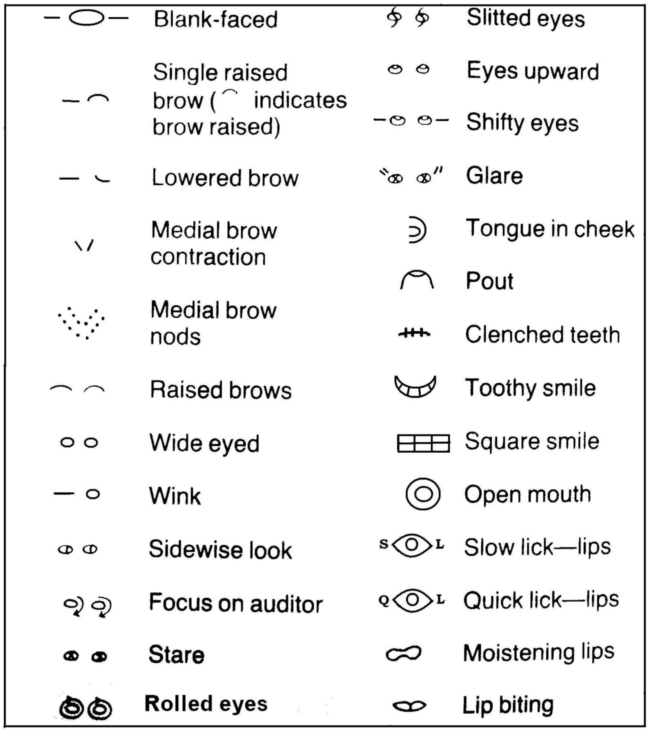

Fig. 1. Some examples of kinegraphs used by Birdwhistell to annotate body language.

language. In particular, we apply the theory of kinesics to define semantic areas as separate kinemes, and then automatically assign to them emotional affiliation.

3.3 Theory of Kinesics

The word *kinesics*, as defined by Vargas [17], refers to all nonverbal 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 first used by Birdwhistell [18], [19], who founded the theory of kinesics. The theory assumes that nonverbal 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 movement, e.g., raising eyebrows or moving the eyes upward. Birdwhistell developed a complex system of *kinegraphs* to annotate kinemes for the research on body language. Some examples of kinemes are given in Fig. 1.

3.3.1 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 purpose of this research, we specified the definition "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 can analyze an emoticon precisely even if some of the areas are absent. The minimal emoticon set considered in this research comprises two eyes (a set represented as "E_L, E_R," e.g., "^ ^" (a happy face)), mouth and an eye ("E_L, M" or "M, E_R," e.g., "(^○)" (a laughing face) and "(_^)" (a smiling face), respectively), or mouth/eye with one element of the additional areas ("$M/E_R, S_3/S_4$" or "$S_1/S_2, E_L/M$," e.g., "(^)/ ~ " (a happy face) and "\(`)" (a sad face), respectively). However, many emoticons contain all or most of the areas, as in the following example showing a crying face. "·゜・(ノД`;)・゜・." See Table 1 for some examples of emoticons and their semantic areas. The analysis of emotive information conveyed in emoticons can therefore be based on annotations of the particular semantic areas grouped in an automatically constructed emoticon database.

4 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. The database development was performed in several steps. First, raw emoticon samples were collected from the Internet. Then, the naming of emotion classes expressed by the emoticons was unified according to Nakamura's [16] classification of emotions. Next, the idea of kinemes was applied in order to divide the extracted emoticons into semantic areas. Finally, the emotive affiliations of the semantic areas were determined by calculating their occurrences in the database.

4.1 Resource Collection

The raw emoticons were extracted from seven online emoticon dictionaries available on seven popular Web pages dedicated to emoticons: Face-mark Party, Kaomojiya, Kaomoji-toshokan, Kaomoji-café, Kaomoji Paradise, Kaomojisyo, and Kaomoji Station.[1] The dictionaries are easily accessible from the Internet.

4.2 Database Naming Unification

The data in each dictionary is divided into numerous categories, such as "greetings," "affirmations," "actions," "hobbies," "expressing emotions," etc. However, the number of categories and their nomenclature is not unified. To unify them, we used Ptaszynski et al.'s [20] 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 of 11,416 emoticons. However, as some of them could appear in more than one collection, we performed filtering to extract only the unique ones. The number of unique emoticons after the filtering was 10,137 (89 percent). Most of the emoticons appearing in all seven collections were unique. Only for the emoticons annotated as expressions of "joy" was a large amount, over one-third, repeated. This means that all of the

1. 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.

TABLE 3
Ratio of Unique Emoticons to All Extracted Emoticons and Their Distribution in the Database According to Emotion Types

joy, delight	liking, fondness	anger	surprise, amazement	sadness, gloom	excite-cite-ment	dis-like	shame, shyness	fear	relief	Over-all	Emoticons
3128	1988	1238	1227	1203	1124	704	526	179	99	**11416**	All extracted
1972	1972	1221	1196	1169	1120	698	511	179	99	**10137**	Unique
63%	99%	99%	97%	97%	99%	99%	97%	100%	100%	**89%**	Ratio

dictionaries from which the emoticons were extracted provided emoticons that did not appear in other collections. On the other hand, the high repeating frequency of emoticons annotated as expressions of "joy" suggests that this emotion type is expressed by Internet users with a certain number of popular emoticons. The emotion types for which the number of extracted emoticons was the highest were, in order, joy, fondness, anger, surprise, gloom, and excitement. This suggests that Internet users express these emotion types more often than the rest, which were, in order, dislike, shame/bashfulness, fear, and relief. The ratio of unique emoticons to all extracted ones and their distribution across the emotion types are shown in Table 3.

4.3 Extraction of Semantic Areas

After gathering the database of raw 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. First, 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). The rule for extracting eye-patterns from triplets goes as follows: If the eyes consist of multiple characters, each eye has the same pattern. If the eyes consist only of one character, they can be the same or different (this was always true among the 10,137 emoticons in our database). Finally, having extracted the E_LME_R triplets and defined the emoticon borders, we extracted all existing additional areas (S1, ..., S4). See Fig. 2 for the details of this process.

