AUTOMATIC EXTRACTION OF TRANSLATION PATTERNS FROM BILINGUAL LEGAL CORPUS

Makoto Ohara†, *Shigeki Matsubara*‡ *and Yasuyoshi Inagaki*§

Graduate School of Engineering, Nagoya University †
Information Technology Center, Nagoya University ‡
Faculty of Information Science and Technology, Aichi Prefectural University §
Furo-cho, Chikusa-ku, Nagoya, 464-8603 Japan
makoto@inagaki.nuie.nagoya-u.ac.jp

ABSTRACT

The multilingualization of legal documents is desirable for promoting the internationalization of the society. Since it is vital to choose proper terms when translating legal documents, which include technical terms and unique patterns, it is desirable to compile bilingual dictionaries for each legal domain. Compiling basic bilingual dictionaries for legal documents, however, is a difficult task because of the great range of legal documents. This paper describes a method for automatically extracting translation patterns for legal document translation by using legal documents and their translated documents. The proposed method extracts translation patterns with Japanese bunsetsu-level units from legal sentences and the translated sentences that are properly aligned with each other. The proposed method utilizes three indexes for pattern extraction: bilingual dictionaries, statistical co-occurrence information on the parallel corpus, and syntactic information based on dependency grammar. We have extracted translation patterns from the Japanese Civil Code and its translation. The result has provided 80.5% precision and 49.1% recall, and the extracted translation patterns will be useful for translating legal documents and helping to construct a Japanese-English legal dictionary.

Keywords: translation pattern, legal document, bilingual corpus, dictionary, statistical alignment, dependency structure.

1. INTRODUCTION

In recent years, the multilingualization of Japanese legal documents, which include laws and judicial precedents, has been desirable to promote the internationalization of the society. Since we need the knowledge not only of both languages, but also of legal operations, while translating legal documents, the task is not easy for non-specialists. Besides, large quantities of legal documents are expected to be translated. This makes it desirable to improve multilingualization support systems, which include electronic dictionaries and machine translation systems to reduce the cost of translation.

In order to accurately translate legal documents by hand or to realize high-quality machine translation, high-quality bilingual dictionaries are necessary. Although large-scale electronic dictionaries have been compiled in recent years [5], they are not necessarily adequate due to the necessity for domain-specific bilingual dictionaries for legal documents.

On the other hand, in recent years, several methods for automatic extraction of translation patterns from bilingual corpora have been proposed as developing NLP technologies (for example, [1, 6, 8]). Maximizing their benefits by taking into account the characteristics of legal documents will enable automatic extraction of the translation patterns for legal documents, along with a reduced development cost. Such extracted translation patterns are useful for both machine translation and bilingual dictionary compilation.

This paper describes a method for automatic extraction of translation patterns from parallel legal corpora. The proposed method extracts the translation patterns by estimating *bunsetsu*-level phrase alignment[1] based on three indexes: existing bilingual dictionaries, stochastic correspondence information and dependency structures.

We have carried out an experiment on Japanese-English translation pattern extraction from the Japanese Civil Code, which is one of the most representative Japanese laws. The result has provided

[1] A *bunsetsu* is one of the linguistic units in Japanese, and roughly corresponds to a basic phrase in English. A bunsetsu consists of one independent word and more than zero ancillary words.

80.5% precision and 49.1% recall on the evaluation with 100 sentences in the Code, suggesting the possibility of support for translating legal documents and compiling a bilingual legal-domain dictionary.

The remainder of this paper is organized as follows: Section 2 explains characteristics of legal documents and the task of translating them. Section 3 describes the indexes for automatic extraction of translation patterns, and Section 4 gives our extraction method. Section 5 reports on the experiment on pattern extraction and provides the results.

2. LEGAL DOCUMENTS AND THE TASK OF TRANSLATING THEM

In this section, we firstly describe the characteristics of legal documents and the issues involved with their translation task. Next, we explain the necessity of a bilingual dictionary for legal translation task and automatic extraction of legal translation patterns.

