Master’s Thesis

Argument Structure Analysis of Event Nouns Based on Noun-verb Co-occurrences and Noun Phrase Patterns

Mamoru Komachi

February 1, 2007

Department of Information Processing
Graduate School of Information Science
Nara Institute of Science and Technology
A Master’s Thesis
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
MASTER of ENGINEERING

Mamoru Komachi

Thesis Committee:

Professor Yuji Matsumoto   (Supervisor)
Professor Shunsuke Uemura  (Co-supervisor)
Associate Professor Kentaro Inui  (Member)
Argument Structure Analysis of Event Nouns Based on Noun-verb Co-occurrences and Noun Phrase Patterns

Mamoru Komachi

Abstract

The study of predicate-argument structure analysis has focused largely on verbs and adjectives. However, to fully understand a sentence such as “His decision was right,” both the predicate “was right” and verbal noun “decision” need to be identified as event denoting expressions. Here, we call verbal nouns event nouns, which refer to events like “he decided something,” and propose argument structure analysis of event nouns. Some event nouns, such as “cook”, do not always refer to an event and word sense disambiguation of event nouns has to be done depending on event nouns in use.

In this thesis, we present a machine learning method to identify eventness using noun-verb co-occurrences and noun phrase patterns from large corpora. We also investigate a method of incorporating noun-verb co-occurrences to label arguments of event nouns, and conduct error analysis for the argument labeling task.

Keywords:

event, noun, argument structure, co-occurrence, large corpora

* Master’s Thesis, Department of Information Processing, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0551054, February 1, 2007.
動詞と格要素の共起と名詞の出現パターンを用いた
事態性名詞の項構造解析*

小町 守

内容梗概

述語項構造解析として、これまで動詞や形容詞を中心とする研究がなされてきたが、文の理解のためには、たとえば文「彼の決断は正しかった」に対して述語「正しい」に関する解析に加え、「彼が(かなを)決断(する)」という事態に対応する名詞の項構造解析も不可欠である。また、事態性名詞の中には「料理」のように文脈によって事態とならないものもあり、文章中の名詞が事態となるか否か(事態性)の判定も必要である。

本論文では、大規模コーパスから獲得した動詞と格要素の共起と事態性名詞の出現パターンを利用し、機械学習の手法を用いて事態性の判定を行う手法を提案する。また、動詞と格要素の共起を用いて項らしさを判定する方法を検討し、項同定に失敗する事例について分析を行う。

キーワード

事態, 名詞, 項構造, 共起, 大規模コーパス

* 奈良先端科学技術大学院大学 情報科学研究科 情報処理学専攻 修士論文。NAIST-IS-MT0551054, 2007年2月1日。
Contents

1. Introduction .................................................. 1
  1.1 Argument Structure Analysis of Event Nouns .............. 2
  1.2 Description of Annotation ............................... 3

2. Identification of Eventness of Nouns ...................... 5
  2.1 Motivation ............................................. 5
  2.2 Learning Patterns of Noun Phrases ...................... 6
  2.3 Experiments .......................................... 6
  2.4 Discussions .......................................... 8

3. Argument Labeling of Event Nouns ......................... 10
  3.1 Motivation ........................................... 10
  3.2 Utility of Noun-verb Co-occurrence ................... 10
  3.3 Argument Labeling Using Predicate-argument Structure Analysis Model ................................. 14
  3.4 Discussion .......................................... 18

4. Conclusion .................................................. 21

Acknowledgements ............................................. 22

References .................................................... 23

Appendix ......................................................... 25

A. Manual pages of SynCha .................................. 25

B. NAIST Text Corpus annotation guideline for event nouns ........................................ 25

List of Figures

1 An example of the noun phrase pattern “shouhin/go1ds torihiki/trade” 7
2 Comparison of co-occurrence scores between an argument and other nouns ........................................ 13
Classification errors of the co-occurrence model of the web with different number of $|Z|$ of PLSI ........................................ 14

List of Tables

1 Inter-annotator agreement of event nouns ....................... 5
2 The distribution of nominative, accusative and dative cases .... 5
3 Examples of rules obtained for eventness classification .......... 8
4 Features used for eventness classification ........................ 9
5 Results of eventness classification ................................. 9
6 Features used for the argument labeling task .................... 15
6 Features used for the argument labeling task .................... 16
6 Features used for the argument labeling task .................... 17
7 Results of binary classification of arguments .................... 18
8 Error types and instances resulted from syntactic patterns ...... 19
1. Introduction

The central problem of understanding a sentence comes from the analysis of the relationship between nouns and verbs. Verbs are considered to be the most prominent semantic component of a sentence. Verbs or *predicates* have so called *arguments*, which express entities that are centrally involved in the activity of verbs. To correctly analyze predicate-argument structure is important in NLP applications such as information extraction, question answering, text summarization, and machine translation.

