

# Extracting Evaluative Conditions from Online Reviews: Toward Enhancing Opinion Mining

Yuki Nakayama

Atsushi Fujii

Department of Computer Science  
Graduate School of Information Science and Engineering  
Tokyo Institute of Technology  
{nakayama.y.aj@m, fujii@cs}.titech.ac.jp

## Abstract

A fundamental issue in opinion mining is to search a corpus for opinion units, which typically comprise the evaluation by an author for a target object from an aspect, such as “This hotel is in a good location”. However, no attempt has been made to address cases where the validity of an evaluation is restricted on a condition in the source text, such as “for traveling with small kids”. In this paper, we propose a method to extract such conditions, namely evaluative conditions, from sentences including opinion units. Our method uses supervised machine learning to determine whether each phrase is a constituent of an evaluative condition. We propose several features associated with lexical and syntactic information, and show their effectiveness experimentally.

## 1 Introduction

Reflecting the rapid growth in the use of opinionated texts on the Web, such as customer reviews, opinion mining has been explored to facilitate utilizing opinions mainly for improving products and decision-making purposes. While in a broad sense opinion mining refers to a process to discover useful knowledge latent in a corpus of opinionated texts, in a narrow sense its purpose is to extract opinions from a corpus. In either case, fundamental issues involve modeling a unit of opinions and searching the corpus for those units, which typically comprise the evaluation by an author for a target object from an aspect.

We take the following review sentence as an example opinionated description.

“I think hotel A is in a good location for traveling with small kids”.

From the above example, existing methods (Pang and Lee, 2008; Seki et al., 2009; Jin et al., 2009; Zhao et al., 2010; He et al., 2011; Liu and Zhang, 2012) for opinion mining extract the following quintuple as an opinion unit.

Target = “hotel A”, Aspect = “location”,  
Evaluation (Polarity) = “good” (positive), Holder = “I (author)”, Time = N/A

Depending on the application, “Evaluation” can be any of a literal evaluation expression (e.g., “good”), a polarity (positive/negative), or a value for multipoint scale rating. However, because this difference is not important in our research, we usually use the term “evaluation”.

Given those structured items extracted from a corpus, it is easy to overview the distribution of values for each element or a combination of elements. Those who intend to improve the quality of hotel A may investigate the distribution of values for “Aspect” in the reviews with “Target=hotel A & Polarity=negative”, while those who look for accommodation may compare the distribution of values for “Aspect & Polarity” in reviews for more than one hotel.

However, no attempt has been made to address cases where the validity of an evaluation is restricted on a condition in the source text. We shall call such a condition “evaluative condition”. In the above example sentence, the evaluation for hotel A (“in a good location”) is valid only “for traveling with small kids”, and it is not clear whether this evaluation is valid irrespective of the situation. The existing methods, which do not analyze evaluative conditions, potentially overestimate or underestimate the utility of hotel A and the quality of opinion mining is decreased accordingly.

To alleviate this problem, we need to introduce evaluative conditions as an element in the opinion unit, such as Condition=“for traveling with small kids”, which enables us to perform deeper

and finer-grained analysis for opinion mining. To avoid any confusion, we consistently use the term “opinion unit” to refer to the traditional quintuple-based unit in which a few elements can be omitted.

Motivated by the above background, in this paper we propose a method to extract evaluative conditions from opinionated corpora. The contribution of our research is introducing the notion of evaluative conditions into opinion mining and proposing a method to extract evaluative conditions from opinionated corpora.

Currently, we target corpora of review text in Japanese. As the first step of research, we focus only on cases where an evaluative condition and an opinion unit are in the same sentence. In addition, we leave the following two research issues as future work.

First, compared with the existing opinion elements, such as Aspect, values for Condition tend to be long and thus it is important to standardize various expressions for the same condition, such as “for traveling with small kids” and “for a family trip with children”. It can be expected that existing methods for paraphrasing alleviate this problem.

