

# Translating with Examples: A New Approach to Machine Translation

Eiichiro SUMITA, Hitoshi IIDA, and Hideo KOHYAMA

ATR Interpreting Telephony Research Laboratories  
Sanpeidani, Inuidani  
Seika-cho, Soraku-gun  
Kyoto 619-02, Japan  
e-mail: sumita%atr-la.atr.co.jp@uunet.uu.net

## Abstract

This paper proposes **Example-Based Machine Translation (EBMT)**. **EBMT** retrieves similar examples (pairs of source texts and their translations) from the database, adapting the examples to translate a new source text.

This paper compares the various costs of **EBMT** and conventional **Rule-Based Machine Translation (RBMT)**. It explains **EBMT**'s new features which **RBMT** lacks, and describes its configuration and the basic computing mechanism. In order to demonstrate **EBMT**'s feasibility, the translation of Japanese noun phrases of the form "N<sub>1</sub> no N<sub>2</sub>" to English noun phrases is explained in detail. Translation of other parts of Japanese sentences, including "da" sentences, aspect, and idiomatic expressions, as well as the integration of **EBMT** with **RBMT** are discussed.

## 1 Introduction

Machine Translation requires handcrafted and complicated knowledge, for example, dictionaries, grammar rules, and rewriting rules<sup>[1]</sup>. It is difficult to scale up from a toy program to a practical system because it is difficult to build the above-mentioned knowledge. It is also difficult to improve translation performance because the effect of adding a new rule is difficult to anticipate and translation using a large-scale rule-based system is time-consuming. In order to conquer these problems in machine translation, it is necessary to study the form of the knowledge, and a translation mechanism for that knowledge.

The use of a database of **examples**(pairs of source texts and their translations) as knowledge has been instituted. The translation mechanism retrieves similar examples from the database, adapting the examples to translate the new source text. This framework is called **Example-Based Machine Translation(EBMT)**. In order to show **EBMT**'s feasibility, a system to translate Japanese noun phrases of the form "N<sub>1</sub> no N<sub>2</sub>" to English noun phrases has been implemented.

This paradigm was initiated by Nagao<sup>[2]</sup> who was the first to emphasize the importance of examples and a thesaurus. He called the method "Translation by Analogy". This is similar to what a human does when he translates using dictionary examples. Sato and Nagao<sup>[3]</sup> developed an experimental system capable of translating simple sentences using this paradigm. Sadler presents a translation simulation with a database of bilingual dependency trees<sup>[4]</sup>. Brown et al. also propose an approach to

machine translation which utilizes statistical techniques of information extraction from a database of bilingual texts <sup>[5]</sup>.

Recent progress in hardware such as vast memory and the computing power of parallel computers makes **EBMT** a practical research goal.

Section 2 compares **EBMT** and **Rule-Based Machine Translation(RBMT)**, section 3 explains **EBMT**'s mechanism, section 4 illustrates the experiment in detail, and section 5 discusses translation of other parts of Japanese sentences as well as the integration of **EBMT** with **RBMT**.

## 2 EBMT and RBMT

This chapter compares the various costs of **EBMT** and **RBMT** and explains **EBMT**'s new features which **RBMT** lacks.

### 2.1 Computational Cost

Computational cost is considerable in **RBMT**. **RBMT** is really a large-scale rule-based system, which consists of analysis, transfer, and generation modules using syntactic rules, semantic restrictions, structural transfer rules, word selections, generation rules, and so on. For example, the Mu system has about 2,000 rewriting rules and word selection rules for about 70,000 lexical items <sup>[6]</sup>. **EBMT** directly returns a translation by adapting the examples without reasoning through a long chain of rules. The computational cost of **EBMT** is less than that of **RBMT**.

### 2.2 Improvement Cost

In **RBMT**, it is too difficult to keep all rules consistent because improvement of translation quality is made by modifying rules that are mutually dependent. **EBMT** has no rules, thus improvement is effected simply by inputting appropriate examples to the database. In other words, **EBMT** is easily upgraded but **RBMT** is not.

### 2.3 System Building Cost

Formulating linguistic rules for **RBMT** is a difficult job and needs a linguistically trained staff. However examples necessary for **EBMT** are easy to obtain because a large number of texts and their translations are available. Moreover, as electronic publishing increases, more and more texts will be machine-readable.