Most arguments are expressed as noun phrases, and it is now generally agreed that co-occurrences and syntactic patterns of nouns differ from those of verbs. However, there are a class of nouns that seem semantically similar to predicates. For example, in a sentence like “His decision was right.” one can assume that the noun “decision” refers to an event “X (makes) decision (of) Y.” We call these kinds of nouns *event nouns* and propose argument structure analysis of event nouns. We define *eventness* that the degree of being a referent to an event, an action or a state. If there is a noun that refers to an event, the noun is said to have eventness. The main focus of our study is on the event nouns which consist of verbal nouns and nouns derived from verbs which have argument structure inherited from original verbs. There are a number of event nouns which do not have direct correspondence to verbs, but they are out of the scope of this thesis.

The beginning of predicate-argument structure analysis using machine learning method goes back to Gildea and Jurafsky [1]. They followed FrameNet’s semantics and defined the semantic role labeling task as finding core and adjunct arguments of predicates or verbs. PropBank is one of the corpora annotated with semantic roles, and thanks to the evaluation campaigns carried out at the CoNLL 2004 and 2005 shared tasks, there have been a number of improvements in machine learning methods with respect to semantic role labeling.

NomBank [2] is annotated with these types of event nouns. It focuses on nominalization of verbs in English, and thus it follows PropBank [3] scheme to add a semantic information layer to Penn Treebank [4]. Penn Chinese TreeBank (CTB)[5] is a Chinese version of Penn TreeBank, and it also provides both syntactic and semantic information.

For the task of argument structure analysis of event nouns, the Kyoto Text
Corpus Version 4.0[6] has been annotated with the noun relations including the argument structure of event nouns. We also built a corpus, NAIST Text Corpus Version 1.2\(\beta\)[7], tagged with predicate-argument structure and coreference. We annotated the whole portion of the Kyoto Text Corpus Version 3.0 while the Kyoto Text Corpus Version 4.0 has annotated relevance tags to only 12% of the corpus.

The difference between other corpora tagged with predicate-argument structure and the NAIST Text corpus is that PropBank and NomBank do not label intra-sentential arguments. Since they focus on shallow semantic parsing, PropBank and NomBank have broader coverage and show high inter-annotator agreement (94%, [3]).

In Japanese, Sasano et al. [8] proposes a method to construct a case frame dictionary of nouns. Kawahara et al. [9] describes how to annotate a relevance-tagged corpus. They both regard the argument structure of nouns in a broader sense, and incorporate a variety of relations such as position, attribute, kinship and so on, while we focus on event nouns and exploit specific patterns for argument structures of event nouns.

The remaining sections of this thesis are organized as follows: we first describe the argument structure analysis of event nouns in section 1.1, and then explain how we created an event-tagged corpus in section 1.2. We divide argument structure analysis of event nouns into two sub-tasks: the event classification task and the argument identification task, and we propose an unsupervised machine learning method to learn structures for the classification task from large corpora in section 2, and a supervised machine learning method using co-occurrence of verbs and nouns for the identification task in section 3. Lastly, we conclude our work and present future directions in section 4.

1.1 Argument Structure Analysis of Event Nouns

Argument labeling of event nouns is very similar to the semantic role labeling (SRL) task for verbs. However, compared to argument structure analysis of predicates, there are two major differences between the argument structure analysis of predicates and event nouns. The first one is the problem of polysemous nouns
in terms of eventness,\(^1\) and the second one is the problem of the unit of analysis.

The first problem is that some event nouns have several meanings including an event reading and other readings. For example, the word “look” could refer to the action of looking and the thing being looked. In the sentence “Mary took a look at the boy,” ‘look’ refers to an event of Mary looking at the boy. On the other hand, in the sentence like “The look of the room suddenly changed,” ‘look’ here means the appearance of the room. To overcome this issue, we divide the argument structure analysis of event nouns into two major tasks. The first task is event classification and the second task is argument labeling. Event classification is the task to determine whether event nouns refer to events or not given a context. Argument labeling is the task of identifying arguments of event nouns which refer to events.

To analyze the argument structure of event nouns, we identify eventness first and then we label the arguments for each event noun that has eventness. The event classification task is a binary classification task which decides whether an occurrence of an event noun refers to an event or not, thus we can exploit a large amount of unambiguous instances with regard to eventness. In the following section, we divide the task into two tasks and propose a method to perform unsupervised learning for the event classification task, and show promising results.

### 1.2 Description of Annotation

To analyze the argument structure of event nouns, we have added argument structure information of event nouns to the Kyoto Text Corpus Version 3.0 [6].\(^2\)

In this corpus, we add argument structure (obligatory cases: nominative, accusative and dative) information to event nouns if they have eventness.

For instance, given a passage:

---

\(^1\) Eventness here does not discriminate between a specific instance and the generic sense of an event, and we do not differentiate them.

\(^2\) In addition to event nouns, we also annotate predicate-argument structures, coreferences, and relations. See http://cl.naist.jp/~ryu-i/coreference_tag.html for annotation guidelines.
Necessity of risk management is strongly exclaimed, but nothing can be done because the reality of the market is not yet known. BIS had been investigated a method of survey since the spring of the last year.

We annotate as follows:

- kanri/management [nominative:exophora, accusative:risuku/risk]
- chousa/investigation [nominative:jittai/reality, accusative:BIS/BIS]

If an event noun has arguments, we annotate them in morphological units. (i.e. we do not regard the segment “risuku kanri no hitsuyousei” as the argument of the event noun “kanri.” Instead, we annotate the morpheme “risk” as the argument.)