2.1. Legal document translation

Translating Japanese legal documents is an important task for increasing international understanding of society. Some problems have been experienced with translation of legal documents to date, including unevenness of the translation quality and the limited use of the translated documents in organizations. In recent years, this fact has made it necessary for some parties to share good-quality translations [3]. The unification of translation patterns is essential for sharing the large-scale translated documents[2]. To achieve it, we need bilingual dictionaries whose entries include words that relate to basic laws. It takes a lot of costs to compile a dictionary, so automatic extraction of translation patterns is required to reduce its cost.

2.2. Linguistic characteristics of legal documents

We have investigated the characteristics of legal documents using the Japanese Civil Code, which is one of the most representative and basic laws. For the analysis, we used 2,129 sentences in this Code[3]. Table 1 shows the average length of sentences, the number of different words/bi-grams, and the perplexity for morphemes. The same types of data of 2,129 sentences in Mainichi newspapers(in 1995), which is one of the most famous newspapers in Japan, are shown in the table for comparison. The length of a sentence is defined as the number of morphemes in it. The basic form is adopted for the number of different morphemes/bi-grams.

Table 1: Comparison between legal documents and newspapers

item	Civil Code	newspapers
average sentence length	30.0	22.5
No. of different morphemes	2000	7197
No. of different bi-grams	10460	26951
word perplexity	159.2	575.6

The sentences in legal documents are generally longer than those in newspapers. On the other hand, the number of different morphemes is only 27.8% that in the newspapers. This supposes legal documents are organized by some constant variation of vocabulary, different from that of newspapers that include many proper nouns, and it also supposes that it is useful to compile a bilingual dictionary of words that frequently occur in legal documents. This expectation is encouraged by the fact that the words used for legal documents are settled in meaning and usage [9]. Moreover, the fact that the perplexity in the legal documents is smaller than that in the newspapers is also encouraged by the existence of some characteristic patterns.

3. INDEXES FOR TRANSLATION PATTERN EXTRACTION

There exist many previous researches for translation pattern extraction. In this section, we describe three indexes used in the previous researches for automatic extraction of translation patterns, based on the linguistic characteristics of legal documents: existing bilingual dictionaries, statistical information for calculating the measurement of whether a particular word-pair is aligned, and the dependency structure for use in extraction.

3.1. Existing dictionaries

Some large-scale bilingual dictionaries have been consolidated and published in recent years [5]. Bilingual dictionaries for legal documents have also been constructed and marketed, although they are small-scale and do not cover all legal terms (for example, [11]). When extracting translation patterns from bilingual corpora, word alignment can be inferred by looking up bilingual dictionaries. The proposed method in this paper aims for high precision and recall of the extracted patterns in the corpus by utilizing existing bilingual dictionaries, as seen in some preceding works(for example, [1]).

[2]Since the selection of words in law is made very carefully, it is also required to carefully choose the translation patterns [9].

[3]The Japanese Civil Code is composed of five Books: General Provisions, Real Rights, Obligations, Relatives, and Succession. Though Books I-III are described in literary style in the original, we used documents that are converted to colloquial style.

3.2. Statistic information of corpus

Preceding researches for extraction of translation patterns from parallel corpora often employ methods for statistical inference of the alignment by occurrences of each word in the corpus (for example, [6, 8]). The statistical methods can accurately extract translation patterns that frequently occur in the corpus. Therefore the proposed method may infer the translation patterns that frequently occur in the legal documents but do not exist in bilingual dictionaries. Thus the proposed method infers translation patterns by statistical information as well as by information from existing dictionaries.

In translated legal documents, few free translations and omissions can exist, since it is required that the information on the source language is represented equally for the properties of a legal translation as we saw in section 2. Besides, the fact that the number of different words is small leads us to expect that words in legal documents occur more frequently than those in newspapers. Thus, in the translated legal documents, most of which are small, statistical information would be useful for inference of word alignment, although the corpora which include over 10,000 sentences have been used for the stochastic word alignment methods.

3.3. Syntactic information based on dependency structure

The proposed method utilizes dependency structure in both source and target sentences as the syntactic information for extraction, just like existing dictionaries and statistical information. Both Japanese and English, each of which belongs to a different language family, have similar dependency structures that mean the modification relation between two phrases, thus can be utilized for extraction.