### 2.4 Context-Sensitive Translation

**RBMT** needs another understanding device in order to translate context-sensitively. Because **EBMT** is a general architecture, incorporating contextual information into example representation provides a way to translate context-sensitively. As for our corpus, i.e., conversation about registering for an international conference, the set of words surrounding examples, the speaker of the examples, and so on, are ready to be used.

### 2.5 Robustness

**RBMT** works on exact-match reasoning, **EBMT** on best-match reasoning. **RBMT** fails to translate when it has no knowledge that matches the input exactly. **EBMT** intrinsically translates in a fail-safe way. If **RBMT** included a fail-safe mechanism to search rules which can translate an expression similar to the input, **RBMT** could then translate by borrowing the found rules.

### 2.6 Reliability factor

In **EBMT**, a reliability factor is assigned to the translation result according to the distance between the input and found similar examples. **RBMT** has no device to compute the reliability of the result. In other words, **EBMT** can tell when its translation is inappropriate.

### 2.7 Example Independency

**EBMT** knowledge consists not of rules based on a particular system as in **RBMT** but rather the linguistic facts themselves. As suggested in Nagao's paper <sup>[2]</sup>, this implies that the knowledge is

completely independent of the system, is usable in other systems and can be analyzed by any linguistic theory.

### 3 EBMT Mechanism

In this section, the **EBMT** system configuration and distance calculation, which are general and applicable to many aspects in machine translation, are shown.

#### 3.1 Configuration

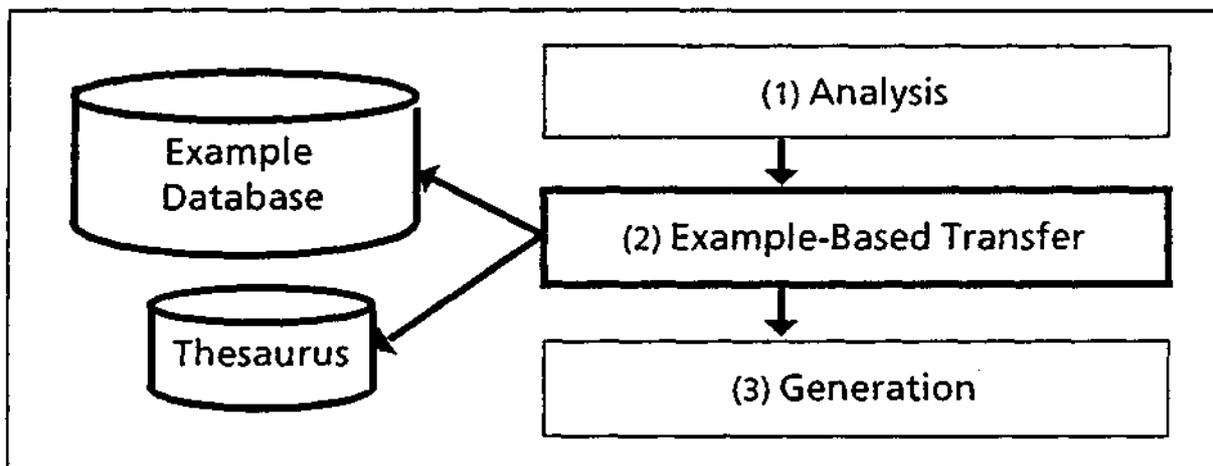


Figure 1 System configuration

The configuration of the **EBMT** system, shown in Figure 1, consists of two databases: an example database and a thesaurus, and three translation modules: analysis, example-based transfer, and generation.

**Example database:** Examples are extracted from a bilingual textbase.

**Thesaurus:** A thesaurus is used to calculate the semantic distance between the content words in the input and those in examples as shown in section 3.2.2.

**Translation:** The following three steps outline the **EBMT** procedure. Step (2) is essential and is explained in detail in section 3.2.

- (1) Conventional Analysis
- (2) Retrieval of examples from the database by distance calculation and Transfer
- (3) Conventional Generation

#### 3.2 Retrieval by Distance calculation

In our paradigm, retrieving the best-match examples to the input is done by measuring the similarity of the input and examples. The essential point is the appropriate definition of distance between the input and examples.