If an event noun has an argument in a document such as the nominative case “jittai/reality” of a predicate “chousa/survey,” we annotate it even if it occurs outside the sentence. Moreover, if obligatory cases are missing, we annotate them as “[exophora].”

As of January 2007, we annotated the entire Kyoto Text Corpus (2,929 articles, 38,384 sentences), and released it as the NAIST Text Corpus 1.2β.

We asked two annotators to annotate the same portion of the corpus, and calculated inter-annotator agreement of the corpus for 30 randomly selected sentences. The results are shown in Table 1.

With regard to inter-annotator agreement, eventness annotation and nominative/accusative case annotation achieve high results, although the annotation of the dative case shows poor inter-annotator agreement. The reason seems that the dative case is less obligatory than other two cases and two annotators do not agree on the annotation criterion. We plan to build and use a case frame dictionary based on verbs to minimize mis-classification of case frames.

---

3 We label exophora into three categories: first person, second person and others.
4 http://cl.naist.jp/nlcldata/corpus/
Table 1. Inter-annotator agreement of event nouns

<table>
<thead>
<tr>
<th></th>
<th>precision</th>
<th>recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>eventness</td>
<td>0.965 (247/256)</td>
<td>0.792 (247/312)</td>
</tr>
<tr>
<td>nominative</td>
<td>0.735 (191/260)</td>
<td>0.743 (191/257)</td>
</tr>
<tr>
<td>accusative</td>
<td>0.827 (86/104)</td>
<td>0.869 (86/99)</td>
</tr>
<tr>
<td>dative</td>
<td>0.389 (7/18)</td>
<td>0.583 (7/12)</td>
</tr>
</tbody>
</table>

Table 2. The distribution of nominative, accusative and dative cases

<table>
<thead>
<tr>
<th></th>
<th>same</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>chunk</td>
<td>intra-sentential</td>
<td>inter-sentential</td>
<td>exophora</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>event</td>
<td>argument</td>
<td>total</td>
<td>total</td>
</tr>
<tr>
<td>nominative</td>
<td>2024(13%)</td>
<td>5220(33%)</td>
<td>1208(8%)</td>
<td>15583</td>
<td>5185</td>
</tr>
<tr>
<td>accusative</td>
<td>5519(55%)</td>
<td>3127(31%)</td>
<td>297(3%)</td>
<td>10045</td>
<td>854</td>
</tr>
<tr>
<td>dative</td>
<td>841(49%)</td>
<td>352(20%)</td>
<td>66(4%)</td>
<td>1729</td>
<td>201</td>
</tr>
</tbody>
</table>

The distribution of each case (nominative/accusative/dative) in the NAIST Text Corpus 1.2β is shown in Table 2.

From Table 2, we can see that nominative cases distribute widely, whereas a large number of accusative cases exist near the event nouns inside the same sentences. For argument labeling task of nominative case, it is necessary to identify inter-sentential arguments for broader coverage.

2. Identification of Eventness of Nouns

2.1 Motivation

Event nouns which refer to events often have their arguments in the same segment or dependent segment, such as “risuku kanri/the risk management” and “kare no
ketsudan/his decision”. If we are able to acquire event noun patterns, we can improve the performance of event classification task.

In order to capture the patterns of nouns, we used a tree classifier called BACT $^5$ [10]. BACT iteratively selects weighted sub-trees effective to classify training data by learning tree structure using the boosting algorithm. If it is given a sentence structure as features, it learns syntactic patterns as rules to classify training data.

2.2 Learning Patterns of Noun Phrases

To learn the patterns of noun phrases, we encoded sequences of morphemes (window size 3) and the sequence of morphemes of dependent segment into a tree structure. Positive instances were defined in terms of eventness. Positive instances should be unambiguous event nouns which only have eventness reading, so we used nouns which exist among the concept node “Noun-Abstract-Eventness-\{Action,Event\}” in Japanese Lexicon[11]. Also, negative instances were defined as proper nouns and general nouns among the node “Noun-Object” in the thesaurus. We used articles from a month of Mainichi Newspapers (positive instances:117,581, negative instances:282,419).

Acquired rules with high weight through these processes are shown in Table 3. If the sum of the weight of all rules is positive BACT outputs the instance as positive, while the sum of the weight of all rules is negative BACT outputs the instance as negative. Therefore, rules with high positive weight are strong indication of eventness and rules with high negative weight are strong indication of not having eventness.

2.3 Experiments

For the event classification task, we used Support Vector Machines $^6$ [12] to learn the context of event nouns, and ran 10-fold cross validation to evaluate the performance of our system.