1.Input: (.*) - STRING OF CHARACTERS
2.Determine emoticon borders:
B_1\{NULL,(, (,<,<,... \}, B_2\{NULL,>,>) ,),...\}:B_1(.*)B_1
3.Localize E_LME_R triplet in the potential emoticon: B_1(.*)E_LME_R(.*)B_2;
4.Separate eyes E_L,E_R and mouth M areas:
5. from E_LME_R take *n* characters from the left n_L and right n_R;
6. if n_L=n_R, n_L is E_L and n_R=E_R; if no match take n-1 characters;
7. if the above fails, take one character form left as E_L and from right as E_R.
8. mouth area M is what is left between E_L and E_R
9.Determine additional areas S_1,...,S_4 according to the regular expression: $S_1B_1S_2E_LME_RS_3B_2S_4$
10.Calculate occurrence frequencies separately for trip-let E_LME_R, pair E_L,E_R, mouth M, additional areas S_1,...,S_4, for all emotion types;

Fig. 2. The flow of the procedure for semantic area extraction.

4.4 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 the 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.

4.5 Database Statistics

The number of unique combined areas of E_LME_R triplets was 6,185. The number of unique eyes (E_L,E_R) was 1,920. The number of unique mouth areas (M) was 1,654. The number of unique additional areas was, respectively, $S_1 = 5{,}169$, $S_2 = 2{,}986$, $S_3 = 3{,}192$, and $S_4 = 8{,}837$ (overall 20,184). The distribution of all area types for which the statistics were calculated is shown in Table 4.

4.6 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 [6]. 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. We aimed at developing an approach as much immune to user creativity as possible. To verify that, we estimated the coverage of the raw emoticon database in comparison to the database of all semantic areas separately. The number of all possible combinations of triplets calculated as $E_L, E_R \times M$, even excluding the additional areas, is equal to 3,175,680 (over 3 million combinations[2]). Therefore, the basic coverage of the raw emoticon database, which contains a somewhat large number of 10,137 unique samples, does not exceed 0.32 percent of the whole coverage of this method. This means that a method based only on a raw emoticon database would lose 99.68 percent of possible coverage, which is retained in our approach.

5 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*. The system performs three main procedures. First, it detects whether input contains any emoticons. Second, if emoticons

2. However, including the additional areas in the calculation gives an overall number of possibilities equal to at least $1.382613544823877 \times 10^{21}$.

TABLE 4
Distribution of All Types of Unique Areas for which Occurrence Statistics Were Calculated across All Emotion Types in the Database

Area type	$E_L M E_R$	S_1	B_1	S_2	E_L E_R	M	S_3	B_2	S_4
joy, delight	1298	1469	--	653	349	336	671	--	2449
anger	741	525	--	321	188	239	330	--	1014
sorrow, sadness, gloom	702	350	--	303	291	170	358	--	730
fear	124	72	--	67	52	62	74	--	133
shame, shyness	315	169	--	121	110	85	123	--	343
liking, 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, amazement	648	405	--	231	183	154	279	--	860
Overall	**6185**	**5169**	**--**	**2986**	**1920**	**1654**	**3192**	**--**	**8837**

were detected, the system extracts all emoticons from the input. Third, 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. $E_L M E_R$ triplets and additional areas $S_1, \ldots, S_4$,
3. separately for the eyes E_L, E_R, mouth M, and the additional areas.

5.1 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.

5.2 Emoticon Extraction from Input

In the emoticon extraction procedure, the system extracts all emoticons from input. This is done in stages, looking for a match with: 1) the raw emoticon database, in case of no match, 2) any $E_L M E_R$ triplet from the triplet database. If a triplet is found, the system matches the rest of the elements of the regular expression: m/[S_1?] [B_1?] [S_2?] [E_L M E_R] [S_3?] [B_2?] [S_4?]/, with the use of all databases of additional areas and emoticon borders, 3) in case the triplet match was not found, the system searches for: 3a) any triplet match from all 3 million $E_L M E_R$ combinations with one of the four possible $E_L M E_R$ patterns matched gradually ([E_L] [M] [E_R], [E_L] [E_R], [M] [E_R], [E_L] [M]), or as a last resort 3b) a match for any of all the areas separately. The flow of this procedure is represented in Fig. 3.

1.Input: "SOME CHARACTERS ･ﾟ(･ﾉД;)･ﾟ･ SOME CHARACTERS"
2.Find match in raw emoticon database: ･ﾟ(･ﾉД;)･ﾟ･
3. If no match, localize $E_L M E_R$ triplet in the $E_L M E_R$ triplet database: ﾉД
4. If no triplet found, look for any $E_L M E_R$ combination;
5. If no combination matched, find any $E_L E_R$ or M from separate semantic area database: ﾉ`, Д
6.Localize emoticon borders B_1,B_2 : (,)
7.Localize additional areas S_1,S_2,S_3,S_4 : ･ﾟ,･;,･ﾟ･
8.Determine the emoticon structure: S_1: ･ﾟ,･ B_1:(, S_2:N/A, $E_L E_R$: ﾉ`, M: Д, S_3:;, B_2:), S_4: ･ﾟ･
9.Look for next emoticon;

Fig. 3. The flow of the procedure for emoticon extraction.

Although the extraction procedure could also function as a detection procedure, it is more time-consuming. 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 over several million entries. With this code improvement, the system skips sentences with no potential emoticons, which shortens the processing time.

5.3 Affect Analysis Procedure

In the affect analysis procedure, the system estimates which emotion types are the most probable for an emoticon to express. This is done by matching the recognized emoticon to the emotions annotated on the database elements and checking their occurrence statistics. This procedure is performed as an extension to the extraction procedure. 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 $E_L M E_R$ triplet. If no match was found, the semantic area databases for eyes $E_L E_R$ 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 Fig. 4. The flow chart of the whole system is presented in Fig. 5.

