Ideas for Using Large-Scale Corpora to Improve Verification of Emotion Appropriateness in Japanese

Michal Ptaszynski† Pawel Dybala† Wenhan Shi† Rafal Rzepka† Kenji Araki†

†Graduate School of Information Science and Technology, Hokkaido University
Address: Kita-ku, Kita 14 Nishi 9, 060-0814 Sapporo, Japan
E-mail: †{ptaszynski,paweldybala,kabura,araki}@media.eng.hokudai.ac.jp

Abstract We present a set of ideas for improving the verification of emotion appropriateness in Japanese language. Emotion appropriateness verification is a new method for discovering not only what are the emotions conveyed by a user in an utterance also whether they are appropriate for the context they are used in. We present ideas to improve this method with the use of several corpora. The corpora we plan to use are Amazon reviews, Web as corpus and two corpora of natural conversations to improve the method and provide conversational strategies for implementation of the method into a conversational agent.

1. Introduction

Affective Computing [1] is a field of study dealing with different aspects of processing of human emotions. The main stream of research in this field is focused on recognizing user emotions in human-computer interaction. However, the recognition is often limited to pairs like joy and sadness [2], or simply to positive or negative emotions.

Ptaszynski [3] presented a prototype method capable of specifying user emotions in a more sophisticated way than simple valence classification. The method is based on a ten type-classification of emotions in the Japanese language proposed by Nakamura [4]. Moreover, it does not only specify what type of emotion was expressed, but also determines whether the emotion is appropriate for the context it appears in. The method uses Ptaszynski's affect analysis and annotation system, ML-Ask [5], for recognition of emotions and Shi's Web mining technique [6] for gathering emotive common sense from the Internet.

However, the method in its primary form, as presented by Ptaszynski [3], was not perfect. The affect analysis system ML-Ask was lacking in database of emotemes (emotiveness indicators), which was gathered manually, as well as emotive expressions, which was based on an out-of-date dictionary [4]. In the Web mining technique, the processing time was several minutes for one utterance therefore it could not be applied as a real time system. This all was the reason the accuracy, even after several improvements [7, 8], was not exceeding 70%. Finally, the method was able to determine whether the emotion expressed by a user was appropriate for the context, but it was not able to determine what conversational strategy should be undertaken in a certain situation.

To solve the mentioned above problems of the intermediary systems, improve the overall accuracy of the emotion appropriateness verification method and to supplement it with a procedure determining the conversational strategy for each case, we decided to gather and present our plans for further improvements. The plans contain using several corpora, such as Amazon reviews, Web as corpus and two corpora of natural conversations, to improve the method and provide conversational strategies for its implementation into a conversational agent.

2. Improving ML-Ask

As it was mentioned above, there were two main problems with the affect analysis system ML-Ask. The first one was lacks in the database of emotemes (syntactical indicators of emotiveness). Since the database was gathered manually, it is possible that some of its elements were classified incorrectly and there might still be many elements not recognized as emotemes. The second problem was with emotive expressions database which, based on an out-of-date dictionary [4], did not contain many of the modern expressions, especially those used in small talks.

2.1. Refining Emotemes

For refining emotemes we plan to perform corpus statistics of online reviews available on online shops, like Amazon [1], in a similar way to Potts and F. Schwarz [9], who extracted n-gram tokens from Amazon online product reviews. The advantage of such corpora is that its already annotated. A five-star rating system of Amazon provides annotation of a product from very bad (1), through neutral (3), to very good (5). The rating is accompanied with a comment review being a linguistic representation of the user's sentiment

Such corpus will be processed by MeCab to extract parts of speech and divided into n-gram phrases. We plan to separate from unigrams to the largest significant n-gram groups. According to Potts, the tokens, with the highest emotive load appear in the largest numbers in 5-star and 1-star reviews, and their concordances form through the corpus a "J-shape", a "U-shape" or a "reversed J-shape", respectively for tokens with a tendency to be positive, balanced and negative.

The operation described above will provide us a list of words and phrases – candidates for emotemes. We will compare this list with the manually gathered database of emotemes and the elements appearing in both collections will be confirmed as certain emotemes and be left in the database. As for the emotemes not appearing on the list of candidates as well as for the new candidates not appearing in the primary database, we will perform a second evaluation. Firstly we plan to extract 30 sentences for one element (10 sentences from each of three different corpora, KWIC by Yoshihira et al. [10] using Web as corpus, BTS Corpus of natural conversations [11], and Nagoya University Conversational Corpus [2]). Having the sentences extracted we will ask laypeople in a questionnaire a question like "Do you think that the author of this utterance made it conveying emotions with it, or was he neutral?". The elements that have passed this evaluation will be included in the database and the rest will be discarded. The operations above will provide us a well refined database of emotemes.