$^5$ http://chasen.org/~taku/software/bact/
$^6$ http://www.chasen.org/~taku/Software/TinySVM/
Figure 1. An example of the noun phrase pattern “shouhin/goods torihiki/trade”
Table 3. Examples of rules obtained for eventness classification

<table>
<thead>
<tr>
<th>Rules with positive weights for eventness</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>the chunk has verbal noun in sequence</td>
<td>1.13</td>
</tr>
<tr>
<td>follows particle and verbal noun</td>
<td>1.01</td>
</tr>
<tr>
<td>follows noun suffix</td>
<td>0.61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rules with negative weights for eventness</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>follows verbal noun</td>
<td>-0.92</td>
</tr>
<tr>
<td>precedes verbal noun</td>
<td>-0.54</td>
</tr>
<tr>
<td>follows particle and precedes verbal noun</td>
<td>-0.50</td>
</tr>
</tbody>
</table>

The baseline system uses some heuristic structures ("A no B/A of B" and compound nouns) as features, and we compared it with the model which takes into account the patterns of event nouns obtained from BACT. The features used in the experiment are summarized in Table 4. Evaluation was performed on 80 articles (800 sentences) from newspapers. The evaluation set contains 1,237 event nouns (590 instances have eventness).

As shown in Table 5, noun phrase patterns improve performance of eventness classification. Using only local information like adjacent morpheme sequences with a three word window and morpheme sequences in the same chunk and dependent chunks dramatically outperforms eventness classification without noun phrase patterns in recall, so it would be effective to include dependency structure within a chunk.

2.4 Discussions

Typical errors for eventness classification are caused by not using co-occurrence information effectively. The current model does not make any distinction between arguments inside and outside a sentence in terms of co-occurrence, and only uses morpheme sequences found in the sentence. This makes it difficult to capture effective combination of co-occurrence and syntactic patterns.
Table 4. Features used for eventness classification

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WF</td>
<td>word form of the event noun</td>
</tr>
<tr>
<td>HEAD_WF</td>
<td>word form of the head word in the same chunk</td>
</tr>
<tr>
<td>HEAD_POS</td>
<td>part-of-speech of the head word in the same chunk</td>
</tr>
<tr>
<td>BGH_ID</td>
<td>category id of the event noun in Bunrui Goi Hyou [13]</td>
</tr>
<tr>
<td>EDR_PATTERN</td>
<td>1 if surrounding noun phrases satisfy selectional restrictions of event nouns from EDR[14]; otherwise 0</td>
</tr>
<tr>
<td>A_NO_B</td>
<td>head word of noun A in the expression of “A no B”</td>
</tr>
<tr>
<td>CHUNK_BGH</td>
<td>category id for each noun in the same chunk if event noun is compound noun.</td>
</tr>
<tr>
<td>PREV_CHUNK_POS</td>
<td>morpheme sequences of previous chunks</td>
</tr>
<tr>
<td>NEXT_CHUNK_POS</td>
<td>morpheme sequences of next chunks</td>
</tr>
<tr>
<td>CHUNK_POS</td>
<td>parts-of-speech of the morphemes in the same chunk</td>
</tr>
<tr>
<td>VN_BEFORE</td>
<td>1 if previous morpheme is verbal noun; otherwise 0</td>
</tr>
<tr>
<td>VN_GENERAL BEFORE</td>
<td>1 if previous morpheme sequence is general noun + verbal noun; otherwise 0</td>
</tr>
<tr>
<td>VN_AFTER</td>
<td>1 if next morpheme is verbal noun; otherwise 0</td>
</tr>
<tr>
<td>VN_PARTICLE_AFTER</td>
<td>1 if next morpheme sequence is verbal noun + particle; otherwise 0</td>
</tr>
<tr>
<td>VN_GENERAL AFTER</td>
<td>1 if next morpheme sequence is verbal noun + general noun; otherwise 0</td>
</tr>
<tr>
<td>VN_COMPOUND</td>
<td>number of verbal nouns in the chunk</td>
</tr>
</tbody>
</table>

Table 5. Results of eventness classification

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without noun phrase patterns</td>
<td>72.3%</td>
<td>58.7%</td>
</tr>
<tr>
<td>With noun phrase patterns</td>
<td>73.3%</td>
<td>80.2%</td>
</tr>
</tbody>
</table>
To classify the former examples, it is necessary to include more contextual information into our argument estimation. Also, to correctly analyze the latter example, it is necessary to utilize co-occurrences between the event noun and its arguments. There is much room for improvements in combining co-occurrence information and syntactic patterns.

3. Argument Labeling of Event Nouns

3.1 Motivation

The second task of argument structure analysis of event nouns is argument labeling. Argument labeling is a task of finding arguments of an event denoted by an event noun.

As shown in Table 2, a half of the accusative cases of event nouns occur in the same chunk of event nouns. Since arguments of event nouns usually do not occur in the head position of the chunk, it is necessary to use information of word composition. Moreover, in comparison to predicates, event nouns have inter-sentential arguments 1.5 times higher than predicates[7]. Thus, it is difficult to label arguments of event nouns putting emphasis on structural information.

To solve this problem, we investigate a method to incorporate information other than structure of a sentence. To begin with, we tried to utilize noun-verb co-occurrences for the argument labeling task.

3.2 Utility of Noun-verb Co-occurrence

To confirm that we can use noun-verb co-occurrence model for the argument labeling task, we explored noun-verb co-occurrence model using only verbal nouns which have verbs of the same surface form. In doing so, we assume that the co-occurrence of noun-verb can be converted into event nouns.