5.4 Output Calculation

After extracting the emotion annotations of emoticons and/or semantic areas, the final emotion ranking output was calculated. In the process of evaluation, we calculated the score in five different ways to specify the most effective method of result calculation.

5.4.1 Occurrence

The processing of one emoticon provides a set of lists—one for each emoticon part (mouth, eyes, additional areas, etc.). Any part of emoticon may appear in databases belonging to

1.Input; (e.g.: ･ﾟ･(ﾉД`;)･ﾟ･)
2.Determine emotion types according to raw emoticon data-
base; (･ﾟ･(ﾉД`;)･ﾟ･ : 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;
(･ﾟ･:..., ;:..., ･ﾟ･:...)
6. Proceed to next emoticon in the character string;
7.If no more emoticons, summarize scores;

Fig. 4. The flow of the procedure for affect analysis of emoticon.

different emotion types (e.g., in the crying emoticon, ･ﾟ･(ﾉД`;)･ﾟ･, element representing "mouth"—appears 53 times in the sorrow database, 52 times in excitement, 28 times in joy, etc. (see Fig. 4 for details). Each of those lists contains emotion types with assigned numbers of occurrences of the element in the database of each emotion type. Having these lists, it is possible to perform different calculations to summarize/generalize them. First, all results can be added and then the emotion type appearing most often will be the most probable for the emoticon to express. In other words, occurrence is the straightforward number of occurrences of an element (emoticon/triplet/semantic area). The higher the occurrence of an element in the emotion-type database, the higher it scored. For more elements, the final score for an emotion type was calculated as the sum of all occurrence scores for all emotion types. The final emotion scores were placed in descending order of the final sums of their occurrences.

5.4.2 Frequency

However, it might be said that to simply add the numbers is not a fair way of score summarization since there are a different number of elements in each database and a database with a small number of elements will have a tendency to lose. To avoid such biases, we divided the emotion score by the number of all elements in the database. Therefore, frequency is 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. For more elements, 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 final scores for each emotion type were placed in descending order of the final sums of their frequencies.

5.4.3 Unique Frequency

It could be further said that just dividing by all elements is also not ideally fair since there are elements appearing more often which are therefore stronger, which will also cause a bias in the results. To avoid this, we also divided the occurrences by the number of all unique elements. Unique frequency is thus calculated similarly to the usual 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 ones.

5.4.4 Position

Position is calculated in the following way. The strings of characters in all databases (raw emoticons, triplets, and

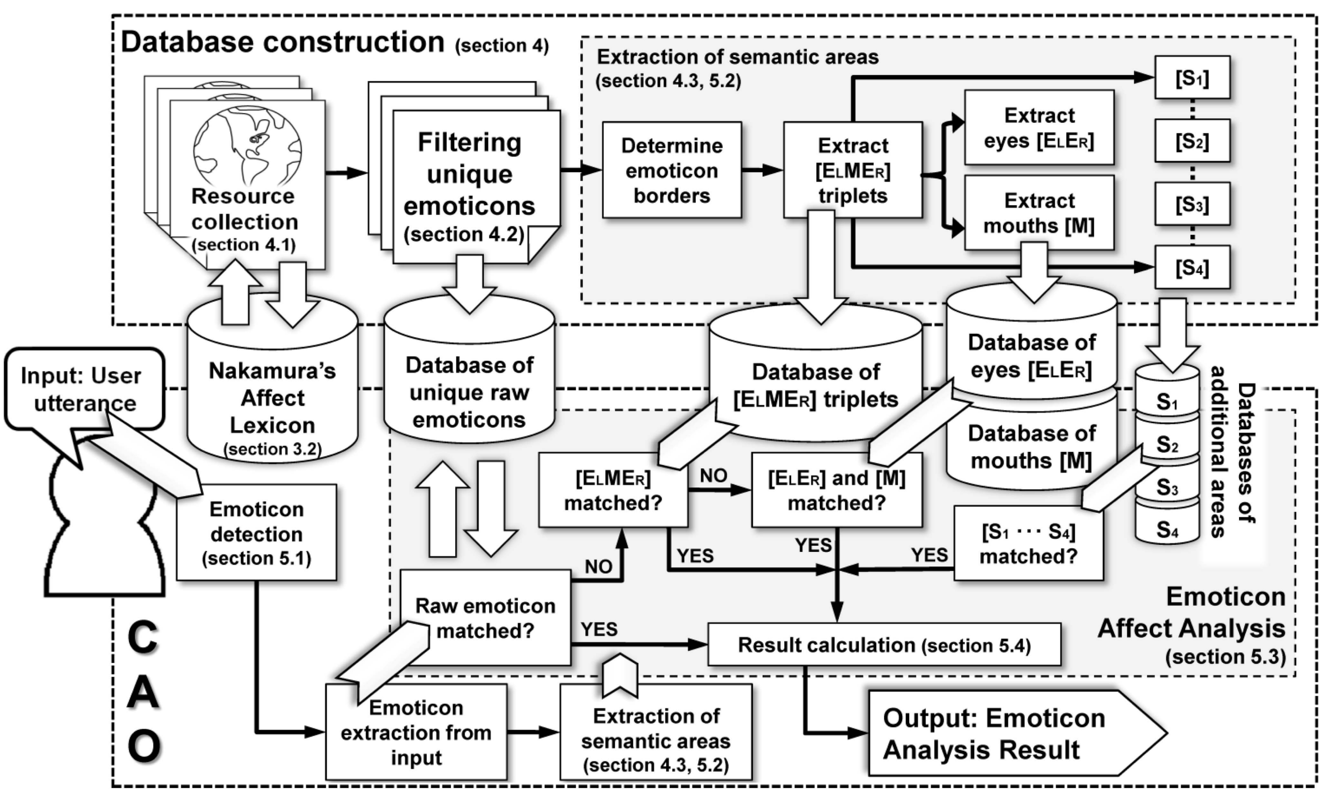

Fig. 5. Flow chart of the database construction and the CAO system.

semantic areas) are sorted by their occurrence in descending order. By position, we mean the place of the matched string in the database. Position is determined by the number of strings, occurrence of which was greater than the occurrence of a given string. For example, in a set of strings with the following occurrences: $n_1 = 5$, $n_2 = 5$, $n_3 = 5$, $n_4 = 3$, $n_5 = 3$, $n_6 = 2$, $n_7 = 2$, $n_8 = 1$, the strings n_6 and n_7 will be in the sixth position. If the string was not matched in a given database, it is assigned a position of the last plus one element from this database.

