A Robust Ontology of Emotion Objects

Michal Ptaszynski †　　Rafal Rzepka ‡　　Kenji Araki ‡　　Yoshio Momouchi §

† High-Tech Research Center, Hokkai-Gakuen University
ptaszynski@hgu.jp
‡ Graduate School of Information Science and Technology, Hokkaido University
{kabura,araki}@media.eng.hokudai.ac.jp
§ Department of Electronics and Information Engineering,
Faculty of Engineering, Hokkai-Gakuen University
momouchi@eli.hokkai-s-u.ac.jp

Abstract

In this paper we present a robust ontology of emotion objects. The ontology was automatically extracted from a large scale blog corpus. We describe the method of extraction and the structure of the ontology. Emotion objects in the ontology are generalized in several ways, according to, e.g, emotion class, valence and activation, number of characters, parts of speech or semantic categories they contain.

1 Introduction

One of the philosophical devises to study emotions is representing them in the form of their formal objects. The formal objects of emotions have been defined as "axiological properties which individuate emotions, make them intelligible and give them correctness conditions" [1, 2, 3]. It has been argued that emotions are sophisticated phenomena and are necessarily intentional, or "about something" (an object) [4, 5, 6]. This leads to the idea that emotions can be described by their formal objects. The linguistic-pragmatic reality proves this: we are afraid/proud **of** something, happy **about** something, etc.. However, the problem on how to specifically define such emotion objects remained. The general definition, saying that "if I am afraid of something then there is necessarily a formal object of my fear", was too unspecific, whereas in practice the idea of formal objects has been always explained on particular examples (e.g., If I am afraid of spiders then a spider (or its formal representation) is the object of my fear; or, if John loves Mary is true, then Mary is the object of John's love). In our assumption formal objects of emotions can be studied on a formalized ontology of those specific, or *particular objects* [3]. Unfortunately, there have been no such collection of particular objects, nor there have been proposed general formalizations of them. In this paper we present the first attempt to automatically create a robust computational ontology of emotion objects as a preface to the further research on formal objects of emotions.

The paper outline is as follows. In section 2 we describe previous attempts to create emotion ontologies or collect emotion objects. We explain the method of extraction of emotion objects from a large corpus in section 3 and construction of the ontology in section 4. Finally, we conclude the paper and list up some of the necessary further work.

2 Previous Research

There have been several attempts to create emotion-related ontologies. Strapparava and Valitutti [7] created WordNet-Affect, an ontology of emotion-related concepts for English. Obrenovic et al [8] created an ontology for description of emotional cues from different modalities (text, speech, etc.). Yan et al. [9] semi-automatically created an emotion ontology for Chinese using HowNet, a Chinese lexical dictionary with inter-concept relations described in a form of a network. Radulovic and Milikic [10] created an ontology of emoticons (smileys) with annotations of emotion classes. Baldoni et al. [11] created an emotion ontology for Italian and proposed an application to emotion-related reasoning.

As for the attempts to collect emotion objects, there have been only a few. Quan and Ren [12] created a Chinese emotion blog corpus, **Ren-CECps1.0**. They collected 500 blog articles from various Chinese blog services, such as sciencenet blog (http://blog.sciencenet.cn/), qq blog (http://blog.qq.com/), etc. The articles were later annotated with a large variety of information, such as emotion class, emotive expressions, and emotion objects (or targets). In our research we performed a similar attempt to collect emotion objects from a large blog corpus in Japanese. The emotion objects were further generalized in several ways and put into an ontology. Apart from the work of Quan and Ren, Wiebe et al. [13], and Hashimoto et al. [14] included emotion object (or cause) annotations in their research on creating corpora of respectively, news articles in English and blog posts in Japanese. However, in both those research the focus was not on particular emotion class, but on one general feature of emotions, namely their valence.