Fujita et al. [15] proposed a noun-verb co-occurrence model of joint probability $P(\langle v, c, n \rangle)$ where noun $n$ depends on verb $v$ via case marker $c$. They regard $\langle v, c, n \rangle$ as co-occurrence of $\langle v, c \rangle$ and $n$, and used PLSI[16] to estimate the joint probability.
\[ P(\langle v, c, n \rangle) = \sum_{z \in Z} P(\langle v, c \rangle | z) P(n | z) P(z) \]

\( Z \) is a random variable which denotes latent semantic classes, and PLSI reduces the dimension of the word matrix into \(|Z|\) using the probability distribution. \( P(\langle v, c \rangle | z), P(n | z), \) and \( P(z) \) can be estimated by EM algorithm.

We used point-wise mutual information\cite{17} for co-occurrence measurement. We can calculate point-wise mutual information given \( P(\langle v, c, n \rangle), P(n | z), \) and \( P(z) \).

\[
PMI(\langle v, c \rangle, n) = \log \frac{P(\langle v, c, n \rangle)}{P(\langle v, c \rangle)P(n)}
\]

For using co-occurrence model, we consider it necessary to compare the argument with other nouns in a sentence in order to estimate the argument, and we paired an argument noun and other nouns and built a binary classifier using co-occurrence model. A binary classifier of arguments given a pair of nouns is independent from the implementation of argument labeling model, and can be incorporated into round-robin competition model that outputs the noun which won the most in the combination of all pairs of event nouns and other nouns, and knockout tournament model that outputs the noun which won the championship of a single-elimination tournament of all the nouns in a sentence.

Experiments were carried out on 137 news articles (1,226 sentences) of the same day from NAIST Text Corpus. We used event nouns which have nominative case in a sentence. We paired an argument and other nouns in the sentence, and plotted a correlation graph in Figure 2.

We were not able to calculate co-occurrence score in 18% of the nouns in the data (mostly proper nouns), so we re-calculate co-occurrence score using named entity class obtained from CaboCha\cite{8} if the score is negative or missing. Even after using named entity class, if the score is still negative or missing, we used POS tags of proper nouns from ChaSen\cite{9}. After these treatment, the out-of-vocabulary rate reduced from 18% to 9%.

\footnote{We compared using co-occurrence probability mutual information, \( \chi^2 \) co-efficient, and Dice co-efficient. We show only mutual information result because mutual information gives the best correlation from our observation.}

\footnote{http://chasen.org/~taku/Software/CaboCha/}

\footnote{http://chasen.naist.jp/hiki/ChaSen/}

11
In Figure 2, instances over $f(x) = x$ are the instances that the score of the argument is greater than the score of other nouns, and instances under $f(x) = x$ are those that the score of other nouns is greater than that of the argument.

The ratio of instances in the second and fourth quadrants in the figure is 71.2% (9,715 instances), and the accuracy achieves 90.0% if we consider the argument to be the noun with greater co-occurrence score. In other words, if we take only the second and fourth quadrants, we can label arguments with high accuracy using only co-occurrence scores.

There are 28.1% (3,839) instances in the first quadrant, and the accuracy is 55.8%. Taking into account that the chance rate is 50% (because of the binary classification), it is clear that we cannot determine the argument in the first quadrant using co-occurrence score alone. Instances in the third quadrant are the pairs where both the argument and the other nouns are not likely to occur with event nouns. However, these examples include low score instances from data sparseness and we can map them into other three quadrants by re-estimating co-occurrence score using clustered nouns with semantic classes like human and artifact. We need to study other means to classify the remainder in both the first and third quadrants.

Furthermore, we compared the co-occurrence model calculated from 30 years of newspapers and the co-occurrence model estimated from extra-huge corpus. We used a web corpus of 500 million sentences collected by Kawahara et al. [18].

The evaluation methodology is the same as newspapers. We compared an argument with other nouns in the sentence, and positive examples are the ones in which the argument has greater co-occurrence score, while negative examples are the ones in which the argument has lower co-occurrence score.

Given that the optimal value of the number of the hidden classes $|Z|$ in PLSI is not obvious, we conducted an experiment with different values of $|Z|$ and plotted classification error rates in Figure 3. The co-occurrence model of the web achieved the best accuracy 87.2% (11,894/13,640) when $|Z| = 4,000$. If we apply the co-occurrence model of newspapers to the problem, we obtained accuracy.

---

10 Data courtesy of Daisuke Kawahara and Sadao Kurohashi

11 We did not use low-frequency words that occur less than 300 times for computational reason. If we use mixed co-occurrence models of newspapers and the web, the out-of-vocabulary rate reduced from 9% to 6%.
Figure 2. Comparison of co-occurrence scores between an argument and other nouns
90.4% (12,335/13,640). Although the performance of the co-occurrence model of the web is worse than that of newspapers, the reason seems that the test set comes from newspapers so that the distribution of the co-occurrence in the test set is different from the training set. We plan to look into more on the resource selection for the construction of co-occurrence model and the domain adaptation of a co-occurrence model in the future.

