Generating Story Reviews Using Phrases Expressing Emotion

Hiroshi Ota and Kazuhide Yamamoto

Nagaoka University of Technology,
1603-1, Kamitomioka, Nagaoka, Niigata 940-2188 Japan
{ota, ykaz}@nlp.nagaokaut.ac.jp

Abstract. This paper presents a method for generating reviews of stories. In this work, we focus on generating sentences that include subjective expressions. First, we constructed lexicons using emotion-emerged expressions that are thought to be the origin of the emotional content. The lexicon consists of syntactic pieces that are proposed as units of syntactic structure to represent suitable units of expression for estimating emotions. We confirmed the effectiveness of the lexicons by estimating the emotional content of blog text. Second, we confirmed the relationship between the naturalness of the sentences and the coherence of the emotion it expresses. Finally, we proposed a method of review generation using the lexicons.

Keywords: emotion, syntactic piece, emotion-emerged expression(E3), review, language generation

1. Introduction

In many situations, we use subjective expressions that convey emotions, feelings and evaluations about someone or something. It is essential to be able to understand such expressions, and in particular, to be able to use them in language generation tasks. One kind of text that involves subjective expressions is a review of consumer goods and services. In this work, we focus on the task of generating review sentences.

This study considers emotional expressions as a subset of subjective expressions. Reviews of stories may include many emotional expressions that emerge in response to the story being reviewed. In this sense, we call the expressions of these emerging emotions ‘emotion-emerged expressions’ (hereafter, E3) in this paper. We first built lexicons for E3, and then made a prototype system to generate story reviews.

2. Related Studies

Sentence generation technology has been approached from the viewpoint of the semantic network (Ozaki et al., 1997) or lexical relations (Seki et al., 1999). However, subjective expressions have not been considered. Sentences should be coherent, as emotional coherence affects the naturalness of a sentence. Consider these three sentences:

- She receives a present that feels joyous.
- She receives a present that is sad.
- She receives a present that is afraid.

We can select the first sentence ‘She receives a present that feels joyous.’ as the most natural. In this study, we confirm the relationship between the naturalness of a sentence and the coherence of the emotion it expresses.

Related studies on emotion analysis include those of Kozareva et al. (2007), Yang et al. (2007) and Mihalcea and Liu (2006). Kozareva et al. propose an emotional classification approach based on frequency and co-occurrence. They calculate an emotional score for each word using mutual information. A word can express several types of emotions, and we aim to compile these expressions for text generation.

Some methods are available for extracting sentences and paragraphs that have emotional content from the blog corpus. Yang et al. used an emoticon as a marker for emotional expression and extracted emotion expressing sentences. Mihalcea and Liu used blog tags such as 'happy' or 'sad' and extracted emotion expressing blogs. However, the tagged blog content that can be collected is limited. One characteristic of our method is considering words used not only in blog text but also in their titles.

Yang el al. built a lexicon of emotions, and Mihalcea and Liu compiled phrases identified as happiness factor. They formed lexicons from 1gram, 2gram or 3gram. We form lexicons from syntactic pieces (Aoki and Yamamoto, 2007). Our purpose is to generate a natural sentence that includes subjective phrases. Therefore, we use syntactic pieces that have structural information and generate sentences effectively.

3. Emotion-Emerged Expression Lexicons (E3 Lexicons)

This section presents now, we construct E3 lexicons that are collections of phrases associated with a certain emotion. We design the lexicons using blogs as a large corpus. Figure 1 illustrates the construction method. First, we identify blogs that express a certain emotion. Second, we form a model of classifier using these blogs. Third, the model estimates the emotional content of new blogs. Finally, we construct the E3 lexicons from these new blogs.

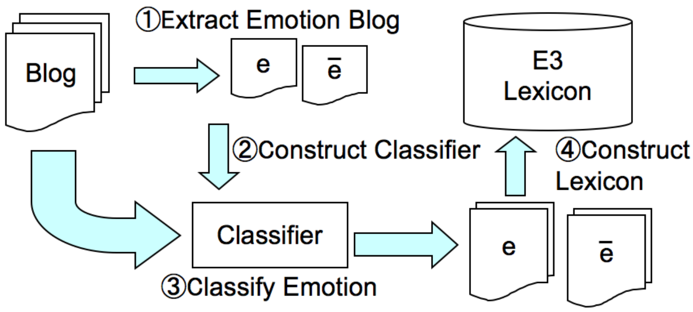

Figure 1: Construction process for an E3 lexicon

3.1. Emotional Word Lexicons

Before constructing the E3 lexicons, we prepared emotional word lexicons. Emotional words clearly represent an emotion; for example, a word 'glad' clearly represents a happy emotion. Therefore, the lexicons provide useful information for extraction of particular emotion expressing blogs(in section 3.2). Emotional words are collected by a human as follows. First, we give adjectives and adverbial words extracted from the IPA thesaurus to 5 examinees. The examinees select the emotion associated with the words from eight possible emotions. Those are defined by Plutchik (1960). We used the words for which there is more than 75% agreement. Table 1 shows the eight possible emotions and the number of words associated with each emotion.

