An Example-Based Disambiguation of Prepositional Phrase Attachment

Eiichiro SUMITA, Osamu FURUSE and Hitoshi IIDA

ATR Interpreting Telephony Research Laboratories
2-2, Hikaridai, Seika, Souraku, Kyoto 619-02, JAPAN
e-mail:{sumita, furuse, iida}@itl.atr.co.jp

Abstract

Spoken language translation is a challenging new application that differs from written language translation in several ways, for instance, 1) human intervention (pre-edit or post-edit) should be avoided; 2) a real-time response is desirable for success. Example-based approaches meet these requirements, that is, they realize accurate structural disambiguation and target word selection, and respond quickly.

This paper concentrates on structural disambiguation, particularly English prepositional . phrase attachment (pp-attachment). Usually, a pp-attachment is hard to determine by syntactic analysis alone and many candidates remain. In machine translation, if a pp-attachment is not likely, the translation of the preposition, indeed, the whole translation, is not likely. In order to select the most likely attachment from many candidates, various methods have been proposed. This paper proposes a new method, Example-Based Disambiguation (EBD) of pp-attachment, which 1) collects examples (prepositional phrase-attachment pairs) from a corpus; 2) computes the semantic distance between an input expression and examples; 3) selects the most likely attachment based on the minimum-distance examples. Through experiments contrasting EBD and conventional methods, the authors show the EBD's superiority from the standpoint of success rates.

1 Introduction

Spoken language translation is a challenging new application that differs from written language translation in several ways, for instance, 1) human intervention (pre-edit or post-edit) should be avoided; 2) a real-time response is desirable for success. Example-based approaches meet these requirements, that is, they realize accurate structural disambiguation and target word selection, and respond quickly. This paper concentrates on structural disambiguation, particularly English prepositional phrase attachment (pp-attachment).

One of the most difficult problems in natural language processing is structural disambiguation.' An English prepositional phrase attachment (pp-attachment) is a typical case of structural ambiguity. Usually, a pp-attachment is hard to determine by syntactic analysis alone and many candidates remain. In sentence (1), the pp "at the conference" can attach to either the verb "present" or the noun "paper". Even if a pp can attach to multiple candidates like this, a human can usually select the most likely one. The attachment to the verb "present" is more likely than to the noun "paper".

(1) I will present a paper *at the conference.*

[1] Structural ambiguity causes a combinatorial explosion when a sentence is long. Many papers have discussed eliminating space and time complexity. This paper does not deal with this problem.

In machine translation, it is of great importance to disambiguate a pp-attachment. For instance, while translating sentence (1) into Japanese, the difference between attachments corresponds to the difference between translations (J1) and (J2).[2] (J1) is more likely than (J2).

(J1)	watashi WA	*kaigi***DE**	ronbun WO	happyou-suru
(J2)	watashi WA	*kaigi***NO**	ronbun WO	happyou-suru
	{I}	{conference}	{paper}	{present}

Thus, in machine translation, if a pp-attachment is not likely, the translation of the preposition, indeed, the whole translation, is not likely.[3] Section 2.1 describes English prepositional phrase attachment and its ambiguity in detail. In order to select the most likely attachment from many candidates, various methods have been proposed. Section 2.2 describes conventional methods for disambiguating a pp-attachment.

The authors have been pursuing the example-based approach in machine translation and have achieved high success rates for target word selection. Using the basic technique of the example-based approach, i.e., a retrieval mechanism using semantic distance calculation, this paper proposes Example-Based Disambiguation (EBD) of pp-attachment, which 1) collects examples (prepositional phrase-attachment pairs) from a corpus; 2) computes a semantic distance between an input expression and examples; 3) selects the most likely attachment based on the minimum-distance examples. Section 3 explains the details.

Section 4 shows the EBD's superiority from the standpoint of success rates through contrastive experiments between EBD and conventional methods.

Section 5.1 discusses the applicability of EBD, integration of EBD with conventional NLP techniques, building an example database and the difference between EBD and related research. Section 5.2 discusses semantic granularity, i.e., the relation among words, thesaurus, and conventional semantic markers. Section 5.3 reviews the word-selection accuracy of the example-based approach. Section 5.4 explains the quick response on massively parallel processors. These discussions lead us to the conclusion that example-based approaches meet the major requirements for spoken language translation.