3.3 Argument Labeling Using Predicate-argument Structure Analysis Model

As we described in the previous section, if an instance is in the first quadrant (i.e. the argument and noun have positive correlation) we need to explore information
other than co-occurrence. Thus, we targeted event nouns which have nominative, accusative and dative cases in a sentence and conducted preliminary experiments on building a binary classifier of the argument and other nouns using the same features proposed by the argument structure analyzer of predicates[19]. After that, we manually examined remaining instances in the first quadrant.

We used NAIST Text Corpus 1.2β for evaluation. Training and test sets were taken from two days of newspapers annotated by two annotators. We discarded tags on which two annotators did not agree. The training set consists of a day of newspapers (137 articles) and the test set consists of another day of newspapers (150 articles). We used Support Vector Machines[12] for the classifier. The implementation we used was TinySVM13. We used polynomial kernel function with second degree (d = 2). Other parameters were set to default.

Features used in this model were almost identical to the model proposed by Iida et al.[19]. Features which were not able to be defined for even nouns were eliminated.14 Since argument labeling is performed on morphological units, we added positional features within a chunk and grammatical features of POS of morphemes in a chunk. Moreover, we added a binary feature to test whether event nouns depend on function verbs15, and a binary feature to test whether the argument depends on function verbs if the last feature is true. All the novel features for argument structure analysis of event nouns are indicated by the asterisk mark ‘*’.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SURFACE</td>
<td>surface form</td>
</tr>
</tbody>
</table>

Table 6: Features used for the argument labeling task

---

12 http://cl.naist.jp/nldata/corpus/

13 http://chasen.org/~taku/Software/TinySVM/

14 Because event nouns syntactically differ from verbs we cannot set voice and adverb features. Also, we did not use heuristic features based on centering theory.

15 Function verbs are the verbs that leave the meaning of a sentence to other nouns and only function grammatically. For example, in a sentence like “kandou/impression wo/[accusative] ataru/give,” a function verb “ataeru/give” does not have its original meaning and it in fact used as a paraphrase of “kandousuru/be impressed.” 128 expressions of such function verbs are taken from Muraki [20]
<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grammatical features</strong></td>
<td></td>
</tr>
<tr>
<td>POS</td>
<td>part-of-speech</td>
</tr>
<tr>
<td>CHUNK_POS*</td>
<td>parts-of-speech of the morphemes in the same chunk</td>
</tr>
<tr>
<td>DEFINITE</td>
<td>1 if NP contains the article corresponding to DEFINITE ‘the,’ such as ‘sore’ or ‘sono’; otherwise 0</td>
</tr>
<tr>
<td>DEMONSTRATIVE</td>
<td>1 if NP contains the article corresponding to ‘that’ or ‘this’, such as ‘kono,’ ‘ano’; otherwise 0</td>
</tr>
<tr>
<td>PARTICLE</td>
<td>particle followed by NP, such as ‘wa (topic),’ ‘ga (subject),’ and ‘o (object)’</td>
</tr>
<tr>
<td><strong>Semantic features</strong></td>
<td></td>
</tr>
<tr>
<td>NEWS_COOC</td>
<td>the score of well-formedness model estimated from a large triplets &lt;Noun, Case, Predicate&gt; obtained from newspapers[15](real number)</td>
</tr>
<tr>
<td>WEB_COOC*</td>
<td>same as above except that triplets are extracted from the web(real number)</td>
</tr>
<tr>
<td>NEWS_COOC_DIFF*</td>
<td>difference of co-occurrence scores between candidate pair (real number)</td>
</tr>
<tr>
<td>WEB_COOC_DIFF*</td>
<td>same as above except that scores are calculated from web co-occurrence(real number)</td>
</tr>
<tr>
<td>SELECT_REST</td>
<td>1 if NP satisfies selectional restrictions in Nihongo Goi Taikei (Japanese Lexicon); otherwise 0</td>
</tr>
<tr>
<td>PRONOUN_TYPE</td>
<td>pronoun type of NP. (e.g. ‘kare(he)’ → PERSON, ‘koko(here)’ → LOCATION, ‘sore (this)’ → OTHERS</td>
</tr>
<tr>
<td>EDR_HUMAN</td>
<td>1 if NP is included among the concept ‘a human being’ or ‘attribute of a human being’ in EDR dictionary; otherwise 0</td>
</tr>
<tr>
<td>NE</td>
<td>named entity of NP: PERSON, ORGANIZATION, LOCATION, ARTIFACT, DATE, TIME, MONEY, PERCENT and N/A</td>
</tr>
</tbody>
</table>

**Positional features**

Table 6: Features used for the argument labeling task
We built a binary classifier that takes a pair of nouns and classifies which one is the argument. We only used event nouns which have their arguments in a sentence. The correct arguments and all the other nouns in a sentence were used for training. To evaluate the system’s performance, each argument in the test set was paired with other nouns in a sentence, and for each pair the classifier outputs which noun is the argument. Table 7 shows the experimental result of the classifier using all the features and the results of the classifiers without using some feature set defined in Table 6. For comparison purpose, the result of a simple classifier that outputs whichever noun has higher co-occurrence score is included in the table.