An example is shown in Fig. 1, which consists of the structures of a Japanese legal sentence "Kono-baai-ni-wa, zenjyou-no kitei-ni-yotte tokubetsu-dairi-nin-wo sennin-shinakereba-naranai," which is translated as "In such cases, a special representative shall be appointed in accordance with the provisions of the preceding Article." In the figure, for instance, the bunsetsu-level translation patterns "tokubetsu-dairi-nin-wo"-"a special representative" and "sennin-shinakereba-naranai"- "shall be appointed," exist. Examining each sentence carefully reveals that the Japanese bunsetsu "tokubetsu-dairi-nin-wo" depends on "sennin-shinakereba-naranai," and English phrase "a special representative" also depends on "shall be appointed." Moreover, for the translation patterns "kitei-ni-yotte"-"in accordance with the provisions"[4] and

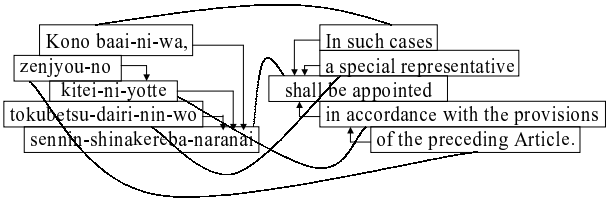


Figure 1: Similarity of dependency structures in Japanese and English

"sennin-shinakereba-naranai"- "shall be appointed," the same is worked out on dependency structure.

Parsing tools with relatively high accuracies enabled by the development of recent statistical parsing techniques have been developed and released [2, 14]. Some previous researches apply these techniques for extracting translation patterns(for example, [1]). The proposed method also utilizes them, and we expect that the tools can parse with reasonable accuracy because of the linguistically faithful structures in legal documents, although legal sentences are generally long and contain particular legal expressions [10].

4. METHOD FOR TRANSLATION PATTERN EXTRACTION

In this section, we propose a method for extracting translation patterns using parallel legal corpora, considering the three indexes described in Section 3.

4.1. Outline of the method

Fig. 2 shows a flow chart of the proposed method, which extracts Japanese bunsetsu-level translation patterns from a Japanese legal document and its translation in which sentences are properly aligned. The process is as follows:

Step 1: Parse both legal sentences in bunsetsus by utilizing dependency parsing tools.

Step 2: Infer word alignment according to dictionaries and Dice coefficient of each word-pair. If the alignment has ambiguities, disambiguate them.

Step 3: Infer bunsetsu-level alignment based on the disambiguated word alignment.

Step 4: Infer the bunsetsu alignment that is not acquired by the word alignment, based on the dependency structure of both languages and the acquired bunsetsu alignment.

[4]The sequence "in accordance with the provisions" consists of two chunks based on the definition in this paper. We take it as one chunk for simplification in Fig. 1.

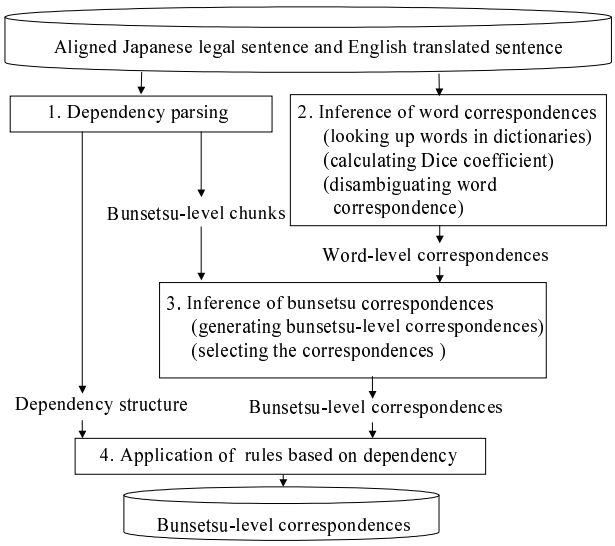


Figure 2: Configuration of the proposed method

We describe these steps in detail in the remainder of this section.