Table 1: The number of words in each emotional word lexicon

Joy	**sadness**	**acceptance**	**disgust**	**surprise**	**anticipation**	**fear**	**anger**
67	100	6	302	28	14	31	6

For brevity, we treat three emotions, joy, sadness and fear, hereafter. We estimate that these three emotions don't very often co-occur. However, we must note that our method does not depend on the number of emotions.

3.2. Extraction of Emotion expressing Blogs

This section proposes a method for extracting blogs that have a particular emotion e. This method considers the words in the blog, from the title and the text. The title is a summary of the article. Our hypothesis is that a blog expresses emotion e when the title has words somehow associated with emotion e. In addition, the method considers words in the blog text. The blog has emotion e when the text contains words associated with emotion e in greater numbers than words in the lexicon of other emotions $\bar{e}$. $\bar{e}$ mean emotions other than e; when emotion e is sadness, emotions $\bar{e}$ are joy and fear. Thus, there are two requirements, below, for a blog to be considered as having emotion e.

Requirement 1: The title includes emotion-e word(s)
Requirement 2: The body includes emotion-e word(s) in greater numbers than words associated with other emotions $\bar{e}$

We performed a preliminary experiment to test the hypothesis. This experiment aimed at extracting blogs expressing sadness. The words in the lexicon sadness were used as emotion-e words, and those in the lexicons joy and fear as emotion-$\bar{e}$ words. We tried the method in 3 patterns: satisfying (1) only requirement 1, (2) only requirement 2 and (3) both requirements.

Table 2: Extraction of blogs expressing sadness

	Requirement 1	**Requirement 2**	**Requirements 1 & 2**
Accuracy (%)	50	60	85

This extraction experiment achieved 85% accuracy when both the requirements were satisfied. As a result, we regarded a blog as being an emotion-e blog when it satisfied both requirements 1 and 2.

3.3. Emotion Classifier Model

This section constructs the emotion classifier model. The classifier classifies the blog according to whether the blog text expresses the emotion-e or -$\bar{e}$. We propose the method for extracting emotion-e blogs in section 3.2. This method obtains good accuracy. However, most blogs fail to fulfil the two requirements. These blogs must be used if we wish to construct a large-scale lexicon(Construct method is written in section 3.4). We want to estimate for all blogs whether the blog emerge the emotion emotion-e or -$\bar{e}$. Therefore, we construct a classifier model by utilizing emotional blogs. This model is binary classifier. We should construct the model for each emotions; joy, sadness and fear. When constructing the model for sadness, emotion-e is sadness, and -$\bar{e}$ is joy and fear. Emotion-e and -$\bar{e}$ blogs, which are collected as described in section 3.2, were used as training data. For our experiments, we used TinySVM, a support-vector machine package. We adopt all the content words of the blogs as the features. Table 3 shows the result of 10-fold cross-validation.

Table 3: Evaluation of the emotion classifier model

	joy	**sadness**	**fear**
Number of Blog text	17270	10810	6460
Accuracy (%)	70.9	71.1	71.1

3.4. Constructing E3 Lexicons

This section proposes an automatic construction method for E3 lexicons. The lexicons are formed using syntactic pieces (Aoki and Yamamoto, 2007), which are defined as pairs consisting of a modifier and a modificand (modified entity) based on dependency analysis. A syntactic piece is the minimum unit of syntactic structure of an expression. This pair is expressed as follows.

syntactic piece : modifier => modificand

A syntactic piece has useful characteristics. In particular, it can deal with a chunk of meaning. Emotion may emerge from a phrase as follows:

・背筋が => こおる
(a chill goes down one's spine)
・腰を => やる
(cause low back pain)

These phrases are associated with fear, which cannot be associated with the individual words. Furthermore, this unit has structure, a characteristic that is effective in generating sentences. We adopt syntactic pieces for the lexicons in this task.