3 Extraction of Emotion Objects

Corpus for Emotion Object Extraction: It has been argued that emotions and their formal objects are necessarily in a causal relation [15]. Such relations could be formulated manually one-by-one, however, we assumed that in the age of large datasets, they should be extractable automatically. It has also been proved that syntactic patterns representing causal relations are applicable in emotion term extraction [16, 17, 18]. Moreover, recently a text source that has come into the focus of opinion mining, or sentiment and affect analysis are blogs [12, 14, 19]. Blogs are open access Internet diaries in which people expressively describe their emotions or opinions. The emotions are often described with regards to situations in which they were experienced. Therefore we decided to extract emotion objects from blogs. A sufficiently large corpus of blogs applicable as the source of emotion object extraction, called YACIS, was developed by Maciejewski et al. [20]. They automatically collected nearly 13 million of blog posts from the pages of Ameba blog service (http://ameblo.jp/). The corpus contains 5.6 billion words within 350 million sentences. It has been annotated with different types of annotations. Firstly, Ptaszynski et al. [21] annotated the corpus with affective information, such as emotive expressions, emotion classes, valence, etc.. Secondly, Ptaszynski et al. [22] added syntactic annotations, such as POS tagging, lemma, dependency parsing, etc.. By utilizing both types of information it is possible to extract from the corpus only those sentences that contain affective information and syntactic morphemes representing causal relations.

Extraction Procedure: We focused only on sentences in which emotion (emotive expression) was causally related with the sentence contents. In Japanese the causal relation is represented by a set of causality morphemes, such as *-kara, -tara*, or *-node*. A typical example of a sentence containing both information is presented below (emotive expression underlined, causality morpheme in bold type font):

彼女に振られたから悲しい...
Kanojo ni furareta ***kara*** *kanashii*...
Girlfriend DAT[1] dump PAS CAUS sad ...
I'm sad **because** my girlfriend dumped me...

The example such as above can be analyzed in the following way. Emotive expression (*kanashii*, "sad") is related with the sentence contents (*Kanojo ni furareta*, "my girlfriend dumped me") with a causality morpheme (***kara***, "**because**"). In such situation the sentence contents represents the object of emotion. A generalization of this structure can be represented as

$$O_E \quad CAUS \quad X_E \,,$$

where O_E=[Emotion object], $CAUS$=[**causality**], and X_E=[expression of emotion].

[1]DAT=dative, PAS=passive, CAUS=causality morpheme;

Table 1: Distribution of separate expressions across emotion classes in Nakamura's dictionary (overall 2100 ex.).

emotion class	number of expressions	emotion class	number of expressions
dislike	532	fondness	197
excitement	269	fear	147
sadness	232	surprise	129
joy	224	relief	106
anger	199	shame	65

Shi et al. [17] distinguished five morphemes that co-occur the most often with words describing emotions in Japanese: *-te, -to, -node, -tara* and *-kara*. To verify that they performed a cross reference of appearance of the eleven morphemes with an emotive expression lexicon [23] using Google search engine (http://www.google.co.jp/). In our research we used both their set of causality morphemes ans the emotive expression lexicon. The lexicon used in Shi et al. is the *Emotive Expression Dictionary* [23], a collection of over two thousand expressions describing emotional states collected manually from a wide range of literature. It uses a classification of emotions into ten classes: 喜 *ki/yorokobi* (joy), 怒 *dō/ikari* (anger), 哀 *ai/aware* (sorrow, sadness, gloom), 怖 *fu/kowagari* (fear), 恥 *chi/haji* (shame, shyness), 好 *kō/suki* (fondness), 厭 *en/iya* (dislike), 昂 *kō/takaburi* (excitement), 安 *an/yasuragi* (relief), and 驚 *kyō/odoroki* (surprise). All expressions in the dictionary are annotated with those emotion classes. The distribution of expressions within emotion classes is represented in table 1.

The expressions from the emotive lexicon [23] were modified with the causality morphemes and the queries in this form were used as seeds in YACIS Corpus. The procedure was as follows:

1. Take all X_E from lexicon [23], e.g. *ureshii* (happy), or *kanashii* (sad),
2. Modify X_E with $CAUS$, where $CAUS$=kara|node|te|to|tara, with optional distance between them up to 5 characters,
3. Query $CAUS(.*\{0,5\})X_E$ in YACIS,
4. For each query extract all near O_E.