4.2. Dependency parsing with bunsetsu level

The proposed method acquires bunsetsu-level dependency structures using parsing tools. Since the translation pattern extraction is based on the Japanese bunsetsu-level chunks, Japanese sentences are segmented into bunsetsus. English sentences are also adjusted with bunsetsu-like chunks using some rules based on parts-of-speech and base NP, as well as dependency analysis.

4.3. Inference of word correspondences

4.3.1. Inference of word alignment by dictionaries

The proposed method infers word-level alignment, before the inference of bunsetsu-level alignment. At first, the word correspondences inferred as words for translation are generated by looking up the relevant words in bilingual dictionaries. Both the surface and basic forms are adopted for the search, and word sequences bounded to five morphemes/words are also looked up. The word correspondence for a sequence is dealt with as one correspondence. A Japanese collocation "houtei dairi nin," for example, is composed of three morphemes "houtei," "dairi" and "nin," so "houtei dairi," "dairi nin" and "houtei dairi nin" are also targets to look up in addition to each morpheme.

Function words are not targets to look up; the ambiguities of word correspondences are increased because the same function words generally exist in the same sentence.

4.3.2. Inference of word alignment by statistical information

There exist many and sophisticated ways to infer word alignment in bilingual corpus(for example, [4, 12]). For simplification of inference, the extraction method we adopted in this paper treats the Dice coefficient as statistical information, which is one of the simplest ways for deciding whether the word-pairs are aligned. It stands for the ratio of co-occurrences of word-pairs and independent occurrences of each word in the corpus, and is based on the assumption that the higher this value, the higher the possibility that the word-pair has aligned. The Dice coefficient $Dice(w_j, w_e)$ can be calculated as follows:

$$Dice(w_j, w_e) = \frac{2 \cdot f_{je}}{(f_j + f_e)}, \quad (1)$$

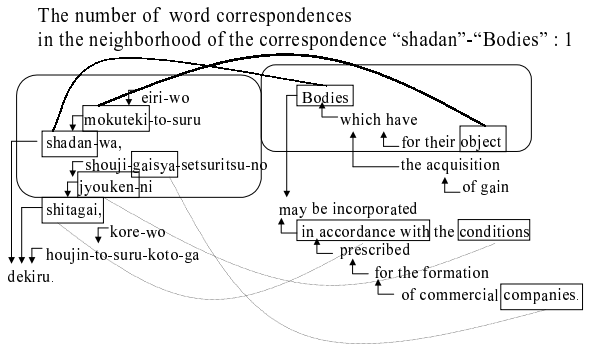
where f_j is the frequency of the Japanese word w_j, and f_e is the frequency of the English word w_e, and f_{je} is the frequency of w_j and w_e that occur on the aligned sentences.

The proposed method acquires word correspondences by inferring that word-pairs with a Dice coefficient above a fixed threshold are aligned.

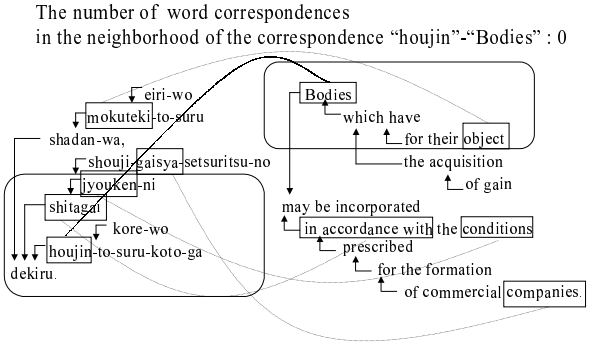
4.3.3. Disambiguation of word correspondences

Word correspondences might have ambiguities in situations where the same word exists more than once in a sentence or some word-pairs for one word have higher Dice coefficient than a fixed threshold. The disambiguation algorithm in the proposed method is based on Aramaki's method [1]. This method supposes that some correspondences exist in the neighborhood of a correspondence where word correspondences are given in the aligned sentences.