2.2. Disambiguation of Emotemes

After refining the emotemes, we need to disambiguate what emotion type each of them express. The ten-type-classification of emotions [4] is mapped on a 2-dimensional model of affect [12] in a way as proposed by Ptaszynski [3]. The previous experiment, apart from emoteme refining provides us the information on whether an emoteme is used with positive or negative bias ("J-shape" or "reversed J-shape" respectively). This will be the first type of information used in disambiguation. The second type of information will be gathered as follows. We will take 300 sentences (100 for every corpus) for each element from the refined and expanded databases of emotemes and emotive expressions. If there is less then 100 sentences for an element, all existing ones will be used.

Having the sentences extracted, we will check: 1) what emotion types appear most often for the sentences with every emoteme; and 2) what emotemes appear most often in the sentences with each emotion type. This will give us the statistics on what are the most probable emotions expressed when using the particular emoteme. The results of the two statistics will be added, since they complement one another. During analysis of an utterance, the system will compare all of the probable emotion type lists for all emotemes appearing in the utterance and choose the emotion types appearing most often as the most probable candidates for the emotions expressed in the utterance.

---

[1] http://www.amazon.co.jp/

[2] http://tell.fll.purdue.edu/chakoshi/public.html

2.3. Refining Emotive Expression Database

To refine the emotive expression database, we will check the concordance of every emotive expression from the database in all four corpora (Amazon, KWIC, BTS and Nagoya Univ.). The expressions not appearing at all in any of them, will be also double checked on the whole Web and the ones which do not appear at all will be discarded from the database. This operation will refine the out-of-date expressions in the present database. However, we will not delete them completely, but put them in an “out-of-date folder”, since they might be useful in analyzing older texts.

2.4. Expanding Emotive Expression Database

To expand the database of emotive expressions [4], which is already out-of-date and does not include many present expressions, we will use the emotemes which are used with only one, specified type of emotion and the ones used consequently with emotions from one part of Russell’s space [12]. Next, we will again use 300 sentences from the 3 corpora for each emoteme, as in section **2.2**. From those sentences we will extract the most frequent syntactical patterns. The unique expressions appearing in the patterns, including the parts of speech and phrases, as listed by Ptaszynski [5], will be the candidates for the new emotive expressions. The sentences including those candidates will be evaluated in a questionnaire in which we will ask laypeople a sentence, like “Do you agree that this sentence expresses [emotion type]?” This evaluation will provide us the refined list of new emotive expressions, which will enrich the database.

3. Improving Web Mining Technique

One of the disadvantages of Shi’s Web mining technique [6] was that it was taking over several minutes to process one sentence. We will deal with this problem by making a standalone version of the system.

With use of HyperEstraier [3], we will gather a large-scale corpus build from the contents of homepages containing the elements from the database of emotive expressions. This should be able to provide us a fast, stand-alone Web mining system with the efficiency comparable to the present one. The use of HyperEstraier might even improve the efficiency since the present system uses as the corpus the Internet, which contains lots of nose.

We also plan to create a separate database of events associated with certain emotions, with expandable context. For example, a context of the phrase “Having a dinner” could be restrained to “Having a dinner with friends” or “-alone” and the list of emotions associating with each restrained context would be different (e.g. joy and loneliness respectively).

4. Extracting Conversational Strategies

Solomon [13] argues that emotions are, first of all, engagements with the world and strategies people learn to perform when dealing with the world around them. When it comes to language and, especially to conversation, the expressions of emotions can be realized by different conversational strategies. Since one of our goals is to implement the system to the conversational agent it is then necessary to verify, what conversational strategies are used to express which types of emotions and which of them are desirable during the interaction in cases when the expressions of emotions are either appropriate or inappropriate. We will use the set of conversational strategies for the Japanese language distinguished by Nakai et al. [14]. The two conversational corpora (BTS and Nagoya Univ.) will be analyzed by ML-Ask and the appropriateness of the specified emotions will be verified by the Web mining technique. The verification will be evaluated by laypeople in a similar way as in Ptaszynski [3]. We will check what conversational strategies interlocutors exchange after what types of emotions including the information about the appropriateness of those emotions. This data will be used to generate utterance patterns in the conversational agent. As a model we plan to use the system by Higuchi et al. [15], including its improvements, e.g. the one by Dybala [16], who successfully added to the baseline system a pun generator.