Evaluation of Extraction Procedure: The overall number of extracted sentences containing emotion objects was 19,459,167. Their distribution among emotion types is shown in table 2. The Spearman's correlation coefficient (ρ [rho]) calculated between the number of expressions in Nakamura's lexicon and the number of emotion objects extracted for each emotion type was low (approximately 0.25), which shows no particular correlation between number of expressions in emotion classes and number of extracted emotion objects. This means that the statistics does not depend on the number of seed phrases, and therefore is reliable.

Table 2: Distribution of number of emotion objects and semantic categories among emotion types.

Emotion type (# of expressions)	Number of extracted: emotion objects	semantic categories
joy (224)	6,123,947	120,068,119
relief (106)	3,321,795	66,605,209
dislike (532)	2,957,596	59,136,703
fondness (197)	2,441,865	45,816,845
fear (147)	1,184,952	23,490,755
excitement (269)	1,104,998	23,102,117
sadness (232)	930,698	18,911,960
surprise (129)	898,138	18,164,806
anger (199)	292,500	5,966,179
shame (65)	202,678	4,006,378
Spearman's ρ	0.24848	0.24848

In the evaluation we randomly extracted a 418 middle sized sentences from the extracted set. The test set contained information on the emotion types expressed in the sentence and the emotion object corresponding to it. The sentences were then evaluated by laypeople, who were asked in a questionnaire about what type of emotion is conveyed and whether the part of the sentence separated as emotion object actually contains the object of emotion. As the result 80% of the cases were evaluated as correct. The result was high, although not ideal. This was mostly due to the fact that some causality morphemes are sometimes used in a function different than causal. The 20% error rate in a nearly 20 million sentence collection scales up to potentially several million failure cases. Therefore in the future it is necessary to optimize the procedure, or additionally filter out the erroneous samples. A few examples of success and failure of the emotion object extraction procedure are represented in table 3.

4 Ontology Construction

The objects of emotion extracted according to the procedure described in section 3 were further composed into an ontology. The ontology was created using Emotion Markup Language (EmotionML)[24] annotation scheme modified for the needs of Japanese language and the emotive expression lexicon used in this research. EmotionML is "a markup language designed to be usable in a broad variety of technological contexts while reflecting concepts from the affective sciences" [24]. It has been used in much research on emotion processing. The objects in the ontology were formalized in five ways:

Emotion class: Emotion classes were annotated according to Nakamura's classification [23].

Emotion Dimensions: Emotion dimensions were annotated according to Russell's two dimensional model of affect [25] with mapping of Nakamura's emotion classes on Russell's model proposed by Ptaszynski et al. [18]. The model assumes that emotions can be represented in two dimensions: valence (positive/negative) and activation (activated/deactivated).

```
<emotionObject name="レディーガガを見る">
   <metadata>
      <origin:sentence name="なぜかレディーガガを見ると恐怖感じる(;´艸`)"/>
      <causality name="と"/>
      <emotiveExpression name="恐怖"/>
      <restOfSentence name="感じる(;´艸`)"/>
   </metadata>
   <emotion>
      <category set="NakamurasEmotions" name="fear"/>
      <dimensions set="valenceActivation">
         <activation value="1"/><!-- high activation -->
         <valence value="-1"/><!-- negative valence -->
      </dimensions>
      <charlength value="9"/>
      <generalSemanticCategory property="見る=精神および行為#2.3#;"/>
      <specificSemanticCategories property="見る=注意・認知・了解#2.3062#;
                                    見る=読み#2.3150#;見る=見る#2.3091#;"/>
      <partsOfSpeech name="N P V"/>
   </emotion>
</emotionObject>
```

Figure 1: An example of emotion object represented in the ontology.