2 Prepositional Phrase Attachment

2.1 English Preposition

Prepositions are the basic devices used in constructing English verb and noun phrases and occur frequently. Prepositions dealt with in this paper, i.e., "of," "to," "for," "in," "on," "at," "from," "by," "with" are the top nine prepositions by frequency in spoken sentences of our domain, *conference registration.* These prepositions also occur frequently in other domains, e.g., technical written language such as computer manuals and general written language such as the *AP news wire.*

English prepositions together with the following nouns modify preceding verbs or nouns. In this paper we call the former *adverbial usage* and the latter *adnominal usage* . For sentence (1), the attachment translated into (J1) is an *adverbial usage* and the attachment

[2] Structural ambiguity is sometimes negligible because translation of the source sentence can be ambiguous as well.

[3] To translate all possible attachments is to make the users pay for the system. A spoken language translation system using such a policy is unlikely.

translated into (J2) is an *adnominal usage.* Many prepositions have two usages and if there are a verb and a noun prior to a preposition, the attachment is ambiguous. The rates of the two usages vary from preposition to preposition (Figure 1). For example, the rate of adnominal usage of "of" and the rates of adverbial usage of "to," "by," and "with" are high (over 90%). Other prepositions are less biased. This paper deals with the above-mentioned nine prepositions as representatives of common prepositions that must be disambiguated structurally.

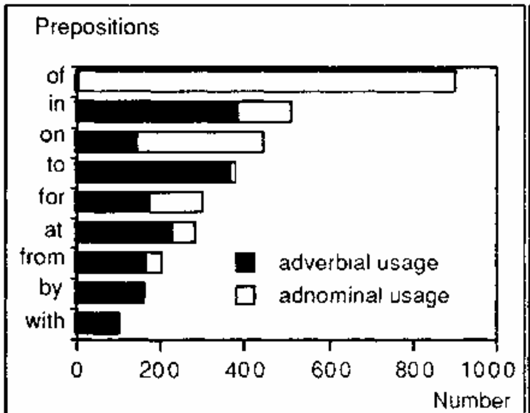

Figure 1
Adverbial/Adnominal Usages

Table 1 Conditions of Experiments

- Domain Conference Registration (E to J)
- Corpus

I. for example
  270K tokens 6.5K types 17K sentences

II. for test (sentences that have at least one ambiguous preposition out of 825 sentences, independent of I)
  1142 tokens 401 types 101 sentences

- Examples 3299 (English parts of EBMT examples extracted from corpus I)
- Test data 131 (ambiguous prepositions and their correct attachments)
- Thesaurus Hierarchy based on LONGMAN LEXICON

2.2 Conventional Methods

This section outlines four groups of conventional methods for disambiguation of pp-attachment. The EBD aims to achieve a high success rate with only a simple computation mechanism like SBD mentioned in section 2.2.3.

2.2.1 Syntactic Methods

Right association which attaches a prepositional phrase immediately to its right and *minimal attachment* which does so in a manner in which the fewest syntactic rules are employed are well known.[1] They are advantageous in that they are simple and general and do not require knowledge other than syntactic knowledge. Recently, however, it is reported that their success rates are not good.[2,3]

2.2.2 Method based on Syntactic Rules and a Dictionary

Many systems specify the preference of pp-attachment in syntactic rules and a dictionary using semantic markers[4]. This method is effective against obligatory cases. However, it is not useful for optional cases exemplified in sentence (1) nor for adnominal usage.

2.2.3 Statistical Methods

Recently, the construction and use of very large corpora has been on the upswing[4] and a variety of statistical NLP research projects have been introduced. Here we outline two.

Method based on Cooccurrence Frequency

Several methods based on cooccurrence have been proposed and have achieved higher success rates than the syntactic methods. Here, we introduce the method proposed by Tsutsumi and Tsutsumi[5]. In this paper, we call the method Statistically-Based Disambiguation

[4] Section 5.2 compares semantic markers, thesaurus, and words.

(SBD). SBD 1) collects 3-tuples, i.e., tuples of verb (or noun) , preposition, and noun [5]; 2) sums up the frequencies of all 3-tuples in the parse tree and stores the totals; 3) selects the parse tree whose total is highest.