We collected syntactic pieces from emotion-e and -$\bar{e}$ blogs (all blogs were classified as either e or $\bar{e}$ by the classifier model) and assigned an emotion-e score to each one. The scoring method used the formula of Fujimura et al. (2004). According to Fujimura, a word that has positive semantic orientation should appear in text expressing a positive opinion. The same can be said even if we replace the positive opinion with emotion e. Furthermore, the same can be said even if we replace a word with a syntactic piece. Based on this hypothesis, we calculate the frequency differences for syntactic pieces between emotion-e and emotion-$\bar{e}$ blogs. In cases of the emotion-sadness score, syntactic pieces expressing sadness have a higher score because they are expected to occur more often in emotion-sadness blogs compared with emotion-joy and emotion-fear blogs. We calculate the emotion-e score as follows:

1) Extract all syntactic pieces from all blog text.
2) Count the occurrence rate of each of the syntactic pieces in emotion-$e/\bar{e}$ blogs.
3) Calculate the $score_e(piece)$ for each of the syntactic pieces with the following expression.

$$score_e(piece) = \frac{P_e(piece) - P_{\bar{e}}(piece)}{P_e(piece) + P_{\bar{e}}(piece)} \qquad (1)$$

$$(-1 \le score_e(piece) \le 1)$$

where $piece$ is a syntactic piece, $score_e(piece)$ is the emotion score of piece for emotion-e, $P_e(piece)$ is the probability of piece appearing in emotion-e blogs and $P_{\bar{e}}(piece)$ is the probability of piece appearing in emotion-$\bar{e}$ blogs.

4) Syntactic pieces that have a positive $score_e(piece)$ are added to the E3 lexicon of emotion-e.

By conducting this process, we can automatically construct the E3 lexicons. This method can be used any emotional classification system.

Table 4: The number of words in each E3 lexicon

joy	**sadness**	**fear**
327,702	37,439	13,238

3.5. Evaluation of E3 Lexicons

To evaluate the constructed E3 lexicons, we extracted 50 syntactic pieces from each of the E3 lexicons of joy, sadness and fear. Three examinees described the emotion of the syntactic pieces as RIGHT, WRONG or OTHER. The result is accepted, if 2 or all examinees agree. RIGHT means the lexicon's identification of the emotion and the judgement of the examinees agree. In the case of WRONG, the examinees judged the emotion to be different from the one in the lexicon. OTHER means 'neutral (no emotion)' or 'cannot evaluate'. 'cannot evaluate' means the syntactic piece does not provide sufficient information. One modified entity can't always keep the phrase. Table 5 shows the results of the evaluation. According to these results, the constructed lexicons include a few WRONG emotions. There are fewer WRONG emotions than RIGHT ones, indicating the accuracy of the lexicons. However, in many cases, a single E3 expression is not clearly associated with an emotion.

Table 5: Results of evaluation of the E3 lexicons

	joy	**sadness**	**fear**
RIGHT	19	8	7
WRONG	0	2	3
OTHER	27	35	30

We show a part of the evaluated result as follows:

・RIGHT

joy: 喉を => 潤す
(wet one's whistle)

sadness: ツキが => ない
(get no luck)

・WRONG: The syntactic piece belongs in the sadness lexicon, but it was classed as joy.

joy: 容疑者が => 逮捕される
(crime suspect is arrested)

・OTHER
sadness: 原因を => 調べる
(examine its cause)

4. Generating Reviews

Figure 2 shows the process of review generation. The system collects material from the corpus. The collected reviews are used as templates to generate phrases. The system's input is syntactic piece, which it uses as a review target. The system decides the relevant emotion and selects a review that has same emotion as the target. Syntactic piece and template are fitted together as review sentence.

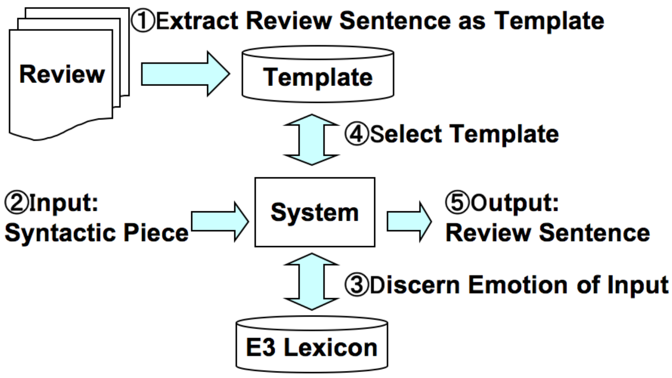

Figure 2: The review generation process

4.1. Review Sentences

In this paper, we define opinion sentences as reviews of stories, such as book or movie reviews. A review consists of four elements: attitude, nominative, object and reason (Nakayama et al., 2005). In their analysis, reviews almost always include the attitude, nominative and object. However, the nominative is not always apparent. Therefore, we regard attitude and object as the basic elements of a review. Nakayama et al. define a sub-element of object as scene, and a sub-element of attitude as emotion. In this study, we consider review sentences, which include object and attitude, especially scene and emotion. The following sentence is an example of a review sentence.