The proposed method disambiguates word correspondences using the number of correspondences in the neighborhood of the ambiguous correspondence, based on the above assumption. We define the neighborhood as the following and preceding two bunsetsus of the indicated bunsetsu. The disambiguation method selects the word correspondence that has the most word correspondences in the neighborhood and rejects the others. Fig. 3 shows an example of an ambiguity in two correspondences, "shadan"-"Bodies" and "houjin"-"Bodies". The number of word correspondences in the neighborhood of "shadan"-"Bodies" is one; the only correspondence "mokuteki"-"object," exists as shown in Fig. 3(a). On the other hand, the one in the neighborhood of "houjin"-"Bodies" is zero; there are no word correspondences in the neighborhood, as shown in Fig. 3(b). Then, the word correspondence "shadan"-"Bodies" is selected and "houjin"-"Bodies" is rejected.



(a)



(b)

Figure 3: Disambiguation of word correspondences

4.4. Inference of bunsetsu correspondences

Bunsetsu correspondences are inferred based on the word correspondences inferred by dictionaries and statistical information. The inference includes two steps: (1) generation of translation pattern candidates, (2) selection of the candidates.

4.4.1. Generation of translation pattern candidates

Bunsetsu-level correspondence candidates are generated based on the inferred word correspondences.

Japanese bunsetsus and English bunsetsu-level phrases that include word correspondences are inferred as candidates for bunsetsu-level correspondence. When a bunsetsu has some word correspondences in more than one bunsetsu in the other language, they are also generated as candidates. Fig. 4 shows an example for generating candidates. In the figure, for a Japanese bunsetsu “yon-bun-no-san-ijyou-no” and two English phrases “of three quarters” and “or more,” three word correspondences “yon-bun-no”-“quarter,” “san”-“three” and “ijyou”-“more” are inferred. In such a case, “of three quarters or more” is a candidate of correspondence for “yon-bun-no-san-ijyou-no” as well as “of three quarters” and “or more”.

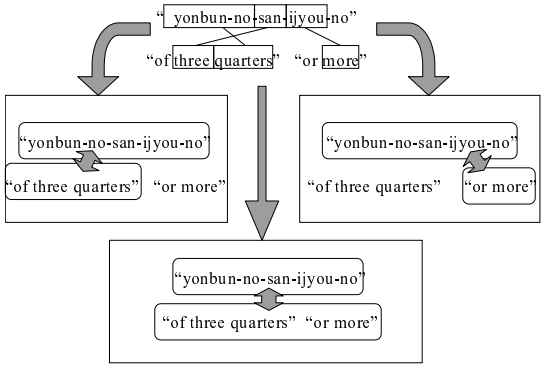


Figure 4: Generation of candidates based on word correspondence

4.4.2. Selection of candidates

Bunsetsu-level correspondences are selected by a measure of the ratio of content words included in word correspondences. The measurement of equivalency E is defined as follows:

$$E = \frac{C(included\ \ word)}{C(Jcontentword) + C(Econtentword)}, \quad (2)$$

where $C(included\ \ word)$ is the number of words included in the word alignment, $C(Jcontentword)$ is the number of Japanese content words in the candidates, and $C(Econtentword)$ is the number of English content words. The method selects the candidate that has the highest measurement, and rejects the other candidates including the selected bunsetsu.

Fig. 5 shows correspondence selection in some candidates. The candidate that has the maximum E is the correspondence “yon-bun no san ijyou no”-“of three quarters or more,” which has three word correspondences, thus $E = (3+3)/(3+3) = 1$, and this correspondence is selected, whereas the others are rejected.