Probabilistic Parsing

Recently, probabilistic parsing has been vigorously studied in order to solve structural ambiguity. Probabilistic parsing optimizes application of grammar rules. Compiling many fine-grained, for instance, word-level, grammar rules is indispensable in order to handle pp-attachment, but it is not easy to write such a grammar. Even though such a grammar may be at hand, training time for optimization will be considerable. The effectiveness of probabilistic parsing in pp-attachment is open to question.

2.2.4 Disambiguation based on Deep Understanding

It is generally held that access to world knowledge and a discourse model is necessary to solve structural ambiguity. There is no doubt that deep understanding of sentences will improve the success rate of disambiguation. There were, however, problems with knowledge acquisition and computational cost. From the experimental results of EBD in section 4, we can conclude that a high success rate can be achieved without deep understanding.

3 Example-Based Disambiguation

Here, the authors outline 1) previous studies of example-based approach for translation, 2) one of the basic techniques of the example-based approaches, i.e., semantic distance calculation, 3) the EBD algorithm, and 4) the SBD algorithm mentioned in section 2.2.3.

3.1 Previous Example-Based Approaches

In the early 1980s, Nagao proposed *analogy-based translation[6]* in order to overcome problems inherent in conventional machine translation. Since the late 1980s, several organizations have begun research along this line. First, it was applied to the translation of parts of sentences such as noun phrases including our first model, Example-Based Machine Translation (EBMT) and high quality translation was demonstrated.[7,8,9]. Several methods to translate a whole sentence have been proposed including our second model, Transfer-Driven Machine Translation (TDMT) and are currently under investigation from various points of view.[7, 10, 11, 12, 13, 14, 15, 16, 17] In this paper, semantic distance calculation is applied for structural disambiguation and called EBD.

3.2 Semantic Distance Calculation

Here, we explain one of the basic techniques of EBMT and EBD, i.e., calculation of the semantic distance of two expressions.[8,9] The input, I, and the source part of the example, E, are n-tuples of words, I_k and E_k, respectively. The semantic distance between expressions, d(I,E), is the sum of the semantic distances between words, $d(I_k,E_k)$, multiplied by weights, w_k as shown in formula (1).

$$d(\mathbf{I},\mathbf{E})=\sum_{k}^{n} d(I_k,E_k) * w_k \qquad (1)$$

[5] In fact, they collect cooccurrence data of the form "verb surface-case-marker noun" and convert the data to the above-mentioned 3-tuple for execution.

The semantic distance between words, $d(I_k,E_k)$ is proportional to the location of the concept in the thesaurus.[6] The weight w_k is adjusted to attain high success rates.[7]

3.3 EBD

First, consider a case in which there is only one ambiguous preposition. Suppose the number of attachment candidates is n, the attachment candidate (verb or noun) is xj(1<=i<=n), the preposition is p, and the object of p is y. The problem is to select the most likely attachment "x_k" of prepositional phrase "p y" for the input "x_1, ..., x_n, p, y." For sentence (1), the input is "present, paper, at, conference," the most likely attachment of "at conference" is "present."

[Algorithm]

0) *Iterate the following best match (1<=i<=n) .*
Compute the semantic distance between "x_jp y" and all examples according to formula (1) in section 3.2 and retrieve the minimum-distance examples. Store the minimum-distance, d_j, and the frequency of the same-distance examples, f_j.
1) If there is only one x_j whose d_j is minimum, return the x_j as the most likely attachment and exit.
2) If there is only one x_j whose dj is minimum and whose f_j is maximum, return the x_j as the most likely attachment and exit.
3) Return the set of x_j whose d_j is minimum, signal that more than one candidate remain and exit.

Case 1 - exits at step 1)

For the pp "at conference" of sentence (1), the minimum-distance of the first candidate, "present" is 0.00, the minimum-distance of the second candidate, "paper" is 0.33.[8] The algorithm selects the first as the most likely attachment.

(1) I will present a paper *at the conference.*

Case 2 - exits at step 2)

For the pp "in 1992" of sentence (2), the distances of the two candidates "have" and "conference" are the same, 0.17. However, the algorithm selects "have" because "have" is more frequent than "conference."