##### 3.2.1 Example Distance

Here we suppose the input and examples in the database are represented in the same data structure, the list of the attribute's values. We refer to them and their *i*-th attribute as **I**, **E** and **I<sub>i</sub>**, **E<sub>i</sub>**, respectively. This is guaranteed by analyzing examples when the database is constructed, using the same analysis program which is used for the input. The total distance between **I** and **E**, **d(I,E)**, is the summation of the distance at each attribute, **d(I<sub>i</sub>,E<sub>i</sub>)**, multiplied by the weight of the attribute, **w<sub>i</sub>**.

$$d(I,E)=\sum d(I_i,E_i) \times w_i$$

### 3.2.2 Attribute Distance

For all attributes other than semantic attributes, the distance is 0 or 1 depending on whether or not they match exactly. For semantic attributes, however, the match is partial and the distance varies between 0 and 1. Semantic distance  $d(0 \leq d \leq 1)$  is determined by the **Most Specific Common Abstraction(MSCA)** [7] from the thesaurus abstraction hierarchy. For example, when the thesaurus is  $(n + 1)$  layered,  $(k/n)$  is assigned to the concepts in the  $k$ -th layer from the leaf layer.

### 3.2.3 Weight of attributes

The weight of the attribute is the degree to which the attribute influences the selection of the translation pattern. We adopt the next expression used by Stanfill for memory-based reasoning [8], to implement the intuition.

$$w_i = \sqrt{\sum (\text{frequency of translation pattern } k \text{ when } E_i = I_i)^2}$$

The calculation of weights is an expensive operation. If the operation is repeated for each calculation of example distance, the total cost is high. Fortunately, the weight depends only on the static frequency in the example database, and this can be determined for each of the attribute's values when the system is built rather than during running time.

## 3.3 Acceleration of retrieval

**EBMT** searches the whole database for best-match examples. Naive implementation will inevitably be slow. There are two solutions to this problem, indexing and parallel computing. As shown in Sumita and Tsutsumi [9], indexing can accelerate the retrieval of syntactically similar sentences. In that system, indexing from the attribute's important values to examples was used, and has proved effective in restricting the search space. Parallel computing is under investigation. This has bright prospects for the future, because the retrieving examples can be divided into independent processes.

In contrast with **RBMT**, **EBMT** can naturally incorporate a bypass which first searches for an exact match. When the input exactly matches examples in the database, translation is quicker because retrieval of the best-match examples is then suppressed.

## 4 Translation of "N<sub>1</sub> no N<sub>2</sub>"

Section 4 describes the importance and difficulty of the problem addressed in the experiment, the configuration, the actual distance calculation, and the results of the experiment.

### 4.1 Difficulty of the problem

Roughly speaking, Japanese noun phrases of the form "N<sub>1</sub> no N<sub>2</sub>" correspond to English noun phrases of the form "N<sub>2</sub> of N<sub>1</sub>". However, "N<sub>2</sub> of N<sub>1</sub>" does not always provide a natural translation as shown in Figure 2. Some translations are too broad in meaning to interpret, others are almost ungrammatical. For example, the third one, "the conference of Kyoto", could be misconstrued as "the conference about Kyoto", and the last one, "hotels of three", is not English. Natural translations often require prepositions other than "of", or no preposition at all. "N<sub>1</sub> no N<sub>2</sub>" is a frequent expression which appears in nearly 50% of all Japanese sentences. In 20-40% of these "N<sub>1</sub> no N<sub>2</sub>" occurrences, "N<sub>2</sub> of N<sub>1</sub>" would be the most appropriate English translation. Selecting the translation for "N<sub>1</sub> no N<sub>2</sub>" is still an important and difficult problem in J-E translation.

In contrast with the preceding research [10], deep semantic analysis of Japanese text is avoided because it is assumed that most translations can be done without deep understanding. As mentioned in section 2.1, **EBMT** directly returns a translation by adapting the examples without reasoning through a long chain of rules. The computational cost of **EBMT** is less than that of **RBMT**.

We can find similar phenomena in other language pairs, for example, Spanish to English. The Spanish preposition "de", with its broad usage like Japanese "no", is also effectively translated by **EBMT**.