Second, it can be useful to subdivide evaluative conditions into general or domain-specific categories, such as “purpose”, “user”, and “situation” in reviews for hotels. For example, those categories can be effective to refine user’s needs in retrieving or recommending products. We show example sentences for several categories, in which the evaluative condition and evaluation expression are in bold and italic faces, respectively.

The room is *large enough* **for a business trip**. (purpose)

The bed is *small* **for people who is 185cm tall**. (user)

**If you stay more than one day**, you will be *tired of* the breakfast. (situation)

I was *content* with the meal **if it was less expensive**. (counterfactual)

**Considering the class of this hotel**, the dinner is *acceptable*. (concession)

## 2 Related Work

Evaluative conditions are related to causes and reasons because all of them have an influence on the validity of the corresponding evaluation in an opinion.

Although causal relations can be divided into inter-sentential and intra-sentential, our current interest is more related to the extraction of intra-sentential relations (Girju, 2003; Chang and Choi, 2004; Inui et al., 2005). These methods generally identify two event-related components in a sentence and determine the type of the causal relationship between those components, if any, such as “precondition”, “cause-effect”, and “consequence”. An event-related component is usually a word, such as “cancer”, or a proposition, such as “he is a heavy smoker”.

However, the above existing methods focused only on specific syntactic patterns, such as “<Clause1, Marker (*tame* in Japanese), Clause2>” (Inui et al., 2005) and “<NP1, Verb, NP2>” (Girju, 2003; Chang and Choi, 2004). In Section 1, none of the example sentences including evaluative conditions matches to those patterns, irrespective of whether in English or in Japanese. Additionally, looking at the examples for “counterfactual” and “concession”, the relation between the evaluation and evaluative condition is different from the causal relation. Besides this, our research is the first attempt to extract cause-like relations in opinion mining.

Kim and Hovy (2006) proposed a method to identify a reason for the evaluation in an opinion, such as “the service was terrible because the staff was rude” and “in a good location close to the station”. However, their purpose is to identify grounds that justify the evaluation, which are different from evaluative conditions.

## 3 Proposed method

### 3.1 Overview

The purpose of our method is to extract one or more evaluative conditions in an opinionated sentence in Japanese. Currently, we assume that both an opinion unit and an evaluative condition are in the input sentence, and that the opinion unit has been identified by an existing automatic method.

Our extraction method follows the BIO chunking classifier, which labels each token in a sentence as being the beginning (B), inside (I), or outside (O) of a span of interest. However, because there is no specific characteristics at the beginning of evaluative conditions in Japanese, we do not use the “B” label. We regard Japanese bunsetsu phrases, which consists of a content word and one or more postpositional particles, as tokens, and ex-

tract a sequence of I-phrases as an evaluative condition. However, phrases in an opinion unit are always classified into O-phrases. We use Support Vector Machine (SVM) to train a binary classifier for bunsetsu phrases and propose several features associated with lexical and syntactic information.

### 3.2 Features for phrase classification

Figure 1 depicts an example of syntactic dependency analysis for a review sentence in Japanese. We used “CaboCha” (Kudo and Matsumoto, 2002) for the dependency analysis. In Figure 1, a rectangle and an arrow denote a phrase and a dependency between two phrases, respectively, and in each phrase we show Romanized Japanese words and their English translations in parentheses.

Looking at Figure 1, by definition the evaluative condition (phrases #3-6) modifies the evaluation expression (phrase #7), but does not modify other opinion elements including the aspect (phrase #2). Also, the evaluative condition ends with specific particles in phrase #6. These properties motivated us to propose the following five features for the binary phrase classification.

**Feature A:** Because an evaluative condition modifies the evaluation expression, they are usually in close proximity to each other. Thus, there should be a pass of dependencies between an I-phrase and the evaluation expression, and a phrase in closer proximity to the evaluation expression is more likely to be an I-phrase. We use the dependency distance (i.e., the number of dependencies) between a phrase in question and the evaluation expression as the value for feature A. The value for a phrase is -1 if there is no pass between that phrase and the evaluation expression. In Figure 1, values for phrases #1, #4, and #8 are 2, 3, and -1, respectively.