---

6 The hierarchy of our thesaurus is three-layered in accordance with the LONGMAN LEXICON[18].

7 In EBMT, the weight is determined based on the k-th word's influence on the translation. In this EBD experiment, the weight is set to 1/n.

8 The candidate most similar to "present at conference" is "present at conference", thus the distance is 0.00. The candidate most similar to "paper at conference" is "speech at conference" and the distance between "paper" and "speech" is 2/3 according to the thesaurus, thus formula (1) gives the distance, 0.33.

(2) We will have the next conference *in 1992.*

Except for step 2), the above-mentioned algorithm is the same one by which the TDMT pattern-matcher disambiguates structural ambiguity using semantic distance.[10] Section 4.2 will show that step 2) which is executed when the distances are equal (we call this step a tiebreaker, hereafter) outperforms the method based only on distance.

We have tested two variations: 1) use 3-tuples; 2) use 2-tuples "x p" (we call them, EBD[xpy] and EBD[xp], respectively, hereafter).

Next, let's move on to a case where there is more than one ambiguous preposition. Just as TDMT selects the combination whose total distance is minimum, the system should select the combination based on distance and frequency of minimum-distance examples. There is an expedient method that selects the combination whose total cost[9] is minimum.

3.4 SBD

The authors implemented SBD by substituting the following exact match for the best match, i.e., step 0) of the algorithm in section 3.3.

---

0) Iterate the following exact match (1<=i<=n) .
Retrieve examples that are the same as "x_ip y". Store the distance, d_i (if found to be 0, otherwise, ∞) and the frequency of the same-distance examples, f_i.

We have tested two variations: (1) use 3-tuples; (2) use 2-tuples "x p" (we call them, SBD[xpy] and SBD[xp], respectively, hereafter). Moreover, for comparison, we have tested a method that determines attachment based only on the adverbial/adnominal rate (we call it the DEFAULT, hereafter).

4 Experimental Results

This section explains the conditions in our experiment, the comparison of the above-mentioned methods, and failures.

4.1 Experimental Conditions

This research is conducted to establish a method to translate spoken language as a part of an interpreting telephony project. Our domain is *conference registration.* As shown in Table 1, the authors make examples and test data from bilingual corpus [19]: 1) The source sides of 3,299 bilingual examples for EBMT are used as examples for EBD; 2) Correct attachments of test data were judged in advance by humans (see appendix). These attachments were used as the standard for comparison.

The algorithms explained in sections 3.3 and 3.4 do not necessarily return a unique

---

[9] The cost is calculated from the semantic distance, d_i, and the frequency of the minimum-distance examples, f_i, according to the next formula $d_i - \frac{f_i}{10^m}$ (2). Here, m depends on the size of the example database.

attachment. The authors use the following rates for analysis of experimental results:

decision-rate(A)

= the number of unique attachments divided by the total number of prepositions;

correct-decision-rate (B)

= the number of correct attachments divided by the number of unique attachments;

correct-rate(C)

= the number of correct attachments divided by the total number of prepositions

= A multiplied by B.

4.2 Effect of Tiebreaking

Figure 2 shows the effect of tiebreaking using frequency, i.e., step 2) of the algorithm presented in section 3.3.

1) Tiebreaking has no adverse side-effects (e.g., converting a correct attachment into an incorrect one.) That is, it does not lower correct-rate(C).

2) It raises decision-rate(A) and has hardly any effect on correct-decision-rate(B). Thus, it lifts correct-rate(C).[10] In the following sections, the authors explain methods using this tiebreaking technique.