### 4.2 Configuration

Japanese	English	translation pattern
八日 廿の 午後 †[8th,afternoon]	the afternoon of the 8th	B of A
会議の 参加料 [conference,application fee]	the application fee for the conference ? the application fee of the conference	B for A
京都 ††での 会議 [Kyoto,conference]	the conference in Kyoto ? the conference of Kyoto	B in A
一週間の 休暇 [a week,holiday]	a week's holiday ? the holiday of a week	A's B
ホテルの 予約 [hotel,reservation]	the hotel reservation ? the reservation of the hotel	A B
三つの ホテル [three,hotel]	three hotels ? hotels of three	A B

Figure 2 Variations in translations of "N<sub>1</sub> no N<sub>2</sub>"

†{N<sub>1</sub>,N<sub>2</sub>} is the list of English words for N<sub>1</sub> and N<sub>2</sub> in "N<sub>1</sub> no N<sub>2</sub>".

††"の" and †††"での" are variants of the Japanese adnominal particle "no".

**Example database:** ATR's linguistic database of spoken Japanese with English Translations is used as the bilingual textbase. The corpus is conversation about registering for an international conference. The size of the database is currently about 100,000 words and is expected to reach about 1,000,000 words<sup>[11]</sup>.

**Thesaurus:** The hierarchy of the thesaurus used is in accordance with the thesaurus of everyday Japanese written by Ohno and Hamanishi<sup>[12]</sup> who classify about 60,000 words decimally and assign a 3-digit semantic code for each class.

**Translation:** Figure 3 illustrates the translation procedure with an actual sample. The distance used when retrieving examples is essential and is explained in detail in section 4.3.

### 4.3 Distance calculation

As explained in section 3.2, distances are calculated using the following two expressions:

$$(1) d(I, E) = \sum d(I_i, E_i) \times w_i$$

$$(2) w_i = \sqrt{\sum (\text{frequency of translation pattern } k \text{ when } E_i = I_i)^2}$$

In this section, an actual distance calculation for "N<sub>1</sub> no N<sub>2</sub>" is illustrated. The attributes of the current target, "N<sub>1</sub> no N<sub>2</sub>", are as follows: for nouns, "N<sub>1</sub>" and "N<sub>2</sub>", lexical subcategory of noun, the existence of prefix or suffix, and its semantic code in the thesaurus; for the adnominal particle "no", the kinds of variants, "の"(no), "での"(deno), "からの"(karano), "までの"(madeno) and so on. Here, for simplicity, only the semantic code and the kind of adnominal are considered.

#### (a) Attribute Distance

For the attribute of the adnominal particle "no", the distance is 0 or 1 depending on whether or not they match exactly, for example, d("での"(deno), "での"(deno))=0 and d("での"(deno), "の"(no))=1.

For example, in the portion of the thesaurus shown in Figure 4, as defined in section 3.2.2, the MSCA(会議[conference], 滞在[stay]) is 行動[actions]. We assign a real number proportional to the location of the MSCA as shown in the following typical examples. In Figure 4, example (1) is indicated with a broken line, example (2) with a dotted line, and distance for each MSCA is parenthesized in the box.

(1) MSCA(会議[conference], 滞在[stay]) = 行動[actions] then d(会議, 滞在) = 2/3

(2) MSCA( 滞在[stay], 到着[arrive]) = 往来[comings & goings] then d( 滞在, 到着) = 1/3

input: 京都での会議 [Kyoto, conference]				
retrieval examples:				
#	distance	Japanese	English	translation pattern
①	0.4	東京での滞在 [Tokyo, stay]	the stay in Tokyo	B in A
②	0.4	香港での滞在 [Hongkong, stay]	the stay in Hongkong	B in A
③	0.4	東京でのご滞在 [Tokyo, stay]	the stay in Tokyo	B in A
④	1.0	大阪の会議 [Osaka, conference]	the conference in Osaka	B in A
⑤	1.0	東京の会議 [Tokyo, conference]	the conference in Tokyo	B in A
output: B in A → the conference in Kyoto				

Figure 3 Translation procedure

Of course, 0 is assigned to  $d$  when the MSCA is the leaf class, for instance,  $MSCA(\text{京都}[\text{Kyoto}], \text{東京}[\text{Tokyo}]) = \text{地名}[\text{place}]$ , and  $MSCA(N, N)$  for each  $N$ .