**Feature B:** Feature A is not robust against errors of the dependency analysis. To complement this weakness of feature A, we roughly estimate the dependency distance by a phrase distance. In practice, we use the difference between the phrase IDs between a phrase in question and the evaluation expression as the value for feature B. If the evaluation expression consists of more than one phrase, we take the minimum difference. Because Japanese sentence has a post modification structure, in which a modifier is followed by its head, a phrase with a negative value for feature B is usu-

ally an O-phrase. In Figure 1, unlike the case for feature A, the values for phrase #1 is 6.

**Feature C:** Because an evaluative condition does not modify any opinion elements other than the evaluation expression, for the value of feature C we take 0 if there is a pass of dependencies between a phrase in question and a non-evaluation opinion element; otherwise 1. In Figure 1, values for phrases #1, #4, and #8 are 0, 1, and 1, respectively.

**Feature D:** Because an evaluative condition often ends with one or more specific particles, we use the existence (1/0) of those particles in a phrase as the value for feature D. Example particles include “ga (the nominative case)”, “no (of)”, “nitottte (for)”, and “nara (if)”. In Figure 1, values for phrases #1 and #6 are 1 and those for the remaining phrases are 0.

**Feature E:** As in Figure 1, an evaluative condition often consists of a phrase whose value for feature D is 1 and one or more preceding phrases. We use the existence (1/0) of a pass of dependencies between a phrase in question and a phrase whose value for feature D is 1. In Figure 1, values for phrases #3-5 are 1 and those for the remaining phrases are 0.

## 4 Experiments

To evaluate the effectiveness of our method, we used the Rakuten Travel data<sup>1</sup>, which consists of approximately 348,564 reviews for hotels in Japanese. From this data set, we randomly selected 675 reviews and manually annotated quintuples for opinion units and evaluative conditions. We found that 182 reviews include evaluative conditions and decomposed those reviews into sentences. We collected 286 sentences including evaluative conditions and used those sentences as the corpus for experiments. The total number of bunsetsu phrases in our corpus is 2,472, which consists of 761 I-phrases and 1,126 O-phrases in which 585 phrases are elements in opinion units.

We performed 10-fold cross-validation and compared different methods in terms of precision (P), recall (R), and F-measure (F). In Table 1, while “Phrase” denotes the result of the binary classification for bunsetsu phrases, “Condition” denotes that of extracting evaluative conditions as a whole using the BIO classifier. The line “Rule” denotes the result of a rule-based method, which is

<sup>1</sup><http://www.nii.ac.jp/cscenter/idr/rakuten/rakuten.html>

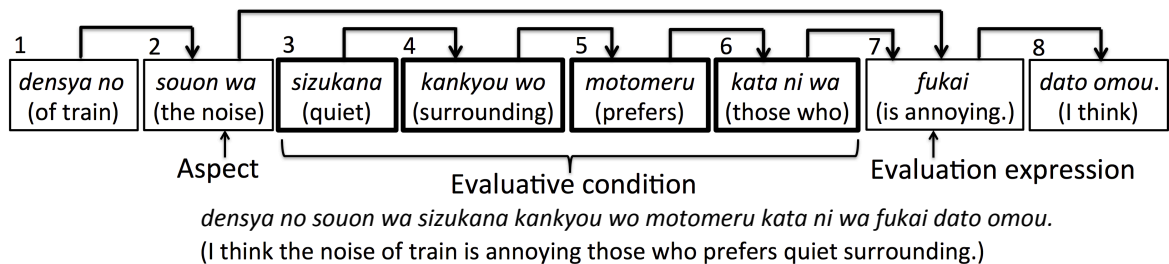


Figure 1: Example of dependency analysis for Japanese.

used as the baseline. This method extracts a bunsetsu phrase ending with one or more specific particles and all phrases from which there is a dependency path to that phrase. For example, in Figure 1 because phrase #6 ends with specific particles, the rule-based method extracts a sequence of phrases #3-#6 as an evaluative condition. The remaining lines denote different combinations of our five features in Section 3.2. While “w/o X” denotes our method without feature X, “All” denotes our complete methods using the five features.