5.4.5 Unique Position

Unique Position is calculated in a similar way to the normal Position, with one difference. Since some strings in the databases have the same number of occurrences, they could be considered as appearing in the same position. Therefore, here we considered the strings with the same occurrences as the ones with the same position. For example, in a set of strings with the following occurrences: $n_1 = 5$, $n_2 = 5$, $n_3 = 5$, $n_4 = 3$, $n_5 = 3$, $n_6 = 2$, $n_7 = 2$, $n_8 = 1$, the strings n_6 and n_7 will be in the third position. If the string was not matched in a given database, it is assigned a position of the last plus one element from this database.

5.5 Two-Dimensional Model of Affect

According to Solomon [21], people sometimes misinterpret specific emotion types, but rarely their valence. One might, for example, confuse such emotions as 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 [22] in his theory of a 2D 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 [23]. For some emotion types, the affiliation is somewhat obvious, e.g., gloom is never positive or activated. However, for other emotion types, the emotion affiliation is not that obvious, e.g., surprise can be both positive as well as negative, dislike can be either activated or deactivated, etc. The emotion types with uncertain affiliation were mapped on all groups they could belong to. However, no emotion type was mapped on more than two adjacent fields. These groups are then used for estimating whether the emotion types extracted by CAO belong to the same quarter. For the details of the mapping of the emotion types, see Fig. 6.

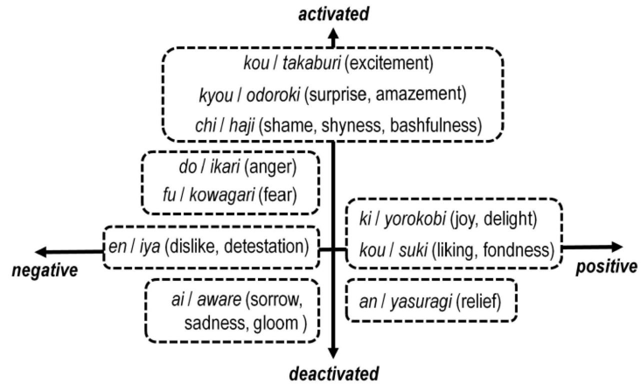

Fig. 6. Grouping Nakamura's classification of emotions on Russell's 2D space.

6 Evaluation of CAO

To fully verify the system's performance, we carried out an exhaustive evaluation. The system was evaluated 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 of emoticons.

6.1 Training Set Evaluation

The training set for the evaluation included all 10,137 unique emoticons from the raw emoticon database. However, to avoid perfect matching with the database (and therefore scoring 100 percent accuracy), we made the system skip the first step—matching 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 five 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 5/5 (1 point). If the name of the emotion database from which the emoticon was taken did not appear in the first place, the score was calculated as the rank number the emotion achieved divided by the number of all emotions annotated. Therefore, if the system annotated five emotion types on an emoticon taken from the "joy" database and the "joy" annotation appeared as the second one on the list of five, the system's score was 4/5 (0.8 point), and so on. These calculations were further performed for all five ways of score calculation.

6.2 Test Set Evaluation

In the test set evaluation, we used Yacis Blog Corpus.

6.2.1 Yacis Blog Corpus

Yacis Blog Corpus is an unannotated corpus consisting of 354,288,529 Japanese sentences. Average sentence length is 28.17 Japanese characters, which fits in the definition of a short sentence in the Japanese language [24]. 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 were 6,421,577 pages containing 50,560,024 comments (7.873 comments per page that contains at least one comment). All pages were obtained between 3rd and 24th of December 2009. We used this corpus as it has been shown before that communication on blogs is rich in emoticons.

6.2.2 Experiment Settings

From Yacis Blog Corpus, we randomly extracted 1,000 middle-sized sentences as the test set; 418 of those sentences included emoticons. Using Cohen's kappa agreement coefficient and balanced F-score, we calculated CAO's performance in detecting emoticons in sentences (with Cohen's agreement coefficient and kappa), and emoticon extraction (including division of emoticons into semantic areas). In the evaluation of the emotion estimation procedure, we asked 42 people to annotate emotions on separate emoticons appearing in the sentences to verify the performance of CAO in specifying emotion types conveyed by particular emoticons (each person annotated 10 sentences/emoticons, except one person, who annotated eight samples). Additionally, we asked the annotators to annotate emotions on the whole sentences with emoticons (however, the emoticon samples appearing in the sentences were different from the ones assigned in only emoticon annotation). This was used in an additional experiment not performed before in other research on emoticons. The usual evaluation only considers recognizing emotions of separate emoticons. We wanted to check how much of the emotive information encapsulated in a sentence could be conveyed with the addition of emoticons and whether it is possible to recognize the emotion expressed by the whole sentence looking only at the emoticons used in the sentence. Emoticons are something like an addition to this meaning. The question was how much does the emoticon match the meaning expressed by the sentence? We checked this appearance of the emotion types and the general emotive features (valence and activation). However, the meaning of written/typed sentences is mostly understood on the basis of lexical information, and we expected these results to be lower than those from only emoticon evaluation.

The system's results were calculated in a similar way to the training set, 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 quarters from Russell's 2D affect space. The calculations were performed for the best three of the five ways of score calculation selected in training set evaluation.