Number of characters: According to Quan and Ren [12] and Matsumoto et al. [26], the number of characters is one of the distinctive features for emotional phrases. For example, shorter phrases tend to be more expressive and emotional, while longer phrases tend to be more descriptive and therefore less emotional.

Parts of Speech: The annotation of parts of speech (POS) on emotion objects was done with MeCab [27], a standard POS tagger for Japanese.

Semantic categories: The semantic formalization of emotion objects was performed using *Bunrui Goihyo* - word list including semantic categories of words [28]. The categories include such labels as: "Abstract objects", "Human Activities", "Natural Objects and Phenomena", "Subject of Actions", etc. Each main category consists also of a number of sub-categories.

One example of emotion object represented in the ontology is represented in figure 1.

5 Conclusions and Future Work

In this paper we presented an ontology of emotion objects. The objects were automatically extracted from a large scale blog corpus. The extraction procedure was evaluated on 80%, which is a promising result. However, the scale of ontology (nearly 20 million of objects) makes it necessary to optimize the procedure or find another way to discard the erroneously extracted objects. The extracted objects were further formalized in several ways and converted into an ontology. The formalizations include emotion classes, emotion dimensions (valence and activation), length in characters, parts of speech and semantic categories. In the near future we plan to thoroughly evaluate the ontology. With the use of the ontology we plan to propose a formalization of particular objects of emotions, and contribute to the research on formal objects of emotions. We also plan to apply the ontology in practical tasks such as affect or sentiment analysis.

Table 3: A few examples of success and failure of the emotion object extraction procedure.

success	failure
友達にアルバム貸してなかなか返ってこなくて怒り気味のむろぴでした-_-笑	怒りと悲しさでｲｯﾊﾟｲです(*_*)
自然災害は予測しにくいので怖いですよね(-.-;)	自分で着てみた時の物悲しさ#kon#(ノ_-。
雨続くとホンマ嫌ですよね(ノД<;。	でも握力が無いので、こそばゆいだけかも～(^-^)/
歯医者って怖いですよね m(__)m	怒られる要因がなくても無理に怒り出すから、もうしばらくはプライベートも静かにしてなきゃ(T-T)
ipodはないと嫌だからすぐに充電だよ！！	
上の娘が来月から小学生なのでちょっと心配です(;^ω^A	

emotion object, **causality form**, emotive expression

Acknowledgments

This research was supported by (JSPS) KAKENHI Grant-in-Aid for JSPS Fellows (Project Number: 22-00358).

References

[1] Wilson, J. R. S. 1972. *Emotion and Object*, Cambridge University Press.

[2] Mulligan, K. 2007. *Intentionality, Knowledge and Formal Objects*, Electronic Festschrift for W. Rabinowicz, T. Ronnow-Rasmussen et al. (eds.). www.fil.lu.se/HommageaWlodek/site/papper/MulliganKevin.pdf

[3] Teroni, F. 2007. "Emotions and Formal Objects", *dialectica*, pp. 395-415.

[4] Kenny, A. 2003. *Action, Emotion and Will*, Routledge, 2003 (reprint from 1963).

[5] Lazarus, R. S. 1991. "Cognition and Motivation in Emotion", *American Psychologist*, Vol. 46, pp. 362-367.

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

[7] Strapparava, C. and Valitutti, A. 2004. "WordNet-Affect: an affective extension of WordNet", In *Proc. of LREC*, Vol. 4, pp. 1083-1086.

[8] Obrenovic, Z., Garay, N., L?ez, J. M., Fajardo, I., Cearreta, I. 2005. "An Ontology for Description of Emotional Cues", In *Proceedings of Proceedings of the First international conference on Affective Computing and Intelligent Interaction (ACII'05)*, pp. 505-512.

[9] Yan, J., Bracewell, D. B., Ren F. and Kuroiwa, S. 2008. "The Creation of a Chinese Emotion Ontology Based on HowNet", *Engineering Letters*, 16:1, EL_16_1_24.

[10] Radulovic, F. and Milikic, N. 2009. "Smiley Ontology", In *Proc. First Int? Workshop Social Networks Interoperability*, Shanghai, China.