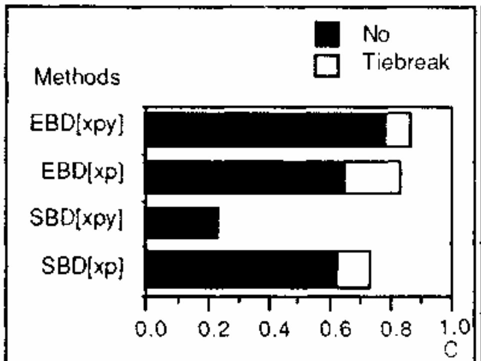

Figure 2
Improvement by Tiebreak

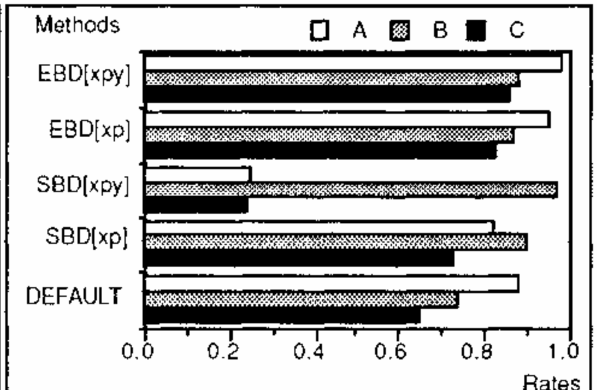

Figure 3
Comparison by Rates: A,B, and C

4.3 Comparison SBD and EBD

4.3.1 Correct-rate(C), Decision-rate(A), Correct-decision-rate(B)

Correct-rate(C) is the product of decision-rate(A) and correct-decision-rate(B) (Figure 3). For instance, correct-rate(C) of SBD[xpy] is lower than those of the EBDs, because the low decision-rate(A) cancels high correct-decision-rate(B). SBD[xp] has a lower decision-rate(A) than do the EBDs and correct-decision-rates(B) are almost the same. Thus, SBD[xp] is lower in correct-rate(C) than the EBDs. Consequently, correct-rates(C) of the SBDs are lower than those of the EBDs.

SBD[xp] is superior to SBD[xpy], partly because the smaller the example unit, the more examples we get, partly because "x p" influences the pp-attachment. This does not directly conclude that the case frame is enough for disambiguation partly because it is difficult to deal with optional cases like "at the conference" in sentence (1) by case frames and partly because

[10] There are too few examples to improve the SBD[xpy] using the tiebreaking technique.

the typical semantic marker is not sufficient for disambiguation (see section 5.2). DEFAULT'S high decision-rate(A) is canceled by its low correct-decision-rate(B). SBD[xp] is better than DEFAULT because "x p" has more information than DEFAULT and, in this experiment, there are enough examples for SBD[xp].

4.3.2 Correct-rate(C) vs. The Example Database Size

We divided the example database into quarters. By incrementing the example database from one quarter to four quarters, the relationship of the example database size and correct-rate(C) is investigated (Figure 4).

1) It is true in EBDs and SBDs that correct-rate(C) increases with the number of examples. The rate of increase of the SBD[xp] is maximum.

2) Within the example database of the experiment[11], EBD's correct-rate(C) is always higher than SBD's. The minimum EBD correct-rate(C) is 0.76 at EBD[xp] with a quarter of the example database. The maximum SBD correct-rate(C) is 0.73 at SBD[xp] with the complete example database. EBDs have accomplished higher correct-rates(C) than SBDs with only a one-quarter database. The larger the differences in correct-rates(C) are, the smaller the size of the example database is.

From these observations, we can conclude the following: 1) compared with SBDs, EBDs achieve high correct-rates (C) with only a small example database; 2) EBDs stand up to unseen data.