6.3 Comparing CAO with Other Systems

We also compared CAO to other emoticon analysis systems where possible. The emoticon extraction was compared to the system developed by Tanaka et al. [4]. Emotion estimation of emoticons was compared to the system developed by Yamada et al. [5], as their approach is similar to ours in the method of exploiting the statistical occurrence of parts of emoticons. The two methods are described in detail below.

6.3.1 Kernel Method for Emoticon Extraction

The system for extraction and analysis of emoticons with kernel methods was proposed by Tanaka et al. [4]. In their method, they used popular tools for processing sentences in Japanese, a POS tagger, ChaSen [25], and a Support Vector Machine-based chunker, *yamcha* [26], to chunk sentences and separate parts of speech from "other areas in the sentence." which they defined as potential emoticons. However, their method was significant as it was the first evaluated attempt to extract emoticons from input. Unfortunately, the method was unable to perform many important tasks. First, as the method is based on a POS tagger, it could not extract emoticons from input other than a chunkable sentence. Therefore, if their system got a nonchunkable input (e.g., a sentence written in a hurry, with spelling mistakes, etc.), the method would not be able to proceed or would give an erroneous output. Moreover, if a spelling mistake appeared inside a parenthesis, a nonemoticon content could be recognized as a potential emoticon. All this made their method highly vulnerable to user creativity, although in a closed test on a set of prepared sentences their best result was somewhat high with 85.5 percent of Precision and 86.7 percent of Recall (balanced F-score = 86 percent).

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. [5].

6.3.2 N-Gram Method for Emoticon Affect Estimation

Yamada et al. [26] used statistics of n-grams to determine emotion types conveyed by 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 simple statistics of all characters occurring in emoticons without differentiating them into semantic areas. Eventually, this caused errors, as some characters were calculated as "eyes" wvwn though they represented "mouths," etc. However, the accuracy of their method still achieved somewhat high scores of about 76-83 percent. For comparison with CAO, we built a second system similar to theirs, but improved it with our emotion-type classification (without this improvement, in our evaluation, their system would always score 0 percent for the lacking emotion types) and emoticon extraction from input, which capability the system of Yamada et al. did not possess. Moreover, we also used our database of raw emoticon samples, which improved the coverage of their system's database to 10,137 from 693 (6.8 percent of the improved database). Improved this way, 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.

7 Results and Discussion

7.1 Training Set Evaluation

7.1.1 Emoticon Extraction from Input

The system 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 (see the example in

TABLE 5
Training Set Evaluation Results for Emotion Estimation of Emoticons for Each Emotion Type with All Five Score Calculations in Comparison to Another System

Emotion type	Yamada et al (2007) improved 1-gram	2-gram	3-gram	CAO: Occurrence	Frequency	Unique Frequency	Position	Unique Position
anger	0.702	0.815	**0.877**	**0.811**	0.771	0.767	0.476	0.476
dislike	0.661	0.809	**0.919**	0.631	**0.800**	0.719	0.556	0.591
excitement	0.700	0.789	**0.846**	0.786	0.769	**0.797**	0.560	0.516
fear	**0.564**	0.409	0.397	0.451	**0.936**	0.858	0.652	0.671
fondness	**0.452**	0.436	0.448	**0.915**	0.778	0.783	0.460	0.389
joy	0.623	0.792	**0.873**	**0.944**	0.802	0.860	0.522	0.421
relief	**1.000**	0.999	**1.000**	0.600	**0.990**	0.985	0.599	0.621
shame	0.921	0.949	**0.976**	0.706	**0.922**	0.910	0.538	0.566
sorrow	0.720	0.861	**0.920**	**0.814**	0.809	0.791	0.553	0.520
surprise	0.805	0.904	**0.940**	0.862	0.866	**0.874**	0.520	0.523
All approx.	0.675	0.751	**0.802**	**0.852**	0.804	0.818	0.517	0.469

Table 1). In primary evaluation of the system [27], approximately 82 percent of all extracted emoticons were extracted correctly. The problem appeared in erroneously extracting additional areas as separate emoticons. We solved this problem by detecting the erroneously extracted additional areas in a postprocedure, using the additional area database and reattaching the erroneously extracted areas with the actual emoticons they belonged to. This optimized the extraction procedure. There were still 73 cases (from 14,570) of erroneously extracting additional areas as emoticons. The analysis of errors showed that these erroneously extracted additional areas contained elements appearing in databases of semantic areas of eyes or mouths and emoticon borders. To solve this problem, the error cases would have to be added as exceptions; however, this would prevent the extraction of such emoticons in the future if they actually appeared as emoticons. Therefore, we agreed to this minimal error rate (0.5 percent), with which the extraction accuracy of CAO is still near ideal (99.5 percent). Finally, the results for the emoticon extraction and division into semantic areas, when represented by the notions of Precision and Recall, were as follows: CAO was able to extract and divide all of the emoticons; therefore, the Recall rate was 100 percent. As for the Precision, 14,497 out of 14,570 were extracted and divided correctly, which gives the rate of 99.5 percent. The balanced F-score for these results equals 99.75 percent, which clearly outperforms the system of Tanaka et al. [4].

7.1.2 Affect Analysis of Emoticons

First, we calculated how many of the extracted emoticons the system was able to annotate any emotions for. This was done with a near-ideal accuracy of 99.5 percent. 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 (100 percent).

Second, we calculated the accuracy in annotation of the particular emotion types on the extracted emoticons. From the five ways of result calculation, two (Position and Unique Position) achieved much lower results than the other three, about 50 percent, and were discarded from further evaluation. All of the other three (Occurrence, Frequency, and Unique Frequency) scored high, from over 80 percent to over 85 percent. The highest overall score in the training set evaluation was achieved by, in order: Occurrence (85.2 percent), Unique Frequency (81.8 percent), and Frequency (80.4 percent). Comparison with the other emoticon analysis system showed that even after the improvements that we made, the best score it achieved (80.2 percent) still did not exceed our worst score (80.4 percent). For details see Table 5.