### (b) Weight of attributes

In Figure 5, all the examples whose  $E_2(\text{adnominal}) = \text{"での"}(\text{deno})$  are translated with the same preposition, "in". This implies that when  $E_2(\text{adnominal}) = \text{"での"}(\text{deno})$ ,  $E_2$  is an attribute which heavily influences the selection of the translation pattern. In contrast to this, the translation patterns of examples whose  $E_1(\text{semantic code}) = \text{"地名}[\text{place}]$ ", are varied. This implies that when  $E_1 = \text{"地名}[\text{place}]$ ",  $E_1$  is an attribute which is less influential on the selection of the translation pattern. According to the expression (2) in the beginning of section 4.3, weights for attributes,  $E_1$ ,  $E_2$ , and  $E_3$  in Figure 5 are as follows:

$$w_1 = \sqrt{((12/27)^2 + (4/27)^2 + \dots + (1/27)^2)} \approx 0.49$$

$$w_2 = \sqrt{((3/3)^2)} = 1.0$$

$$w_3 = \sqrt{((9/24)^2 + (9/24)^2 + \dots + (1/24)^2)} \approx 0.54$$

### (c) Example distances

The example distances shown in Figure 3 are calculated using the weights in section 4.3 (b), attribute distances as explained in section 4.3 (a) and the expression (1) in the beginning of section 4.3.

$$\textcircled{1} d(\text{京都での会議}, \text{東京での滞在}) = d(\text{京都}, \text{東京}) \times 0.49 + d(\text{での}, \text{での}) \times 1.0 + d(\text{会議}, \text{滞在}) \times 0.54$$

$$\begin{array}{l} \text{[Kyoto, conference]} \quad \text{[Tokyo, stay]} \\ \text{[Kyoto][Tokyo]} \quad \text{[conference][stay]} \end{array}$$

$$= 0 \times 0.49 + 0 \times 1.0 + 2/3 \times 0.54 = 0.4$$

$$\textcircled{2} d(\text{京都での会議}, \text{東京の会議}) = d(\text{京都}, \text{東京}) \times 0.49 + d(\text{での}, \text{の}) \times 1.0 + d(\text{会議}, \text{会議}) \times 0.54$$

$$\begin{array}{l} \text{[Kyoto, conference]} \quad \text{[Tokyo, conference]} \\ \text{[Kyoto][Tokyo]} \quad \text{[conference][conference]} \end{array}$$

