社団法人 情報処理学会　研究報告
IPSJ SIG Technical Report

2004－ICS－138　(33)
2004／12／7

FrameNet-Based Shallow Semantic Parsing with a POS Tagger

Nobukazu Shibui, Akito Sakurai
shibui@ae.keio.ac.jp
Faculty of Science and Technology, Keio University
3－14－1 Hiyoshi, Kouhoku-ku, Yokohama 223-0061, Japan

In this paper we propose a FrameNet-based shallow semantic parsing without syntactic parsing. Previous studies on shallow semantic parsing utilize the results of syntactic parsing of input sentences as input data. However, syntactic parsing has well-known shortfalls, such as large amount of computation and insufficient accuracy etc... Furthermore, when use of syntactic parsing is premised, it limits applicable languages, since good syntactic parser is rarely available. To prevent such undesirable consequences in shallow semantic parsing, we propose to use POS tagger instead of syntactic parsing. Our experiments using FrameNetII data as training and test data showed the same level performance as existing methods using syntactic parsing.

1. Introduction

Shallow semantic parsing is “the process of annotating texts with semantic roles specified either using predicate specific labels.” (Baker, Fillmore, & Lowe1998) or predicate independent labels (Kingsbury, Palmer, & Marcus 2002)

Existing Methods of shallow semantic parsing use a probabilistic model (Gildea and Jurafsky [2]), SVM (Sameer Pradjan, Kadri Hacioglu, Wayne Ward, James H. Martin, and Daniel Jurafsky [4]), a maximum entropy model (Michael Fleischman, Namhee Kwon, Eduard Hovy [5]), and a hidden Markov model (Cynthia A. Thompson, Roger Levy, and Christopher D. Manning).

We use SVM, as with Jurafsky et al. [4]. However, in their method using SVM, they created most of input features from the result of syntactic parsing, though use of syntactic parsing passes the problems of syntactic parsing to shallow semantic parsing. Problems of syntactic parsing include the large amount of computation and resulting long lead time and insufficient accuracy. Furthermore, the level of performance of syntactic parsing greatly differs among languages, and when syntactic parsing is premised, it will hinder application of shallow semantic parsing to many languages. To avoid such problems and realize widely applicable shallow semantic parsing, it is necessary to find a method that does not use syntactic parsing. In this paper, we aimed to realize shallow semantic parsing that does not use syntactic parsing.

2. Semantic role and Dataset

Firstly we provide a brief explanation of the semantic role. Then we explain the FrameNet project [1].

To understand the meaning of a sentence, it is necessary to know the interrelationship between words and phrases in the sentence. The semantic role is a unit to represent this interrelationship. To grasp the outline of a sentence, which is the first step to know the meaning of a sentence, it is not necessary to know the interrelationships among all words. It is sufficient to know relationships of words with the verb that composes the core of the meaning of a sentence. In shallow semantic parsing, it is to obtain semantic roles of words to the designated verb and.

This paper uses FrameNetII data for training and testing our method. The FrameNet project is a project conducted by University of California, Berkeley, in which they try to assign semantic roles to words in sentences based on frame semantics and create corpus annotated with semantic roles. At present, FrameNet II Corpus is composed of more than 100000 sentences with semantic roles and other features (e.g. part of speech).

In FrameNet framework relationships between subjects are necessary as background knowledge to understand the meaning of a sentence. We call it frame and represent it by using relationships designated by semantic roles. Each sentence belongs to one of numerous Frames. The semantic role is called Frame Element (FE) in FrameNet and defined for each Frame. In this paper, we use this FE as the semantic role.

3. System Architecture

Fig. 1 shows the system architecture.

We created this method with the method of Jurafsky et al. [3], [4] as a guide. The significant difference between out method and the method of Jurafsky et al. is that our method does not use syntactic parsing and features extracted from its result. By not using syntactic parsing, our method is expected to shorten the required lead time to apply to other languages.

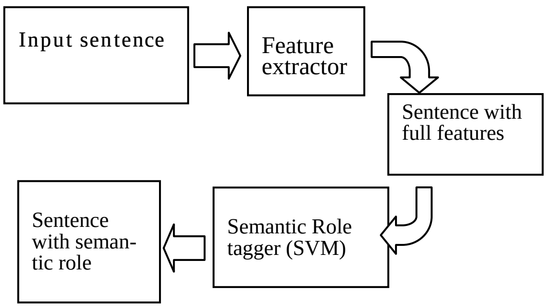

Fig. 1 System architecture

The input data for this system include words, part of speech of each word, the core word of the frame (predicate), and the frame the sentence belongs to. The remaining features necessary to assign semantic roles are obtained by the feature extractor in Fig.1. Each feature will be described in the following section. This system uses all the features obtained this way to assign semantic roles type or "null" to each word. "Null" is an interim tag given to words that do not have a semantic role.

We used YamCha, a generic chunker based on SVM [7], as the classifier to assign the semantic roles. In SVM training, a different model is created for each frame. We use these models to assign semantic roles to sentences that belong to the corresponding frame. Lastly we output sentences with the semantic roles.

3.1 Features

This system uses the following features as classifier input to assign the semantic roles. Fig. 2 shows examples of input data.

a) Predicate
It is the word which composes the core of a Frame. The "Target word" tag is assigned to this in the FrameNet corpus.

b) Part of speech (POS)
It is the part of speech tag of the word.

c) Distance from predicate
This feature indicates the distance from the predicate. The distance index is the number of words between the predicate and the word. When the word is placed before the predicate, the number has a negative sign, and when placed after the predicate, the expression has a positive sign. The value of this feature assigned to the predicate shall be zero.

d) Frame
It is the name of the frame the sentence belongs to.