7.2 Test Set Evaluation

7.2.1 Emoticon Detection in Input

The system correctly detected the presence or absence of emoticons in 976 out of 1,000 sentences (97.6 percent). In 24 cases (2.4 percent of all sentences), the system failed to detect that an emoticon appeared in the sentence. However, the system achieved an ideal score in detecting the absence of emoticons. This means that there are no errors in the detecting procedure itself, but that the database does not cover all possibilities of human creativity. However, it can be reasonably assumed that if our system, with the database coverage of over 3 million possibilities, still has 2.4 percent of error in emoticon detection, the methods based on smaller databases would fail even more often in similar tasks. The strength of the Cohen's coefficient of agreement with human annotators was considered to be very good ($\mathrm{kappa} = 0.95$). The results are summarized in Table 7.

7.2.2 Emoticon Extraction from Input

From 418 sentences containing emoticons, CAO extracted 394 ($\mathrm{Recall} = 94.3\%$). All of them were correctly extracted and divided into semantic areas ($\mathrm{Precision} = 100\%$), which gave

TABLE 6
Results of the CAO System in Affect Analysis of Emoticons

Emotion Estimation on Separate Emoticons									
Yamada et al. (2007)			CAO						
1-gram	2-gram	3-gram	Occurrence		Frequency		Unique Frequency		
			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						
1-gram	2-gram	3-gram	Occurrence		Frequency		Unique Frequency		
			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**	

The results summarize three ways of score calculation, specific emotion types, and 2D affect space. The CAO system shown in comparison to another system.

an overall extraction score of over 97.1 percent 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, a capability not present in the compared system.

As an interesting remark, it should be noticed that in the evaluation on the training set, the Recall scored perfectly, but the Precision did not, and in the evaluation on the test set it was the opposite. This suggests that sophisticated emoticons, which CAO had problems detecting, do not appear very often in the corpora of natural language such as blog contents, and the database applied in CAO is sufficient for the tasks of emoticon extraction from input and emoticon division into semantic areas. However, as human creativity is never perfectly predictable, sporadically (in at least 2.4 percent of cases), new emoticons still appear which the system is not able to extract correctly. This problem could be solved by frequent updates of the database. The race against human creativity is always an uphill task, although, with close to ideal extraction (over 97 percent), CAO is already a large step forward. The results are summarized in Table 7.

7.2.3 Affect Analysis of Separate Emoticons

The highest score was achieved by, in order: Unique Frequency (93.5 percent for specific emotion types and 97.4 percent for estimating groups of emotions mapped on Russell's affect space model), Frequency (93.4 percent and 97.1 percent), and Occurrence (89.1 percent and 96.7 percent). The compared system by Yamada et al. [5], despite the numerous improvements we made to this system, did not score well, achieving its best score (for trigrams) far below our worst score (Occurrence/Types). The scores are shown in the top part of Table 6. The best score was achieved by Unique Frequency, which, in training set evaluation, achieved the second highest score. This method of score calculation will therefore be used as default score calculation in the system. However, to confirm this, we also checked the results of evaluation of affect analysis of sentences with CAO.

TABLE 7
Results of the CAO System in Emoticon Detection, Extraction from Input, and Estimation of Emotions

		Detection		Extraction		
		System		R	P	F-score
		Emoticon	No emoticon	94.3%	100%	97.1%
Users	Emoticon	394	24			
	No emoticon	0	582	$\left(\frac{394}{418}\right)$	$\left(\frac{394}{394}\right)$	$\frac{2*P*R}{(P+R)}$
No. of agreements=976 (97.6%), Kappa= 0.95						

7.2.4 Affect Analysis of Emoticons in Sentences

The highest score was achieved by, in order: Unique Frequency (80.2 percent for specific emotion types and 94.6 percent for estimating groups of emotions mapped on Russell's affect space model), Frequency (80 percent and 94 percent), and Occurrence (75.5 percent and 90.8 percent). It is the same score order, although the evaluation was not of estimating emotions of separate emoticons, but of whole sentences with the use of CAO. This proves that Unique Frequency is the most efficient method of output calculation for our system. The compared system scored poorly here as well, achieving only one score (for bigrams) higher than our worst score (Occurrence/Types). The scores are shown in the bottom part of Table 6.

The score for specific emotion-type determination was, as we expected, not ideal (from 75.5 percent to 80.2 percent). This confirms that, using only emoticons, affect analysis of sentences can be performed at a reasonable level (80.2 percent). However, as the emotive information conveyed in sentences also consists of other lexical and contextual information, it is difficult to achieve a result close to ideal. Although the results for 2D affect space were close to ideal (up to nearly 95 percent), which means that the emotion types for which human annotators and the system did not agree still had the same general features (valence polarity and activation), this also confirms the statement from Section 5.5 that people sometimes misinterpret (or use interchangeably) the specific emotion types of which general features remain the same (in the test data people annotated, e.g., "fondness" on sentences with emoticons expressing "joy," or "surprise" on "excitement," etc., but never, e.g., "joy" on "fear"). The above can also be interpreted as further proof for the statement from Section 3.2, where emoticons are defined as expressions used in online communication as representations of body language. In direct communication, body language is also often used to convey a supportive meaning for the contents conveyed through language. Moreover, some sets of behavior (or kinemes) can be used to express different specific meanings for which the general

TABLE 8
Examples of Analysis Performed by CAO

Example 1: *Chakku-shime wasure-san ga ooidesu ne, watashi mo tama ni yarakashite hitori sekimen* (;^_^A
Translation: Many people forget to close their fly. I sometimes do that too and when I notice, I get all red (;^_^A

S_1	B_1	S_2	$E_L M E_R$	S_3	B_2	S_4
N/A	(	;	^_^	A	N/A	N/A

CAO	Human Annotation: emoticon	Human Annotation: sentence
fear / anxiety (0.06450746)	fear / anxiety	fear / anxiety, shame
...		