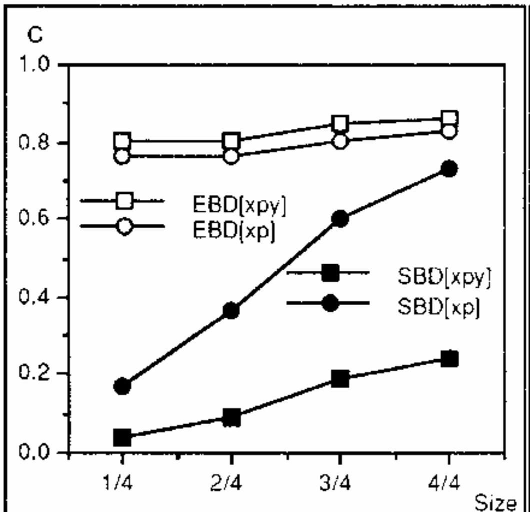

Figure 4 Success vs. DB size

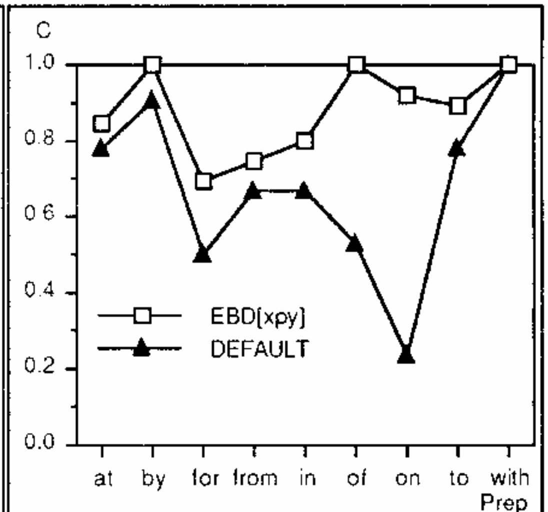

Figure 5 Comparison by Preposition

4.3.3 Observations by Prepositions

Observing differences in EBD[xpy]'s correct-rate(C) by preposition (Figure 5), we can see that the average correct-rate(C) of prepositions whose adverbial/adnominal rate is biased, i.e., "of," "to," "by," "with" is 0.96 and the average correct-rate(C) of prepositions whose adverbial/adnominal rate is less biased, i.e., "for," "in," "on," "at," "from" is 0.81. The result reflects somewhat the bias in the adverbial/adnominal rate (Figure 1).

However, correct-rate(C) of EBD[xpy], 0.86, is much higher than correct-rate(C) of DEFAULT, 0.65. Failures of DEFAULT are as follows: 1) e.g., "of." When the candidates are all

[11] If the size of the example database continues to increase, the correct-rates (C) will converge.

adnominal, DEFAULT cannot differentiate them; 2) e.g., "on." The difference in the global preference ("p") and the specific preference ("x p y") is large. There are many noun phrases of the form "conference on something," i.e., *adnominal usage* in this corpus. However, the test set has more adverbial usages than adnominal usages.

4.4 Failures and Countermeasures

This section examines the failures of the most accurate method, EBD[xpy] and presents countermeasures[12].

Lack of Similar Examples

In sentence (3), the most similar example for the correct attachment "facility at conference" was "speaker at conference." The semantic distance is relatively low, 0.5.[13] These failures can be eliminated by adding appropriate examples.[14]

(3) I want to ask about facilities *at the conference.*

Questions in Example Unit

EBD[xpy] uses 3-tuples,"x p y" as an example unit, "x" (and "y") represents the head of a noun phrase or verb phrase. The example unit cannot capture the influence of a verb's case frame, prepositions other than the prepositions of the example unit, nor a modifier of heads in the example unit. As exemplified in sentence (4), it is likely that "from" and "to" attach to the same candidate. There are two countermeasures: 1) to use a larger example unit; 2) to integrate EBD with a conventional method based on rules and a dictionary.

(4) There are busses running *from Kyoto to Kitaoji.*

5 Discussions

Section 5.1 mentions the applicability of EBD, integration of EBD with conventional methods, building an example database and the difference between EBD and related research. Section 5.2 compares semantic distance calculation based on thesaurus and other conventional semantic handling. Section 5.3 reviews accurate word selection of EBMT. Section 5.4 explains the quick response of the example-based approach on massively parallel processors.

5.1 Prospects for and Related Research on EBD

EBD, which is based on semantic distance and frequency, is not a specialized method for English pp-attachment. It is effective for other structural ambiguity caused by, for instance, infinitives, relative pronouns, subordinate conjunctions and so on.[15] The authors have already implemented structural disambiguation[16] based on semantic distance in TDMT, which translates

[12] The causes of failures are intricate. Here, the authors determine countermeasures under the principle that any change in the thesaurus and semantic distance definition is to be avoided.

[13] We have shown that, in general, the smaller the distance, the better the quality[8].

[14] We have shown that, in general, the more examples we have, the better the quality[8].

[15] In the case of ambiguous scope caused by coordinate conjunctions, it is necessary to integrate EBD with a parsing method like the one proposed by Kurohashi[22] which is based on dynamic programming.

[16] In TDMT, tiebreaking has not been implemented.

Japanese spoken sentence into English. The experimental results demonstrated that EBD in TDMT can properly deal with many kind of structural ambiguity in Japanese.

EBD is easily incorporated into conventional parsers by using EBD as a subroutine. 1) During parsing, by augmenting rules for pp-attachment, or, 2) After parsing, by traversing parse trees, 3-tuples, "x p y", are passed to the EBD subroutine. This integration can utilize the parser's global constraints, such as *no cross modification,* and *lexical constraint,* such as the obligatory case of the verb's case frame, thus improving correct-rate(C).