Looking at Table 1, one can see that any variation of our method outperformed the rule-based method irrespective of the configuration, and that our complete method outperformed the remaining of our methods in terms of F-measure. We used the two-tailed paired t-test for statistical testing and found that the differences of “Rule” and “All” in F-measure for “Phrase” and “Condition” were significant at the 1% level. Thus, we conclude that each of our five features was independently effective for extracting evaluative conditions in review sentences and that when used together the improvement was even greater. At the same time, because values for P, R, and F in “Condition” were substantially smaller than those in “Phrase”, we need to improve methods to combine I-phrases and determine the final evaluative condition.

	Phrase			Condition		
	P	R	F	P	R	F
Rule	.539	.614	.553	.407	.412	.410
w/o A	.733	.797	.734	.505	.541	.517
w/o B	.598	.685	.609	.410	.460	.426
w/o C	.719	.789	.725	.490	.524	.500
w/o D	.732	.787	.732	.522	.554	.531
w/o E	.745	.756	.713	.456	.496	.468
All	.730	.792	.735	.538	.571	.548

Table 1: Results for experiments.

## 5 Conclusion

Although a number of methods have been proposed to search an opinionated corpus for opinion units, no attempt has been made to address cases where the validity of an evaluation in an opinion is restricted on a condition in the source text. We proposed a method to extract such conditions, namely evaluative conditions, from sentences including opinion units. Our method performed supervised binary classification to determine whether each phrase is a constituent of an evaluative condition. We proposed five features associated with lexical and syntactic information for Japanese, and show their effectiveness using reviews for hotels. Future work includes addressing research issues discussed in Section 1.

## References

- Du-Seong Chang and Key-Sun Choi. 2004. Causal relation extraction using cue phrase and lexical pair probabilities. In *Proceedings of 1st International Joint Conference on Natural Language Processing*, pages 61–70.
- Roxana Girju. 2003. Automatic detection of causal relations for question answering. In *Proceedings of the ACL 2003 workshop on Multilingual summarization and question answering*, pages 76–83.
- Yulan He, Chenghua Lin, and Harith Alani. 2011. Automatically extracting polarity-bearing topics for cross-domain sentiment classification. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 123–131.
- Takashi Inui, Kentaro Inui, and Yuji Matsumoto. 2005. Acquiring causal knowledge from text using the connective marker *tame*. *ACM Transactions on Asian Language Information Processing*, 4(4):435–474.
- Wei Jin, Hung Hay Ho, and Rohini K. Srihari. 2009. Opinionminer: A novel machine learning system for web opinion mining and extraction. In *Proceedings of the 15th ACM SIGKDD*, pages 1195–1204.

- Soo-Min Kim and Eduard Hovy. 2006. Automatic identification of pro and con reasons in online reviews. In *Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions*, pages 483–490.
- Taku Kudo and Yuji Matsumoto. 2002. Japanese dependency analysis using cascaded chunking. In *Proceedings of the 6th Conference on Natural Language Learning*, pages 1–7.
- Bing Liu and Lei Zhang. 2012. A survey of opinion mining and sentiment analysis. In C.C. Aggarwal and C.X.Zhai, editors, *Mining Text Data*, pages 415–463. Springer.
- Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Journal Foundations and Trends in Information Retrieval*, 2(1–2):1–135.
- Yohei Seki, Noriko Kando, and Masaki Aono. 2009. Multilingual opinion holder identification using author and authority viewpoints. *Journal of the Information Processing and Management*, 45(2):189–199.
- Wayne Xin Zhao, Jing Jiang, Hongfei Yan, and Xiaoming Li. 2010. Jointly modeling aspects and opinions with a maxent-lda hybrid. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 56–65.