From this experiment, we can see that positional features, especially the fea-

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SENTNUM</td>
<td>distance between NP and PRED</td>
</tr>
<tr>
<td>SENT_Begin</td>
<td>1 if NP is located in the first chunk of sentence; otherwise 0</td>
</tr>
<tr>
<td>SENT_END</td>
<td>1 if NP is located in the last chunk of sentence; otherwise 0</td>
</tr>
<tr>
<td>NP_Begin</td>
<td>1 if NP is located in the beginning of chunk; otherwise 0</td>
</tr>
<tr>
<td>NP_End</td>
<td>1 if NP is located in the end of chunk; otherwise 0</td>
</tr>
<tr>
<td>PRED_NP</td>
<td>1 if PRED precedes NP; otherwise 0</td>
</tr>
<tr>
<td>NP_PRED</td>
<td>1 if NP precedes PRED; otherwise 0</td>
</tr>
<tr>
<td>DEP_PRED</td>
<td>1 if NP depends on PRED; otherwise 0</td>
</tr>
<tr>
<td>DEP_PRED_PATH2NP*</td>
<td>1 if NP’s ancestor depends on PRED; otherwise 0</td>
</tr>
<tr>
<td>DEP_NP</td>
<td>1 if PRED depends on NP; otherwise 0</td>
</tr>
<tr>
<td>DEP_NP_PATH2PRED*</td>
<td>1 if PRED’s ancestor depends on NP; otherwise 0</td>
</tr>
<tr>
<td>DEP_FUNC_PRED*</td>
<td>1 if PRED depends on function verbs[20]; otherwise 0</td>
</tr>
<tr>
<td>DEP_FUNC_NP*</td>
<td>1 if NP depends on function verbs; otherwise 0</td>
</tr>
<tr>
<td>IN_QUOTE</td>
<td>1 if NP exists in the quoted text; otherwise 0</td>
</tr>
</tbody>
</table>

Table 6: Features used for the argument labeling task
Table 7. Results of binary classification of arguments

<table>
<thead>
<tr>
<th>Feature</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>All features</td>
<td>476 (3.5%)</td>
</tr>
<tr>
<td>- Lexical features</td>
<td>510 (3.7%)</td>
</tr>
<tr>
<td>- Web co-occurrence features</td>
<td>502 (3.7%)</td>
</tr>
<tr>
<td>- Newspaper and web co-occurrence features</td>
<td>684 (5.0%)</td>
</tr>
<tr>
<td>- Semantic features</td>
<td>510 (3.7%)</td>
</tr>
<tr>
<td>- Grammatical features</td>
<td>644 (4.7%)</td>
</tr>
<tr>
<td>- Positional features</td>
<td>700 (5.1%)</td>
</tr>
<tr>
<td>Only newspaper co-occurrence features</td>
<td>1,305 (9.6%)</td>
</tr>
</tbody>
</table>

ture that refers to the position of the argument and event nouns in the chunk, were most effective for the argument labeling task of event nouns. In addition, co-occurrence features were the second effective. It turned out that co-occurrence information has more impact than grammatical information on argument labeling.

3.4 Discussion

We examined 112 instances which exist in the first quadrant in the co-occurrence model of newspapers and whose argument score is lower than those of the other noun. We found that most of the errors come from not capturing syntactic patterns specific to event nouns.

Inability of capturing syntactic patterns (62 instances)

The most frequent errors were the instances failing to capture syntactic patterns of event nouns.

This is a similar phenomena described in Meyers et al.[21]. They report that event nouns have certain syntactic patterns in English. For example, in a light verb construction like “Mary took a walk.” the event noun “walk” shares the
Table 8. Error types and instances resulted from syntactic patterns

<table>
<thead>
<tr>
<th>Error type</th>
<th># of instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharing the argument with verbs</td>
<td>19</td>
</tr>
<tr>
<td>Case alternation of arguments</td>
<td>13</td>
</tr>
<tr>
<td>Arguments in compound nouns</td>
<td>12</td>
</tr>
<tr>
<td>Arguments in adnominal clauses</td>
<td>6</td>
</tr>
<tr>
<td>Arguments in the form of “A no B/A of B”</td>
<td>6</td>
</tr>
<tr>
<td>Long-distance dependency</td>
<td>6</td>
</tr>
</tbody>
</table>

argument with the support verb “took.” By making use of syntactic patterns of event nouns that share arguments with the predicates, we can utilize the result of the argument structure analysis of predicates to the argument structure analysis of event nouns.

For this experiment, although we added the function verbs listed in Muraki[20], expressions like “tenkai suru/to develop” were left out. According to his definition, function verbs are said to carry little meaning and put more emphasis on event nouns while changing the aspect, voice and modality of the event nouns. However, there seems to exist a broader class of verbs including function verbs that share arguments with event nouns. Because of the number of mis-classified examples it is most important to enumerate and encode the syntactic patterns of event nouns, and we are going to exploit the syntactic patterns of this kind first.