In our experiments, the examples are collected by hand. Kaji *et al.* [20] proposed a method that automates collection of bilingual examples from a corpus. The authors consider that automation of example collection is indispensable for large-scale implementation and plan to devise a method in the near future.

In order to disambiguate pp-attachment, Jensen *et al*.[21] proposed a method that parses definition sentences in a dictionary and extracts necessary information by pattern-matching against the parse trees. Their method requires many rules that extract semantic relations and reliability factors from the parse tree pattern for each preposition, relation by relation. However, the EBD mechanism can deal with every preposition in a uniform way.

EBD attaches importance to the individuality of linguistic phenomena. For the time being, it does not aim to extract any abstract knowledge from examples, but to make use of examples directly. To generalize examples[23] or to compress the example database with no drop in system performance is of importance from the standpoint of space and time complexity.

5.2 Semantic Granularity

As explained below, thesauri give appropriate semantic granularity for problems of natural language processing.

In the pp-attachment problem, compared with SBD which utilizes the frequency of word cooccurrence (3-tuples of words), EBD which utilizes examples (3-tuples of words) and a thesaurus can achieve higher correct-rate(C) with far fewer examples (a much smaller corpus) as shown in section 4.3.2. This implies that best *match* based on the thesaurus compensates for the notorious sparse data problem.

In the case frame selection problem, Nagao [24] compared an example-based method that utilizes example sentences and a thesaurus with a conventional method that utilizes semantic markers that consist of several categories. The former is significantly better in the selection of case frame than the latter.

5.3 Accurate Word Selection (EBMT)

As shown in our previous papers[8,9], EBMT has achieved high success rates, about 80-90% on average, for target word selection. Frequent and polysemous linguistic phenomena: 1) content words such as Japanese verbs, 2) function words such as Japanese case particles and English prepositions, were tested. EBMT can achieve accurate word selection not only for content words but also function words. EBMT is not language specific because we have tested at least two very different languages.

5.4 Distance Calculation on Massively-Parallel Processors

Semantic distance calculation is performed against every example in the database. Thus, the calculation with a vast example database is computationally demanding. Each calculation, however, can be done independently of other calculations. The response time of semantic distance calculation can be lessened drastically by massively parallel processors. When each example is assigned to a Processing Element (PE) of a massively parallel processor, each PE's computation cost and the communication cost between PEs, are very small. Unlike many natural language processing techniques (e.g., unification) semantic distance calculation used in EBD, EBMT and TDMT is the best fit for massively parallel processors. Experimental results[25] showed that semantic distance calculation on a massively parallel associative processor, IXM2, exhibited the best performance and attained a response speed that would suffice for real-time spoken language translation such as interpreting telephony.

6 Conclusion

This paper has proposed EBD, an example-based method to disambiguate pp-attachment, and demonstrated EBD's superiority from the standpoint of success rate, compared with conventional methods. In addition, our previous papers have shown that EBMT achieved accurate word selection and that semantic distance calculation on massively parallel processors attained a response speed that would suffice for real-time translation. Thus, example-based approaches meet major requirements for spoken language translation

Future work includes realization of spoken language translation using example-based approaches and automation of example collection.

References

[1]Frazier, L. & Fodor, J.: "The Sausage Machine. A New Two-Stage Parsing Model," Cognition 6, pp.191-325, (1979).

[2]Whittemore, G, Ferrara, K. and Brunner, H.: "Empirical study of predictive powers of simple attachment schemes for post-modifier prepositional phrases," Proc. of 28th ACL, pp.23-30, (1990).

[3]Hindle, D. and Rooth, Mats: "Structural ambiguity and lexical relations," Proc. of 30th ACL, pp.229-2236, (June 1991).

[4]Walker, D.: "The Ecology of Language," Proc. of International Workshop on Electronic Dictionaries, Oiso, Japan, pp.10-22, (Nov. 1990).

[5]Tsutsumi, Y. and Tsutsumi, T.:"Disambiguation Method Based on the Stochastic Data for Natural Language Parsing," Trans. IEICE, J72-D-IL9, pp.1448-1458, (1989)(in Japanese).