4.5. Correspondence inference by syntactic information

Some bunsetsus that have not yet been inferred as correspondences are extracted as translation patterns, using dependency structure and bunsetsu correspondences inferred by word correspondences. The proposed method uses five heuristic rules for legal sentences to extract new correspondences or combines

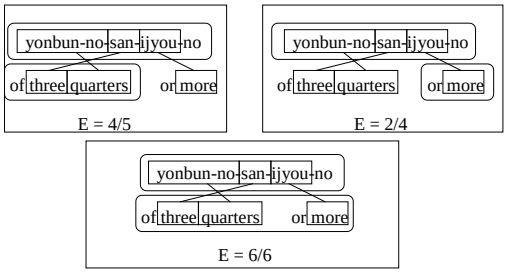


Figure 5: Measurement of equivalency

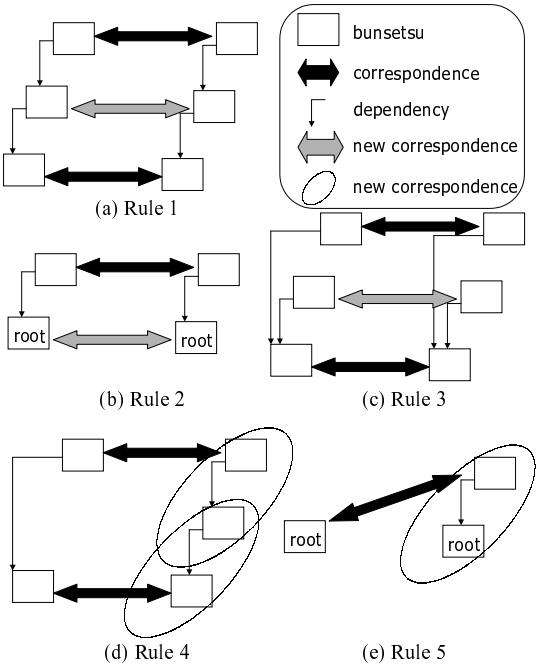


Figure 6: Rules

bunsetsus with existing correspondences. The proposed method aims to extract translation patterns with high precision, using both syntactic information and the already-extracted correspondences.

Rules 1, 2 and 3 extract new correspondences between bunsetsus, whereas rules 4 and 5 combine bunsetsu with existing correspondences. Details of the rules and illustrated figures are as follows:

Rule 1: When dependents of each candidate bunsetsu are inferred as correspondences, and at least one correspondence that depends on each candidate bunsetsu is inferred (Fig. 6(a)).

Rule 2: When both candidate bunsetsus are the roots of sentences and at least one correspondence that depends on each candidate bunsetsu is inferred (Fig. 6(b)).

Rule 3: When dependents of each candidate bunsetsu are inferred as correspondences, and the other bunsetsus that depend on the same bunsetsu as the candidate have already been inferred as correspondences (Fig. 6(c)).

Rule 4: When the dependent of the candidate and a bunsetsu that depends on the candidate are inferred as correspondences, and these correspondences have a dependency relation, combine the candidate with either correspondence. The selection is based on the two measurements of equivalency, as shown in section 4.6.2. (Fig. 6(d)).

Rule 5: When the root bunsetsu in a sentence and a non-root bunsetsu in another are inferred as a correspondence, and the non-head bunsetsu and the root bunsetsu in the other have a dependency relation, combine the non-root with root bunsetsu (Fig. 6(e)).

5. EXPERIMENTS

5.1. Conditions

We performed translation pattern extraction experiments using the Civil Code, which is one of the most representative laws in the system of Japanese law. In these experiments, we utilized 1994 aligned Japanese/English sentences that can be prepared and translated, for the Dice coefficient. The threshold for the Dice coefficient was decided according to a preexperiment and set as 0.6, a score that can maintain the precision of the extracted patterns.

The dictionaries we used for word correspondences were EDR bilingual dictionaries [5], which features approximately 360,000 entries for Japanese-English and about 290,000 entries for English-Japanese, and a bilingual dictionary for business and legal terms by Inter Press [11], which contains about 30,000 entries for both Japanese-English and English-Japanese.

For the Japanese dependency analysis, we used a Japanese morphological analyzer ChaSen [15], a multipurpose chunker Yamcha [13] and a Japanese dependency analyzer Cabocha [14]. Cabocha has a 91% dependency accuracy for newspapers.