Example 2: *Itsumo, "Mac, ne-----" tte shibui kao sareru n desu. Windows to kurabete meccha katami ga semai desu* (ノ Д`):･ﾟ +:･。
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`):･ﾟ +:･。

S_1	B_1	S_2	$E_L M E_R$	S_3	B_2	S_4
N/A	(	N/A	ノ Д`	N/A	)	:･ﾟ +:･。

CAO	Human Annotation: emoticon	Human Annotation: sentence
sadness / sorrow (0.00698324)	sadness / sorrow	sadness / sorrow, dislike
excitement (0.004484305)		
dislike (0.001897533)		
...		

Example 3: >*Aki-san, eee,* (ﾉﾟ oﾟ)ﾉ *ipod wa nai to iya dakara sugu ni juden da yo!*!
Translation: >>Aki-san, What!? (ﾉﾟ oﾟ)ﾉ I couldn't imagine a day without my ipod! Recharge your battery at once!

S_1	B_1	S_2	$E_L M E_R$	S_3	B_2	S_4
N/A	(	ﾉ	ﾟ oﾟ	N/A	)	ﾉ

CAO	Human Annotation: emoticon	Human Annotation: sentence
surprise (0.02686763)	surprise	surprise
joy (0.02679939)		
excitement (0.02238806)		
...		

Example 4: *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	Human Annotation: sentence
surprise (0.4215457)	surprise	surprise , dislike
...		

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

emotive feature remains the same. For example, wide opened eyes and mouth might suggest emotions like fear, surprise, or excitement; although the specificity of the emotion is determined by the context of a situation, the main feature (activation) remains the same. In our evaluation, the differences in the results for specific emotions types and 2D affect model prove this phenomenon. Some examples illustrating this have been presented in Table 8.

8 CONCLUSIONS

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 10,000 of unique emoticons collected from the Internet. These emoticons were automatically distributed into emotion-type databases with the use of an affect analysis system developed by Ptaszynski [20]. Finally, the emoticons were automatically 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 [19] idea of kinemes as minimal meaningful elements in body language. The database applied in CAO contains over 10,000 raw emoticons and several thousands of elements for each unique semantic area (mouths, eyes, etc.). This gave the system coverage of over 3 million combinations. With such a coverage, the system is capable of automatically annotating potential emotion types of any emoticon. There are a finite number of semantic areas used by users in emoticons generated during online communication. The number CAO can match over 3 million emoticon face (eye-mouth-eye) triplets and is sufficient to cover most possibilities.

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, which, calculated using balanced F-score, reached over 97 percent. Among the five methods of calculating emotion rank score we compared in evaluation of emotion estimation of emoticons, the highest and the most balanced score was based on Unique Frequency and this method of score calculation will be used as a default setting in CAO. Using Unique Frequency, the system estimated emotions of separate emoticons with an accuracy of 93.5 percent for the specific emotion types and 97.3 percent for groups of emotions belonging to the same 2D affect space [22]. There were some minor errors, however not exceeding the standard error level, which can be solved by optimization of CAO's procedures during future usage. Also, in affect analysis of whole sentences, CAO annotated the expressed emotions with a high accuracy of over 80 percent for specific emotion types and nearly 95 percent for 2D affect space.

9 Future Work

At present, CAO is the most accurate and reliable system for emoticon analysis known to the authors. In the near future, we plan to apply it to numerous tasks. Beginning with a 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 with other people. 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 also be 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 [20], CAO could also contribute to preserving online security [28], which has been an urgent problem for several years. To standardize emoticon interpretation, we plan to contribute to the Smiley Ontology Project [29]. Finally, we plan to annotate large corpora of online communication, like Yacis Corpus, to contribute to linguistic research on emotions in language.

References

[1] N. Suzuki and K. Tsuda, "Express Emoticons Choice Method for Smooth Communication of e-Business," *Proc. 10th Int'l Conf. Knowledge-Based and Intelligent Information and Eng. Systems,* pp. 296-302, 2006.

[2] D. Derks, A.E.R. Bos, and J. von Grumbkow, "Emoticons and Social Interaction on the Internet: The Importance of Social Context," *Computers in Human Behavior,* vol. 23, pp. 842-849, 2007.

[3] K.C. Chiu, "Explorations in the Effect of Emoticon on Negotiation Process from the Aspect of Communication," master's thesis, Dept. of Information Management, Nat'l Sun Yat-sen Univ., 2007.

[4] Y. Tanaka, H. Takamura, and M. Okumura, "Extraction and Classification of Facemarks with Kernel Methods," *Proc. 10th Int'l Conf. Intelligent User Interfaces,* Jan. 2005.

[5] T. Yamada, S. Tsuchiya, S. Kuroiwa, and F. Ren, "Classification of Facemarks Using N-Gram," *Proc. Int'l Conf. Natural Language Processing and Knowledge Eng.,* pp. 322-327, 2007.

[6] M. Kawakami, "The Database of 31 Japanese Emoticon with Their Emotions and Emphases," *The Human Science Research Bull. Osaka Shoin Women's Univ.,* vol. 7, pp. 67-82, 2008.

[7] A. Ip, "The Impact of Emoticons on Affect Interpretation in Instant Messaging," http://amysmile.com/pastprj/emoticon_paper.pdf, 20110.

[8] A. Wolf, "Emotional Expression Online: Gender Differences in Emoticon Use," *CyberPsychology and Behavior,* vol. 3, no. 5, pp. 827-833, 2004.

[9] J.M. Maness, "A Linguistic Analysis of Chat Reference Conversations with 18-24 Year-Old College Students," *The J. Academic Librarianship,* vol. 34, no. 1, pp. 31-38, Jan. 2008.

[10] J. Nakamura, T. Ikeda, N. Inui, and Y. Kotani, "Learning Face Mark for Natural Language Dialogue System," *Proc. Conf. IEEE Int'l Conf. Natural Language Processing and Knowledge Eng.,* pp. 180-185, 2003.