$$= 0 \times 0.49 + 1 \times 1.0 + 0 \times 0.54 = 1.0$$

## 4.4 Experiment

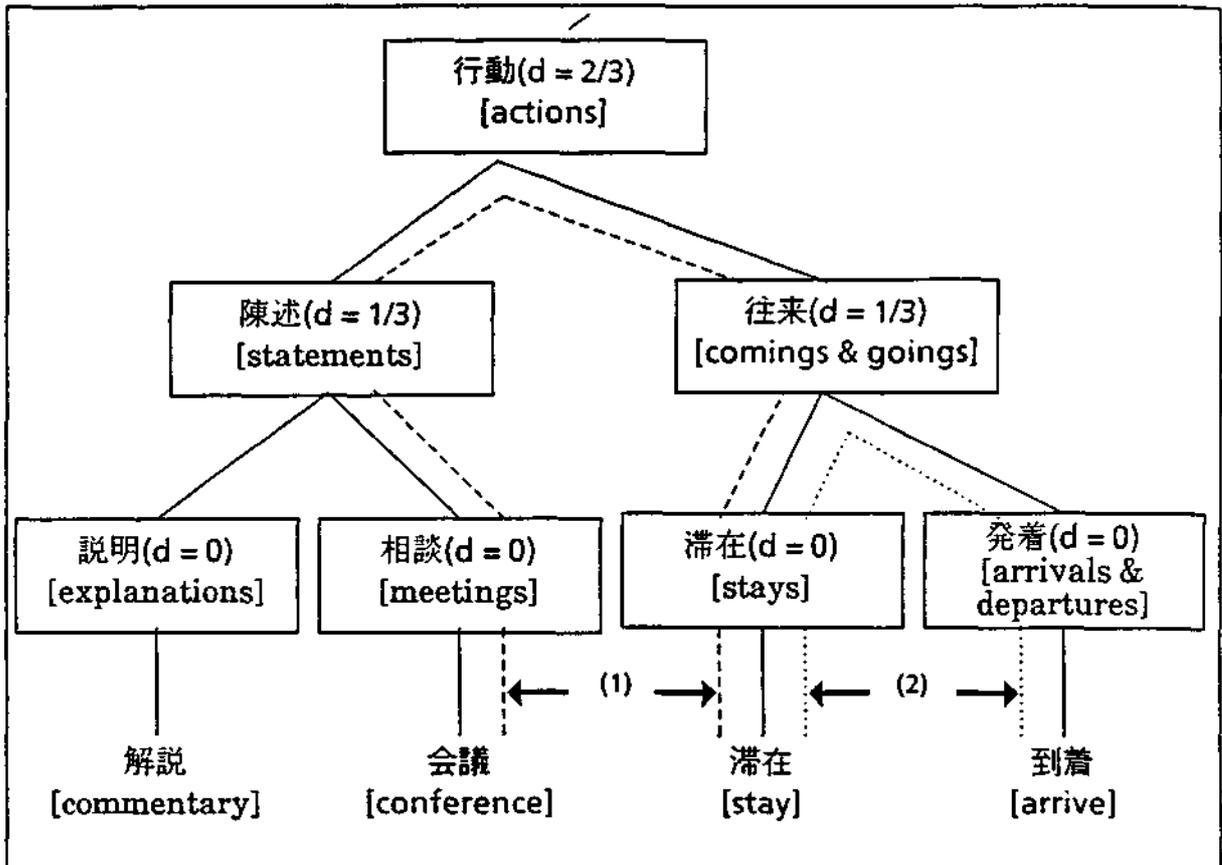


Figure 4 Thesaurus(portion)

$W_1 = 0.49$		$W_2 = 1$		$W_3 = 0.54$	
k	frequency	k	frequency	k	frequency
B in A	12/27	B in A	3/3	B	9/24
↑^A B	4/27	$E_2 = \text{での}(\text{deno})$ [variant of adnominal]		↑^A B	9/24
B from A	2/27			B in A	2/24
B A	2/27			A's B	1/24
...	...			...	...
B to A	1/27			A B	1/24
$E_1 = \text{地名}$ [place]				$E_3 = \text{相談}$ [meeting]	

Figure 5 Weight of the i-th attribute

↑^A means the adjectival form of A

About 700 examples of "N<sub>1</sub> no N<sub>2</sub>" are included in the corpus. The collection of examples continues. Their number is estimated to be about 5,000. The current 700 examples are divided into two groups: (1) Japanese parts of 100 examples are used as inputs of translation; (2) The remaining 600 examples are registered in the example database. The failure ratio is 42%. This seems rather high. However, about 90% of the failures (38 failures out of 42) are caused by a lack of similar examples. In other words, they are easily solved by adding appropriate examples.

**EBMT has made some interesting observations. For instance, some kinds of nouns, such as "方"(kata), used for politeness in Japanese are always omitted in English. In other words, "N<sub>1</sub> no N<sub>2</sub>" is expressed only by the English word for N<sub>1</sub>.**

## 5 Discussions and Comparison with Similar Approaches

### 5.1 Phenomena other than "N<sub>1</sub> no N<sub>2</sub>"

It can be said that when one of the following conditions is satisfied, the phenomena is suitable for **EBMT**.

- (1) It is difficult or time-consuming to formulate translation rules for the phenomena.
- (2) Rule application is too expensive even though translation rules can be formulated for the phenomena.