For English, after chunking with Japanese bunsetsu-like chunks, we calculate dependency relation between the chunks. These processes have been performed based on the phrase structures of sentences generated by Charniak's parser [2]. Charniak's parser has 90.1% of label precision/recall. The proposed method calculates English dependency structures from phrase structures by determining the head child in each

Table 2: Experimental results

item	precision (%)	recall (%)
proposed method	80.5(43.5)	49.1(26.5)
dictionaries+syntax	79.5(43.4)	48.0(26.1)
dictionaries+statistics	80.7(42.0)	46.3(23.7)
dictionaries	80.2(43.1)	44.9(23.7)
statistics	83.5(52.2)	16.6(10.2)

CFG rewriting rule [7]. For chunking English sentences, we adopt Yamamoto's method[6].

5.2. Results

We evaluated the precision and recall of extracted translation patterns with 100 sentences, which were manually aligned by bunsetsu-level. Table 2 shows the experimental results. In this paper, we define that a correct answer is the one that completely or partly matches with one correspondence that is manually aligned. This definition takes into consideration of manual modification. Parenthesized numbers in the table are the recall and precision based on the complete matches. Precision is defined as the ratio of extracted translation patterns that are included in the correct correspondences, while recall is defined as the ratio of the correct correspondence that matches the extracted translation patterns. The words "dictionaries+syntax/statistics" mean, in this result, that the results of the experiment using existing dictionaries and syntactic/statistical information, but not using statistical/syntactic information, respectively. The word "dictionaries" means that extraction has been done only via looking up dictionaries, and the word "statistics" means that looking up was performed only with the Dice coefficient.

The precision of the proposed method is better than those of other related researches, and the recall for higher than that obtained by other statistical extraction methods [1, 6, 8]. Moreover, the use of statistical and syntactic information caused a 4.2% increase of recall for dictionaries only, retaining its high precision.

In the case that translation patterns are extracted with only statistical information, inference of correspondences can be done with relatively high precision, although its recall is lower. This result suggests that the Dice coefficient has a beneficial effort regardless of the small corpus and its simple inference method. Besides, the statistical information gives an increase of 1.4% in recall in comparison with extraction with only dictionaries.

Using the five rules based on dependency structure, a rise of 2.8% recall has been achieved with similar high precision, in the case that bunsetsu-level alignment is given with about 80% precision and 45% recall.

Table 3: Examples of extracted patterns

Japanese legal term	Its translation
koukyou-no (fukushi-ni)	to the public welfare
kono houritsu-wa	this Code
kaisyaku-shi-nakereba-naranai	shall be construed
hani-nai-de	within the scope

Table 4: Extraction by each dictionary

item	precision (%)	recall (%)
both	80.2	44.9
EDR[5]	78.2	38.8
Inter Press[11]	81.2	36.7

Examples of extracted patterns are shown in Table 3. The proposed method can acquire the patterns that have several bunsetsus/basic phrases, as shown an example "kono houritsu-wa" and "this Code," which has two bunsetsus in the Japanese pattern. Besides, relatively longer translation patterns are acquired although each words exist in dictionaries, as shown in the example "hani-nai-de" and "within the scope." In this pattern, "han'i-nai-de"-"within the scope" doesn't exist in the dictionaries, although "han'i"-"scope" and "nai"-"within" exist as translation words in EDR dictionary[5].

5.3. Relation between two dictionaries

To investigate each dictionary's capabilities, we extracted translation patterns using only one dictionary. Table 4 shows the results.

Each of dictionary has almost the same contribution in this extraction, although EDR dictionaries have higher recall and Inter Press dictionary has higher precision.

5.4. Evaluation for disambiguation algorithm

To evaluate the disambiguation algorithm of word correspondences adopted by the proposed method, we compared this method with random choice in selecting ambiguous alignment. Figure 5 shows the precision and recall, where they use the Dice coefficient and the dictionaries, disambiguation algorithm in each condition.

As shown in the table, 3.3% higher precision and 0.5% higher recall is achieved by utilizing disambigua-

[5]The proposed method infers a word correspondence "nai"-"within" based on looking up "nai," although doesn't looking up "within" because it is a function word.