[11] N. Suzuki and K. Tsuda, "Automatic Emoticon Generation Method for Web Community," *IADIS Int'l Conf. Web Based Communities,* pp. 331-334, 2006.

[12] K. Takami, R. Yamashita, K. Tani, Y. Honma, and S. Goto, "Deducing a User's State of Mind from Analysis of the Pictographic Characters and Emoticons Used in Mobile Phone Emails for Personal Content Delivery Services," *Int'l J. Advances in Telecomm.,* vol. 2, no. 1, pp. 37-46, 2009.

[13] J. Read, "Using Emoticons to Reduce Dependency in Machine Learning Techniques for Sentiment Classification," *Proc. ACL Student Research Workshop,* pp. 43-48, 2005.

[14] C. Yang, K. Hsin-Yih Lin, and H.-H. Chen, "Building Emotion Lexicon from Weblog Corpora," *Proc. ACL Demo and Poster Sessions,* pp. 133-136, 2007.

[15] *Handbook of Emotions,* M. Lewis, J.M. Haviland-Jones, L. Feldman Barrett, eds. Guilford Press, 2008.

[16] A. Nakamura, *Kanjo Hyogen Jiten,* [Dictionary of Emotive Expressions] (in Japanese). Tokyodo Publishing, 1993.

[17] M.F. Vargas, *Louder than Words: An Introduction to Nonverbal Communication.* Iowa State Press, 1986.

[18] R.L. Birdwhistell, *Introduction to Kinesics: An Annotation System for Analysis of Body Motion and Gesture.* Univ. of Kentucky Press, 1952.

[19] R.L. Birdwhistell, *Kinesics and Context.* Univ. of Pennsylvania Press, 1970.

[20] M. Ptaszynski, P. Dybala, R. Rzepka, and K. Araki, "Affecting Corpora: Experiments with Automatic Affect Annotation System—A Case Study of the 2 Channel Forum," *Proc. Pacific Assoc. for Computation Linguistics,* pp. 223-228, 2009.

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

[22] J.A. Russell, "A Circumplex Model of Affect," *J. Personality and Social Psychology,* vol. 39, no. 6, pp. 1161-1178, 1980.

[23] M. Ptaszynski, P. Dybala, W. Shi, R. Rzepka, and K. Araki, "Towards Context Aware Emotional Intelligence in Machines: Computing Contextual Appropriateness of Affective States," *Proc. 21st Int'l Joint Conf. Artifical Intelligence,* pp. 1469-1474, 2009.

[24] H. Kubota, K. Yamashita, T. Fukuhara, and T. Nihsida, "POC Caster: Broadcasting Agent Using Conversational Representation for Internet Community," *Trans. JSAI,* [in Japanese], AI-17, pp. 313-321, 2002.

[25] Y. Matsumoto, A. Kitauchi, T. Yamashita, Y. Hirano, H. Matsuda, K. Takaoka, and M. Asahara, "Japanese Morphological Analysis System ChaSen Version 2.2.1," 2000.

[26] T. Kudo and Y. Matsumoto, "Chunking with Support Vector Machines," *Proc. Second Meeting North Am. Chapter Assoc. for Computational Linguistics,* pp. 192-199, 2001.

[27] M. Ptaszynski, P. Dybala, R. Rzepka, and K. Araki, "Towards Fully Automatic Emoticon Analysis System (ˆoˆ) ," *Proc. 16th Ann. Meeting Assoc. for Natural Language Processing,* pp. 583-586, 2010.

[28] A. Abbasi and H. Chen, "Affect Intensity Analysis of Dark Web Forums," *Intelligence and Security Informatics,* pp. 282-288, 2007.

[29] F. Radulovic and N. Milikic, "Smiley Ontology," *Proc. First Int'l Workshop Social Networks Interoperability,* 2009.

Michal Ptaszynski received the MA degree from the University of Adam Mickiewicz, Poznan, Poland, in 2006. He was a research student at the Otaru University of Commerce. Since 2007, he has been working toward the PhD Degree at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include natural language processing, dialogue processing, affect analysis, sentiment analysis, HCI, and information retrieval. He is a member of the AAAI, the HUMAINE, the SOFT, the JSAI, and the NLP and a student member of the IEEE.

Jacek Maciejewski studied for the MA degree in computer science at the University of Adam Mickiewicz, Poznan, Poland. He was awarded a scholarship to Hokkaido University, Japan, for the period 2008-2010. During his scholarship, he participated in research activities at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include software engineering, natural language processing, Web mining, and information retrieval.

Pawel Dybala received the MA degree from Jagiellonian University, Krakow, Poland, in 2006. He was a research student at Hokkaido University, Japan. Since 2007, he has been working toward the PhD degree at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include natural language processing, dialogue processing, humor processing, HCI, and information retrieval.

Rafal Rzepka received the MA degree from the University of Adam Mickiewicz, Poznan, Poland, in 1999, and the PhD degree from Hokkaido University, Japan, in 2004. Currently, he is an assistant professor in the Graduate School of Information Science and Technology at Hokkaido University. His research interests include natural language processing, Web mining, common sense retrieval, dialogue processing, language acquisition, affect analysis, and sentiment analysis. He is a member of the AAAI, the ACL, the JSAI, the IPSJ, the IEICE, the JCSS, and the NLP.

Kenji Araki received the BE, ME, and PhD degrees in electronics engineering from Hokkaido University, Sapporo, Japan, in 1982, 1985, and 1988, respectively. In April 1988, he joined Hokkai Gakuen University, Sapporo, Japan, where he was a professor. He joined Hokkaido University in 1998 as an associate professor in the Division of Electronics and Information Engineering and became a professor in 2002. Presently, he is a professor in the Division of Media and Network Technologies at Hokkaido University. His research interests include natural language processing, spoken dialogue processing, machine translation, and language acquisition. He is a member of the AAAI, the IEEE, the JSAI, the IPSJ, the IEICE, and the JCSS.