[11] Baldoni, M., Baroglio, C., Patti V. and Rena P. 2011. "From Tags to Emotions: Ontology-driven Sentiment Analysis in the Social SemanticWeb", In *Proceedings of the 5th International Workshop on New Challenges in Distributed Information Filtering and Retrieval*, Palermo, Italy.

[12] Quan, C., Ren, F. 2010. "A blog emotion corpus for emotional expression analysis in Chinese", *Computer Speech & Language*, Vol. 24, Issue 4, pp. 726-749.

[13] Wiebe, J., Wilson, T. and Cardie, C. 2005. "Annotating expressions of opinions and emotions in language". *Language Resources and Evaluation*, Vol. 39, Issue 2-3, pp. 165-210.

[14] Hashimoto, C., Kurohashi, S., Kawahara, D., Shinzato, K. and Nagata, M. 2011. "Construction of a Blog Corpus with Syntactic, Anaphoric, and Sentiment Annotations" [in Japanese], *Journal of Natural Language Processing*, Vol 18, No. 2, pp. 175-201.

[15] Aquila, R. E. 1974. "Emotions, Objects and Causal Relations", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 26, No. 3/4, pp. 279-285.

[16] Tokuhisa, R., Inui K., Matsumoto, Y. 2008. "Emotion Classification Using Massive Examples Extracted from the Web", In *Proc. of Coling 2008*, pp.881-888.

[17] Wenhan, S., Ptaszynski, M., Rzepka, R. and Araki, K. 2008. "Emotive Information Discovery from User Textual Input Using Emotion Expression Element and Web Mining", *Technical Report of IEICE*, Thought and language, 108(353), pp.31-34.

[18] Ptaszynski, M., Dybala, P., Shi, W., Rzepka R. and Araki, K. 2009. "A System for Affect Analysis of Utterances in Japanese Supported with Web Mining", *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, Special Issue on Kansei Retrieval, Vol. 21, No. 2, pp. 30-49 (194-213).

[19] Li, J. and Ren, F. 2008. "Emotion Recognition from Blog Articles", *IEEE NLP-KE 2008*, pp. 355-362.

[20] Maciejewski, J., Ptaszynski, M., Dybala, P. 2010. "Developing a Large-Scale Corpus for Natural Language Processing and Emotion Processing Research in Japanese", In *Proceedings of the International Workshop on Modern Science and Technology (IWMST)*, pp. 192-195.

[21] Ptaszynski, M., Dybala, P., Rzepka, R., Araki, K. and Momouchi, Y. 2012. "Annotating Affective Information on 5.5 Billion Word Corpus of Japanese Blogs", In *Proceedings of The 18th Annual Meeting of The Association for Natural Language Processing (NLP-2012)*.

[22] Ptaszynski, M., Rzepka, R., Araki, K. and Momouchi, Y. 2012. "Annotating Syntactic Information on 5.5 Billion Word Corpus of Japanese Blogs", In *Proceedings of The 18th Annual Meeting of The Association for Natural Language Processing (NLP-2012)*.

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

[24] Schroder, M., Baggia, P., Burkhardt, F., Oltramari, A., Pelachaud, C., Peter, C., Zovato, E. 2010. "Emotion markup language (EmotionML) 1.0.", http://www.w3.org/TR/emotionml/

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

[26] Matsumoto, K., Konishi, Y., Sayama, H., Ren, F. 2011. "Analysis of Wakamono Kotoba Emotion Corpus and Its Application in Emotion Estimation", *International Journal of Advanced Intelligence*, Vol. 3, No. 1, pp.1-24.

[27] Kudo, T. 2001. "MeCab: Yet Another Part-of-Speech and Morphological Analyzer", 2001. http://mecab.sourceforge.net/

[28] The National Institute for Japanese Language, "Bunrui Goihyo (Word List by Semantic Principles, Revised and Enlarged Edition)", http://www. kokken.go.jp/en/publications/bunrui_goihyo/