Table 5: Evaluation of disambiguation method

item	precision(%)	recall(%)
neighborhood:1	80.4	46.8
neighborhood:2	80.7	46.3
neighborhood:3	79.9	46.1
random choice	77.4	45.8

tion algorithm.

6. CONCLUSION

In this paper, we have described a method for automatic extraction of translation patterns from parallel legal corpora. For the indexes of extraction, we utilized bilingual dictionaries, the Dice coefficient in a corpus, and syntactic information based on dependency structure, concerning the characteristics of legal documents and their translation. Using existing bilingual dictionaries and parsing tools, we extracted translation patterns with high recall, while maintaining high precision. By manually modifying the acquired patterns, they become useful for legal translation tasks. The proposed method has utilized the Dice coefficient and the rules applied rather infrequently for dependency structures, thus future work will make more practical use of statistical and syntactic information. We also plan to compile a dictionary for legal translation.

7. ACKNOWLEDGEMENT

The authors are grateful to the reviewers of this paper for their helpful comments. The authors would like to thank Prof. Yoshiharu Matsuura, Associate Prof. Frank Bennett and Associate Prof. Tokuyasu Kakuta for their beneficial instructions and discussions on legal informatics, and also thank Associate Prof. Katsuhiko Toyama and Dr. Yasuhiro Ogawa for their valuable advice. This research is partially aided by the Grant-in-Aid for Young Scientists of the Ministry of Education, Science, Sports, and Culture, Japan and the Nakajima Foundation.

References

[1] E. Aramaki, S. Kurohashi, S. Sato, and H. Watanabe, “Finding Translation Correspondences from Parallel Parsed Corpus for Example-based Translation,” *Proc. of MT Summit-2001*, pp. 27-32, 2001.

[2] E. Charniak, “A Maximum-Entropy-Inspired Parser,” Technical Report CS-99-12, Brown University, 1999.

[3] F. G. Bennett, “Distributed Legal Translation Project,” http://www.nomolog.nagoya-u.ac.jp/˜bennett/

[4] I. Dan Melamed, “Models of Translational Equivalence among Words,” *Computational Linguistics*, Vol. 26, No. 2, pp. 221-249, 2000.

[5] Japan Electronic Dictionary Research Institute, Ltd., “EDR Electronic Dictionary Technical Guide,” http://www.jsa.co.jp/EDR/

[6] K. Yamamoto, and Y. Matsumoto, “Acquisition of Phrase-level Bilingual Correspondence using Dependency Structure,” *Proc. of COLING-2000*, pp. 933-939, 2000.

[7] M. J. Collins, “A New Statistical Parser Based on Bigram Lexical Dependencies,” *Proc. of ACL-96*, pp. 184-191, 1996.

[8] M. Kitamura, and Y. Matsumoto, “Automatic Extraction of Word Sequence Correspondences in Parallel Corpora,” *Proc. of WVLC-96*, pp. 79-87, 1996.

[9] M. Maeda, “Housei-Shitsumu,” Gyousei, 1983 (in Japanese).

[10] M. Ohara, S. Matsubara, and Y. Inagaki, “Legal Sentence Parsing Based on Syntactic Patterns,” *Proc. of IEICE General Conference*, 2003 (In Japanese).

[11] Nichigai Associates. Inc., “INTER PRESS Japanese/English legal dictionary,” 2000.

[12] P. F. Brown, S. A. Della Pietra, V. J. Della Pietra and R. L. Mercer, “The Mathematics of Machine Translation: Parameter Estimation,” *Computational Linguistics*, Vol. 19, No. 2, pp.263-311, 1993.

[13] T. Kudo, and Y. Matsumoto, “Chunking with Support Vector Machines,” *Proc. of NAACL-2001*, pp. 192-199, 2001.

[14] T. Kudo, and Y. Matsumoto, “Japanese Dependency Analysis using Cascaded Chunking,” *Proc. of CoNLL-2002*, pp. 63-69, 2002.

[15] Y. Matsumoto, et al., “Morphological Analysis System ChaSen version 2.2.1 Manual,” http://chasen.aist-nara.ac.jp/