There are many other phenomena in J-E translation, which are suitable for **EBMT**:

(1) The "da" sentence is typical. Its form is "N<sub>1</sub> wa N<sub>2</sub> da". Here "N<sub>1</sub>" and "N<sub>2</sub>" are nouns, "wa" is a topical particle, and "da" is a kind of verb which, roughly speaking, is the English copula "be". In other words, "N<sub>1</sub> wa N<sub>2</sub> da" corresponds to "N<sub>1</sub> be N<sub>2</sub>" but, of course, the correspondence is more complicated. When "N<sub>2</sub>" is the nominalization of some verb, denominalization is often required. Moreover, some "da" sentences are contractions of normal sentences made by omitting the verb. In that case, the English translation can be rendered only by finding the missing object. As can be seen, the basic structures of the "da" sentence and "N<sub>1</sub> no N<sub>2</sub>" are the same, both consist of two nouns and function words. Thus, the same mechanism is applicable.

**(2) A Japanese aspectual phrase, i.e., a verb with an aspectual marker, is another candidate. Its structure is simpler than that of "N<sub>1</sub> no N<sub>2</sub>" and can be translated using EBMT with a thesaurus of verbs. The Japanese aspectual marker "ている"(teiru), depending on the semantic category of its verb, is translated by the English aspectual markers, "ing"(progressive), "have + pp"(present perfect), or "(simple). However, the Japanese aspectual marker "はじめる"(hajimeru) is always translated by the English aspectual marker "begin to".**

(3) Translation of idiomatic expressions from a composite of the translations of their elements is not possible. This implies that they are not suitable for **RBMT**, but are suitable for **EBMT**. Furthermore, translation of an idiomatic expression can only be used to translate the same idiomatic expression; it cannot be used to translate a similar expression. A mark indicating an example is idiomatic must be added to the example attributes in order to prevent its over-use.

(4) Sato and Nagao [3] show that simple sentences which have one verb with several nouns as its arguments can be translated in the same manner. They use a semantic distance based on word cooccurrence in the example database. The calculation requires huge numbers of examples. As mentioned section 3.2.2, we adopt semantic distance based on the thesaurus to avoid this difficulty.

Concerning the reverse, i.e., English to Japanese, translation of compound nouns is a promising candidate.

### 5.2 Integration with Rule-Based Paradigm

There are two research groups aiming at transferring total sentences by example-based paradigm only: (1) Sato and Nagao recently proposed a representation of matching between the dependency tree of the input sentence and fragments from dependency trees of multiple examples [13]. (2) Sadler presents a simulated translation by retrieval of fragments from examples, while traversing the input

dependency tree in top-down fashion<sup>[4]</sup>. His group also assumes that analysis can be done by similar methods<sup>[4]</sup>.

In contrast with this, we have adopted another easily implemented way to prove the usefulness of **EBMT** by finding a way of integrating **EBMT** with conventional **RBMT**. Because it is not yet clear whether **EBMT** can/should deal with the whole process of translation. We assume that there are many kinds of phenomena: some are suitable for **EBMT** and others are not. In other words, they are suitable for **RBMT**. Thus, it is more acceptable for users if **RBMT** is first introduced as a base system which can translate totally, then its translation performance can be improved incrementally by attaching **EBMT** components as soon as suitable phenomena for **EBMT** are recognized. This agrees with the line Nagao proposed in the first paper<sup>[2]</sup>.

In addition to the naive architecture that calls **EBMT** as subroutines of transfer, the next architecture is proposed. During analysis, **EBMT** is called and returns the translation which is then hidden in memory until generation needs it and the information necessary for the rest of the analysis process, such as syntactic category or structure. This architecture provides **RBMT** with **EBMT** as a bypass when necessary.

**This is done by using cue words which notify that EBMT may be applicable. Cue words are determined from phenomena which are dealt with by EBMT, for example, adonominal case particles "の"(no), "での"(deno), "からの"(karano), "までの"(maden) and so on for "N<sub>1</sub> no N<sub>2</sub>" and "だ"(da) for "da" sentences (section 5.1(1)). Suppose the input is "私[I]はドイツ[Germany]の車[car]を買う[buy]". In this case the cue word is "の"(no). EBMT for "N<sub>1</sub> no N<sub>2</sub>" is called with the input "ドイツ[Germany]の車[car]". EBMT returns syntactic category NP with the translation "German car". The former is used in analysis and the latter is kept in memory until generation requires the target expression.**

This requires only cheap modifications of **RBMT** and **EBMT** as follows.