・別れの場面が悲しかった。
(The parting scene is sad.)

Review sentences exist in review articles. We decide on a rule for extracting them heuristically; the rule is based on specific nouns, which exhibit the point of the scene in the review. '場面 (the scene)' is one of the specific nouns. In Japanese, the scene is written before the noun and the impression is written after the noun using a word expressing emotion. We extract the review sentences by the rule from Amazon customer reviews. Amazon customer reviews provide an easy-to-use review corpus. Of the categories used by Amazon, we selected the categories of books and DVDs. A review sentence is used as a template for generating text. Each review sentence includes an emotion word. We describe the requirements for a review sentence as follows:

・The review sentence has a specific noun that exhibits the scene: '場面(scene/locale)', 'シーン(scene)', 'ところ(where)' and 'くだり(line)'.

・The particle (‘が’, ‘は’ or ‘で’) are joined to specific nouns. The particles suggest the object.
・The specific noun modifies the emotion word.

Table 6 shows the number of review sentence templates for each of the emotions we consider. We also show examples of the collected review sentences that are used as templates.

Table 6: The number of review sentence templates for each emotion

joy	**sadness**	**fear**
987	105	49

・ joy: バスでの掛け合いの場面が楽しい。
(The scene of the dialogue on the bus is pleasant.)
・sadness: 誰も信じたくないというくだりは悲しいです。
(The scene where he said ‘I do not believe anyone!’ is sad.)
・fear: 刀でザックリのシーンはちょっと怖いですね。
(The scene where they were ‘slashed with a sword’ is a bit scary.)

4.2. The Naturalness of the Review Sentences

We confirm the relationship between emotional coherence and the naturalness of a sentence. In this paper, we regard E3 as one scene. (We treat the accepted E3 elements in section 3.5.) That is, each E3 is associated with one emotion. Three review sentences are generated based on one scene, each containing a different emotion word (joy, sadness or fear). Each emotion is expressed by ‘喜ばしい(joyous)’, ‘悲しい(sad)’ or ‘恐い(scared)’. For example, the sentences in section 2 were generated by using the scene ‘to receive a present’. We evaluate 90 sentences generated from 30 scenes: 30 sentences are emotionally coherent, and 60 sentences are incoherent. All the coherent sentences, but only 11 of the incoherent sentences, were evaluated as natural. This shows that coherence of emotion affects the naturalness of the sentence. Therefore, review sentences described in section 4.3 are generated from scenes and templates that agree emotionally.

4.3. Review Sentence Generation

To generate emotionally coherent review sentences, a ‘scene’ is input to the system to obtain a review sentence about the scene. A phrase from E3 is used to express the scene as a part of story. A review sentence is made from the phrase and the template, which have the same emotional content, as described in section 4.2. The review sentence is generated by the following processes:

1) Input the syntactic piece into the system.
2) Discern the emotion of input by E3 lexicons.
3) Select a review sentence template that has the same emotion as the E3 phrase.
4) In the selected template, replace the word before the specific noun phrase with the E3 phrase.
5) Output the resulting sentence as the review sentence.
We show examples of generated sentences below.

・Input syntactic piece (E3 phrase): メダルを逃す
(miss the medal)
・Selected template: <u>涙をこらえる</u>シーンは切なすぎる
(The scene where he <u>holds back his tears</u> is agonizing.)

・Output review sentence: メダルを逃すシーンは切なすぎる
(The scene where he misses the medal is agonizing.)

5. Discussions

5.1. Estimating Emotion in Blogs

This section confirms the effectiveness of the E3 lexicons by estimating the emotions expressed in blogs. We treat the emotion of sadness below. We selected 120 blogs, including 45 sad blogs, for which the emotional content is evaluated.