Next recurring examples were case assignment errors of dependent nouns. There are a number of case alternation studies in linguistics, and we will investigate more on this phenomena in the future.

Other instances in this type include arguments in compound nouns, adnominal clauses and “A no B/A of B”, and these are the issues we plan to work next. All the classified sub-items are summed up in Table 8.
Failure in resolving coreference (13 instances)

There are 8 instances that compared semantically equivalent nouns, which are not necessary to compare in the first place if we had performed coreference resolution. Still, there are 5 instances that refer to different entities even though they have the same surface form.

In examples like:

\[
\ldots \text{maruchi media shijou/market wo “shijou/market kibo hyaku ni juu san chou en, koyou soushutsu/creation ga ni hyaku yon juu man nin” to yosou shita ga \ldots}
\]

(\ldots although the market of multimedia was expected to “the market size was 2.3 trillion Japanese yen and the creation of employment was 1.4 million people”\ldots)

The topic of the market does not refer to an entity but a class of market, so we do not need to make distinction between shijou/market and shijou/market.

However, in examples like:

\[
\ldots \text{roshia gun butai/troop wa itsuka, shuto no suukasho de dyudaefu seiken butai/troop to hageshii sentou/battle wo kurikaeshita.}
\]

(\ldots on fifth, Russian army troops had severe fierce battles repeatedly with Dudaev troops on several places. )

\[
\text{two instances butai/troops and butai/troops refer to different entity and thus they need to be treated as separate expressions. The former examples occur in 8 instances and the latter examples occur in 5 instances.}
\]

Others (37 instances)

Various reasons are found in the remaining examples. There are 10 examples that need some kind of information outside the document like bridging reference.

For other 27 instances, 15 instances can be handled easily like two nouns are linked with a conjunction and it is not necessary to compare them, while 12 instances seems difficult to solve immediately. One of such examples is the event noun occurs in a quotation and the argument exists outside the quotation structure.
4. Conclusion

In this thesis we explained the argument structure analysis of event nouns. We described our annotation scheme of event nouns and defined two tasks for argument structure analysis. We conducted two experiments for each task.

First, we proposed an unsupervised method for identifying the eventness of nouns using noun phrase patterns. From our experiments, the precision and recall of identification of eventness were 76.6% and 79.6%.

Second, we reported preliminary results on the argument labeling task of event nouns using noun-verb co-occurrences from large corpora, and examined the errors of a binary classifier of arguments. We successfully showed that the co-occurrence information is effective for the argument labeling task. While the accuracy of the binary classifier of arguments using all features achieved 96.5%, we found that most errors outside the scope of co-occurrence model were caused by not capturing syntactic patterns typical for event nouns that predicates share arguments with event nouns.

It is still an open question that how we can incorporate a co-occurrence model into the argument structure analyser of event nouns. In this thesis, we used two co-occurrence models and syntactic patterns as features of machine learning methods, but they could be used in other ways. For example, we could use co-occurrence scores to filter noun candidates for argument labeling, and if there are certain syntactic patterns we could directly determine arguments without comparing the argument with other nouns. We plan to investigate more on the model of argument structure analysis and evaluate the whole system after implementing an argument structure analyzer of event nouns.
Acknowledgements

I am grateful to my advisor, other collaborators, and fellow students. I am especially fortunate to have been advised by Yuji Matsumoto. He has taught me how enjoyable research would be. I am also very happy to be supervised by Kentaro Inui. The work in this thesis was shaped by numerous discussions with him.

Many thanks to Ryu Iida for teaching almost everything about doing research and being part of the academic community. The last several months of my master’s course could not be better without the collaboration with him. Also I am very happy to be a member of paraphrasing group to get acquainted with Atsushi Fujita. He gave me valuable advice on multiple occasions.

I am also grateful to Takahiko Ito for showing me how to do original work. I learned from him the diligence and effort finally smile. And thanks to Taku Kudo for opening my eyes to the profound machine learning technique and programming. I sometimes feel like following the steps he walked long time ago.

Thanks to Eric Nicholes and Noah Evans for helping me writing English. Especially Eric is an excellent person who not only corrects my English but also makes insightful comments to my work. I enjoyed talking about “geek” topics.

I am very happy to have been able to finish the SICP study meeting with Shuya Abe, Ai Azuma, Ken’ichi Ishibashi, Hiroki Kuzuno and Osamu Matsumoto. SICP study meeting was the most exciting and invaluable time of my experience as a programmer.

Words cannot express my gratitude to Nozomi Kobayashi for having merry and happy discussions and tea times. Without her, the second year of my master would have been miserable. I cannot imagine the day she will leave the laboratory.

I would also like to thank to Hanako Aoyama. She taught me the way to enjoy my life. The last two years of my life could be the happiest without her. I will miss everything we shared.

Finally, many thanks to my parents Hiroshi and Kumie, my brothers Jun, Manabu and Susumu, for their support and welcoming me at every and any time I come home.

I gratefully acknowledge the financial support through the Excellence Student Program funded by Nara Institute of Science and Technology.
References


Appendix

A. Manual pages of SynCha

B. NAIST Text Corpus annotation guideline for event nouns