Additional mechanism for **RBMT**:

- (1) mark **cue words** in the lexicon in advance
- (2) call **EBMT** when a word is marked as a **cue word**

Additional mechanism for **EBMT**:

- (1) search around the **cue word**
- (2) return not only the translations but also the information for the remaining process

When **EBMT** fails to translate, i.e., no similar examples are found, a general rule-based process is called up. This guarantees that the system is fail-safe.

## 5.3 Related Comments

### (a)Sublanguage Bias

**EBMT** is intrinsically biased toward an example database. This is a good feature because it provides a way of automatically tuning to sublanguage. No system has dealt with general language, nor has any system tuned to sublanguage automatically.

### (4) No Rule Generation

Automatic generation of rules from example databases is another solution to problems associated with **RBMT**. However, rule generation has proved to require extensive searches and there are no good **RBMT**-oriented constraints which restrict search space. **EBMT** currently does not support automatic generation of rules.

## 6 Concluding Remarks

**EBMT (Example-Based Machine Translation)** has been proposed. **EBMT** retrieves similar examples (pairs of source texts and their translations), adapting the examples to translate a new source text.

The feasibility of **EBMT** has been shown by implementing a system which translates Japanese noun phrases of the form "N<sub>1</sub> no N<sub>2</sub>" to English noun phrases. The result of the experiment to translate

Japanese noun phrases of the form "N<sub>1</sub> no N<sub>2</sub>" to English noun phrases was encouraging. The more examples obtained, the higher the quality of the translations achieved. The system has been written in Common Lisp, and is running on Genera 7.2 at ATR.

### Future work

- (1) To build systems for other parts of Japanese sentences: "da" sentences, aspect, and idiomatic expressions in order to investigate the strengths and weaknesses of **EBMT**
- (2) To integrate an **EBMT** system with a rule-based translation system
- (3) To incorporate contextual information into **EBMT**

## Acknowledgement

The authors wish to express their appreciation to the management of the ATR Interpreting Telephony Research Laboratories for permission to publish this paper, and also to members of the staff who assisted in its preparation.

## References

- [1] S. Nirenburg: "Machine Translation", p.350, Cambridge University Press, 1987.
- [2] M. Nagao: "A framework of a mechanical translation between Japanese and English by analogy principle", in Artificial and Human Intelligence, ed. A. Elithorn and R. Banerji, pp. 173-180, North-Holland, 1984.
- [3] S. Sato and M. Nagao: "Memory-Based Translation", reprint of WGNL70-9, IPSJ, 1989(in Japanese).
- [4] V. Sadler: "Translating with a Simulated Bilingual Knowledge Bank (BKB)", BSO/Research, 1989.
- [5] P. Brown, J. Cocke, S. Della Pietra, V. Della Pietra, F. Jelinek, R. Mercer, and P. Roossin: "A statistical approach to language translation", Proc. of Coling '88, pp. 71-76, 1988.
- [6] M. Nagao, J. Tsujii, J. Nakamura: "Machine Translation from Japanese into English", Proceedings of the IEEE, 74, 7, 1986.
- [7] J. Kolodner and C. Riesbeck: "Case-Based Reasoning", tutorial textbook of 11th IJCAI, 1989.
- [8] C. Stanfill and D. Waltz: "Toward Memory-Based Reasoning", CACM, 29-12, pp. 1213-1228, 1986.
- [9] E. Sumita and Y. Tsutsumi: "A translation aid system using flexible text retrieval based on syntax-matching", Proc. of the second international conference on theoretical and methodological issues in machine translation of natural languages, CMU, Pittsburgh, 1988.
- [10] A. Shimazu, S. Naito, and H. Nomura: "Semantic structure analysis of Japanese noun phrases with adnominal particles", pp.123-130, Proc. of 25th ACL, 1987.
- [11] K.Ogura, K. Hashimoto, and T. Morimoto: "Object-Oriented User Interface for Linguistic Database", Proc. of Working Conference on Data and Knowledge Base Integration, University of Keele, England, 1989.
- [12] S. Ohno and M. Hamanishi: "Ruigo-Shin-Jiten", p.932, Kadokawa, 1984(in Japanese).
- [13] S. Sato and M. Nagao: "Toward Memory-Based Translation", Coling '90,1990(to appear).
- [14] V. Sadler: "Working with Analogical Semantics", Foris Publications, 1989.