This experiment applies requirement 2 in section 3.2 using the E3 lexicons. That is, the text includes words from the E3 lexicon for emotion e in greater numbers than E3 lexicon words for other emotions. As a result, 20 of the 120 blogs were identified as sad, using the E3 lexicons. Of the 20 blogs, 14 are identified correctly, for an accuracy of 0.70. Moreover, the examined blogs included 3 that were not identified using the emotion lexicons. Therefore, the E3 lexicons have different characteristic compared with the emotion lexicons. The accuracy of 0.70 is not bad, however the E3 lexicons estimated 14 of 45; we obtained the recall of 0.31.

As a result, we can not say whether the E3 lexicons are good for estimating emotion, because the recall is low. However, the obtained result suggests the possibility of effective estimation when we use the E3 lexicons together with emotion lexicons.

5.2. Generated Review Sentences

Generated review sentences are evaluated as natural, because the template used was made from reviews that were collected by strict rules. The review template has emotion words, which are constructed in section 3.1. Emotion word lexicons express clear emotions that do not differ between examinees. Therefore, the collected reviews can be used as templates. Collected templates are nice; there is no incorrect. However, we have a problem. We cannot select the best template. We doesn’t consider the selecting the template in this paper. Therefore, when input one syntactic piece, multiple reviews are generated, one for each template. Take the example of the following generated review.

・初戦を落とすシーンはつらいです。
(The scene where he drops the 1st round is heartbreaking.)
・初戦を落とすシーンが悲しいです。
(The scene where he drops the 1st round is sad.)

Both sentences are natural reviews. We cannot select the better one because we lack input information. In this experiment, input information is one syntactic piece containing only two segments. Therefore, the expressed scene is not clear. In future work, we will input multiple syntactic pieces and generate a review. Increasing the information should make the review target clear.

6. Conclusion

This paper presents a method for generating story reviews using emotion-emerged expressions (E3) that are thought to be the origin of the emotional content. In this work, we have confirmed that the coherence of emotions is an important factor in generating natural sentences. Future work will include increasing the information about the review target to generate more natural sentences.

We have also constructed E3 lexicons using blogs. The lexicons are made up of syntactic pieces that are proposed as units of syntactic structure to represent suitable units of expression for estimating emotions. We finally illustrate the possibility of effective estimation when we use the E3 lexicons together with emotion lexicons.

Tools and language resources

1) IPA Thesaurus, Ver.2.7.0, Matsumoto Lab., Nara Institute of Science and Technology. http://chasen.naist.jp/stable/ipadic
2) Amazon.co.jp. http://www.amazon.co.jp/
3) CaboCha, Ver.0.53, Matsumoto Lab., Nara Institute of Science and Technology. http://chasen.org/˜taku/software/cabocha/
4) TinySVM, Ver.0.09, Matsumoto Lab., Nara Institute of Science and Technology. http://chasen.org/˜taku/software/TinySVM/
5) livedoorBlog. http://blog.livedoor.com

References

Suguru Aoki and Kazuhide Yamamoto. 2007. Opinion extraction based on syntactic pieces. Proc. of PACLIC21, pages 76–86.

Shigeru Fujimura, Masashi Toyota, and Masaru Kitsuregawa. 2004. A consideration of extracting reputations and evaluative expressions from the web. Technical Report of The Institute of Electronics, Information and Communication Engineers, 104:141–146.

Zornitsa Kozareva, Borja Navarro, Sonia Vazquez, and Andres Montoyo. 2007. A headline emotion classification through web information. 4th International Workshop on Semantic Evaluations, pages 334–337.

Rada Mihalcea and Hugo Liu. 2006. A corpus-based approach to finding happiness. Proceedings of AAAI 2006 Spring Symposium on Computational Approaches to Analyzing Weblogs, pages 139–144.

Norio Nakayama, Kouji Eguchi, and Noriko Kando. 2005. Emotion expression model. 11th Annual Meeting of the Association for Natural Language Processing, pages 149–152. (in Japanese)

Syoutarou Ozaki, Sadao Kurohashi, and Makoto Nagao. 1997. Text generation using semantic network. IPSJ SIG Notes, 97(69):133–140.

Robert Plutchik. 1960. The multifactor-analytic theory of emotion. The Journal of Psychology, 50:153–171.

Yohei Seki, Tadashi Iijima, and Ken'ichi Harada. 1999. Semantic focus-driven sentence generation. IPSJ SIG Notes, 99(49):79–85. (in Japanese)

Changhua Yang, Kevin Hsin-Yih Lin, and Hsin-Hsi Chen. 2007. Building emotion lexicon from weblog corpora. Proc. of 45th Meeting of the Association for Computational Linguistics, pages 133–136