[3] http://hyperestraier.sourceforge.net/index.html

5. Conclusions

In this paper we presented a set of ideas for improving a new method for discovering what are the emotions conveyed by a user in an utterance and whether they are appropriate for the context they are used in. The planned improvements include using of different corpora, such as Amazon reviews, Web as corpus and two corpora of natural conversations. By the use of the conversational corpora we also plan to extract conversational strategies for further implementation of the system into a conversational agent.

After implementing the improvements described in this paper we also plan to use our system to compare different conversations, to check the differences in use of the emotive language between people in the same and different age, sex, or on the same and different social levels (e.g. between a teacher and a student and between two students).

References

[1] R. W. Picard, Affective Computing, The MIT Press, Cambridge, 1997.

[2] J. Teixeira, V. Vinhas, E. Oliveira and L. Reis, “A New Approach to Emotion Assessment Based on Biometric Data.” In Proc. of WI-IAT’08, pp. 459-500, 2008.

[3] M. Ptaszynski, P. Dybala, W. Shi, R. Rzepka and K. Araki, “Towards Context Aware Emotional Intelligence in Machines: Computing Contextual Appropriateness of Affective States.” In Proceedings of IJCAI-09, Pasadena, CA, USA, pp. 1469-1474, 2009.

[4] A. Nakamura, Kanjo hyogen jiten, Tokyodo, 1993.

[5] M. Ptaszynski, P. Dybala, R. Rzepka and K. Araki, “Affecting Corpora: Experiments with Automatic Affect Annotation System -A Case Study of the 2channel Forum-,” In Proc. of PACLING-09, Japan, 2009.

[6] W. Shi, R. Rzepka and K. Araki, “Emotive Information Discovery from User Textual Input Using Causal Associations from the Internet,” FIT-08, pp. 267-268, 2008.

[7] M. Ptaszynski, P. Dybala, W. Shi, R. Rzepka and K. Araki, “Shifting Valence Helps Verify Contextual Appropriateness of Emotions,” In Workshop Notes of the IJCAI-09 Workshop on Automated Reasoning about Context and Ontology Evolution (ARCOE-09), Pasadena, California, USA, pp. 19-21, 2009.

[8] M. Ptaszynski, P. Dybala, W. Shi, R. Rzepka and K. Araki, “Conscience of Blogs: Verifying Contextual Appropriateness of Emotions Basing on Blog Contents,” In Proceedings of CI-09, Honolulu, USA, 2009.

[9] C. Potts and F. Schwarz. “Exclamatives and heightened emotion: Extracting pragmatic generalizations from large corpora.” Ms., UMass Amherst, 2008.

[10] K. Yoshihira, T. Takeda, and S. Sekine, “KWIC system for Web Documents” (in Japanese), In Proceedings of NLP-2004, pp. 137-139, 2004.

[11] M. Usami (Ed.), “BTS ni yoru nihongo hanshikotoba kopasu 1 (hatsu tamen, yuujin; zatsudan, touron, sasoi) [Conversation corpus of spoken Japanese using the Basic Transcription System (first meeting, friend's conversation, small talk, discussion, invitation)],” Tokyo University of Foreign Studies, Tokyo, Japan, 2007.

[12] J. A. Russell, A circumplex model of affect, Journal of Personality and Social Psychology, Vol. 39, No. 6, pp. 1161-1178, 1980.

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

[14] Y. Nakai, M. Ooba, M. Doi, “Danwa reberu de no kaiwa kyoiku ni okeru shido komoku no teian - danwa / kaiwa bunseki teki apurochi no kanten kara mita danwa gino no komoku,” Sekai no nihongo kyouiku, 14, 2004.

[15] S. Higuchi, R. Rzepka, & K. Araki, "A Casual Conversation System Using Modality and Word Associations Retrieved from the Web," Proc. of EMNLP-08, Honolulu, USA, pp. 382-390, 2008.

[16] P. Dybala, M. Ptaszynski, R. Rzepka and K. Araki, "Humoroids - Talking Agents That Induce Positive Emotions with Humor," In Proc. of AAMAS’09, 2009.