[6]Nagao, M.: "A Framework of a Mechanical Translation between Japanese and English by Analogy Principle," in Artificial and Human Intelligence, eds. A. Elithorn and R. Banerji, North-Holland, pp.173-180, (1984).

[7]Sato, S.: "Example-Based Machine Translation," Doctoral Thesis, Kyoto University, (1991).

[8]Sumita, E., and Iida, H.: "Example-Based Transfer of Japanese Adnominal Particles into English," IEICE TRANS. INF. & SYST, VOL. E75-D, No. 4, pp.585-594, (1992).

[9]Sumita, E., and Iida, H.: "Example-Based NLP Techniques - A Case Study of Machine Translation - Statistically-Based Natural Language Processing Techniques-Papers from the 1992 Workshop, Technical Report W-92-01, AAAI Press, (1992).

[10]Furuse, 0. and Iida, H.: "Cooperation Between Transfer and Analysis in Example-Based Framework," Proc. of Coling '92, pp.645-651, (1992).

[11]Watanabe, H.: "A Similarity-Driven Transfer System," Proc. of Coling'92, Nantes, vol. 2, pp.770-776, (1992).

[12]Sadler, V.: "Working with Analogical Semantics," Foris Publications, (1989).
[13]Kitano, H." DMDIALOG, An Experimental Speech-to-Speech Dialog Translation System," IEEE Computer, June, pp.36-50, (1991).
[14]Kitano, H. and Higuchi, T.:"Massively Parallel Memory-Based Parsing," Proc. of IJCAI'91, pp.918-924, (1991).
[15]Maruyama, H. and Watanabe, H.: "Tree Cover Search Algorithm tor Example-Based Translation," Proc. of TMI-MT'92, pp.173-184, (1992).
[16]Jones, D.: "Non-Hybrid Example-Based Machine Translation Architectures," Proc. of TMI-MT'92, pp.163-171, (1992).
[17]McLEAN, I. J.: "Example-Based Machine Translation Using Connectionist Matching," Proc. of TMI-MT'92, pp.35-43, (1992).
[18]McArthur, T.:"Longman Lexicon of Contemporary English," London:Longman Group, (1981).
[19]Ehara, T., Ogura, K., Shinozaki, N.. Morimoto, T, Kurematsu, A:"Construction of a Dialogue Database (ADD) from Dialogues through Telephone or Keyboard Channels," Trans. of IPSJ.Vol.33,No.4,pp.448-456(1992)(in Japanese).
[20]Kaji, H., Kida ,Y. and Morimoto, Y: "Learning Translation Templates from Bilingual Text," Proc. of Coling'92, Nantes, pp.672-678, (1992).
[21]Jensen, K. and Binot, J.:"Disambiguating Prepositional Phrase Attachments by Using On-Line Dictionary Definitions," Computational Linguistics, Vol.13, No.3-4, pp.251-260, (1987).
[22]Kurohashi, S. and Nagao, M.:"Dynamic Programming Method for Analyzing Conjunctive Structures in Japanese," Proc. of Coling'92, Nantes, pp.170-176, (1992).
[23]Nomiyama, H.: "Machine Translation by Case Generalization," Proc. of Coling'92, Nantes, pp.714-720, (1992).
[24]Nagao, M.: "Some Rationales and Methodologies for Example-Based Approach," Proc. of FGNLP'92, pp.82-94, (1992).
[25]Sumita, E., Oi, K., Furuse, O., Iida, H., Higuchi, T, Takahashi, N., and Kitano, H.: "Example-Based Machine Translation on Massively Parallel Processors," Proc. of IJCAI'93, (to appear), (1993).

Appendix

Judging attachment is not simple and is a disputable problem. In this experiment, we have determined the following standard. Prepositional phrases are shown in italic, attachment candidates are underlined, and correct attachments are shown in boldface.

- In expressions like light verb construction, a pp attaches to the verb.

make arrangement	*at hotel*
take the subway	*to Kyoto station*
give our best regards	*to Prof. X*

- If a pp is a recipient, a locative, etc., different attachments are interpreted as having almost the same meaning. In this case, multiple attachments are accepted as correct (about 10%).

arrange limousine	*for you*
is a change	*in your plans*
tell me **the details**	*on the conference*