Word	Frame	pos	Distance	Predicate	FE
her	Clothing	DPS	-2	dress	Wearer
wedding	Clothing	NN1	-1	dress	Use
dress	Clothing	NN1-VVB	0	dress	Garment
for	Clothing	PRP	1	dress	
many	Clothing	DT0	2	dress	
years	Clothing	NN2	3	dress	
waiting	Clothing	VVG	4	dress	

Fig. 2 Examples of input data

3.2 Classification

As we mentioned in the previous section, SVM is trained frame by frame. When we use it we first identify frame and then the SVM gives us semantic role or "null" to each word. In the following, assigning the semantic role to each word is also called classifying the word into the semantic role class.

To classify a word, we use five words and accompanying features, and the semantic role two words before the word to be classified. The five words include the word to be classified and two words before and after the word each. The parts enclosed in Fig.2 are the features to be used to classify "for."

4. Experiment

In this paper, we conducted two types of experiments. Firstly, we conducted a baseline experiment, which used the word and its part of speech only to assign the semantic role. In the second experiment, we used all features.

4.1 Data description

In this experiment, we used sentences that belong to Clothing, Containers, and Quantity Frame from the FrameNet data.

For the Predicate, POS, and Frame features, we used the value in the FrameNet data without modification.

4.2 Method of experiment

For the experiment method, we used the ten-fold cross validation.

5. Result

We used precision, recall, and sentence correctness ratio (in the following we use an abbreviation "Sent") as the evaluation criteria. Precision indicates the correctness the ratio of semantic roles correctly assigned by the system among all semantic roles which the system has assigned. Recall indicates coverage, which is the ratio of semantic roles correctly assigned by the system among all correct answers. The sentence correctness ratio indicates the percentage of sentences where all semantic roles in one sentence are correct among all sentences in the test data. Please note that we did not count the words which semantic role is "null" in the result and answer data set, because the number of "null" is so large that our result would be distorted.

The result of the baseline experiment is shown in Table 1 and the result of the experiment using all features is shown in Table 2.

Table.1 The result of baseline experiment

	Precision	Recall	Sent
Clothing	0.595	0.408	0.218
Containers	0.754	0.531	0.379
Quantity	0.840	0.808	0.515

Comparison of the baseline experiment and the experiment based on all features shows that the experiment based on all features performed far better. We can say that the features used in these experiments (especially distance information) are valid and effective.

Table.2 The result of the experiment using all characteristics

	Precision	Recall	Sent
Clothing	0.823	0.72	0.488
Containers	0.843	0.77	0.589
Quantity	0.868	0.857	0.592

6. Discussion

To evaluate the result of this experiment, we compared the result of conventional methods and that of our experiment.

In Table.3, the row SVM shows the result of the experiment conducted by Pradjan, Hacioglu, Ward, Martin, and Jurafsky using SVM [4]. ME shows the result of the experiment conducted by Hovy, Kwon, and Fleischman using a maximum entropy method [5]. HMM shows the result of the experiment conducted by Mannin, Thompson, and Levy using a Hidden Markov Models [6].

Table.3 Comparison with previous studies

	Precision	Recall	Sent
SVM	0.860	0.810	
ME	0.600	0.554	
HMM	0.739		0.637
Proposed method	0.836	0.755	0.536

The last row indicates the average of the result of this experiment weighted by the number of sentences in the test data for each frame.
Please note that while other studies evaluate the result by frame element, this paper evaluated the result by word.

Table 3 indicates that this method showed the same level performance as other methods, even though it did not use syntactic parsing. This result proved that syntactic parsing might not be necessary in shallow semantic parsing.

7. Conclusion

The experiment result indicates the success of our attempt of performing shallow semantic parsing without syntactic parsing. Our future challenges will include assigning the POS tag automatically to input sentences, evaluating the result not by word but by frame element, and creating a system incorporating these results.

References

[1] Colin F. Baker, Charles J. Filmore, and John B. Lowe, "The Berkley FrameNet Project," Proceedings of Coling/ACL, pp. 86-90, 1998

[2] Daniel Gildea and Daniel Jurafsky, "Automatic Labeling of Semantic Roles," Computational Linguistics, 28:3, pp. 245-288, 2002

[3] Kadri Hacioglu, Sameer Pradhan, Wayne Ward, James H. Martin, and Daniel Jurafsky, "Shallow Semantic Parsing using Support Vector Machines," Technical Report TR-CSLR-2003-1, Center for spoken Language Research, Boulder, Colorado, 2003
[4] Sameer Pradjan, Kadri Hacioglu, Wayne Ward, James H. Martin, and Daniel Jurafsky "Semantic Role Parsing: Adding Semantic Structure to Unstructured Text," Center for Spoken Language Research, Boulder , Colorado, 2003
[5]Michael Fleischman, Namhee Kwon and Eduard Hovy, "Maximum Entropy Models for FrameNet Classification," USC Information Sciences Institute, 2003
[6]Cynthia A. Thompson, Roger Levy, and Christopher D. Manning, "A Generative Models for Semantic Role Labeling," School of Computing, University of Utah and Departments of Linguistics and Computer Science Stanford University, 2003
[7] Taku Kudo, Yuji Matsumoto "Fast Methods for Kernel-Based Text Analysis," ACL 2003