Feedback of Correcting Information in Postediting to a Machine Translation System

Fujio NISHIDA, Shinobu TAKAMATSU, Tadaaki TANI and Tsunehisa DOI

Department of Electrical Engineering,
Faculty of Engineering, University of Osaka Prefecture
Mozu, Sakai, Osaka, 591, Japan

Abstract

This paper presents an attempt to construct a feedback system PECOF which improves a Japanese-English Machine Translation system by feedback of correcting information given by posteditors. PECOF analyzes the error-correcting information by using an English-Japanese Machine Translation system which works in the reverse direction to the original MT system, compares the intermediate expressions of the corrected patterns with those of the erroneous parts of the original MT output at every transfer stage and identifies the responsible parts of the original Japanese-English MT system. Then PECOF corrects the irrelevant parts of the database or adds error correcting patterns to a document of postediting to ask users for further examinations for corrections.

1. Introduction

In recent years, studies of machine translation have been rapidly developed and tend to be put to practical use in various specific fields. [5] However, it is expected that the output sentences from machine translation systems need post-editing, more or less, over a long years for practical use as seen in the report on the practical experience of the Systran Machine Translation system and the posteditors' experience. [1]-[4]

As can be easily seen, feedback of correcting information given by posteditors to the original MT system undoubtedly will bring a remarkable improvement of the translation proficiency to the system itself. However, it does not seem that adequate discussion about feedback of the information for improving of the MT system has been done so far. One of the main reasons will be that it is difficult to identify the part to be corrected in the MT system only by using brief correcting information.

This paper presents an attempt to identify the responsible parts of a machine translation system in the case of Japanese to English translation. The part to be corrected in the MT system is identified by applying an English-Japanese Machine Translation system to the postedited output in the reverse direction to the original translation. The English-Japanese MT system is assumed here to be capable of analyzing and comprehending the postedited output at least by using the fundamental and the general linguistic knowledge of the target language. Associated with the assumption,

many of erroneous patterns arise due to lack of information about specific usages of words rather than lack of general linguistic information. Accordingly, such kind of correcting information given by posteditors can be understood from the general linguistic information.

PECOF(abbreviation of a PostEditing COrrecting information Feedback system) analyzes the correcting information by using the English-Japanese MT system and tries to perform feedback of it to the MT systems. In the following sections, the basic idea and the construction of PECOF are described in some details and correcting of typical error patterns is illustrated with some examples.
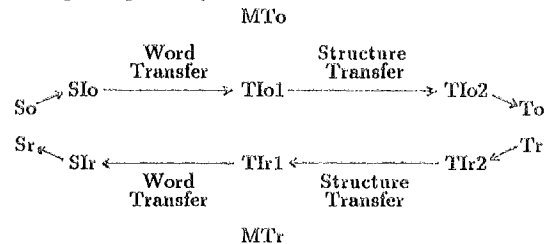
2. The principle of system construction

Fig.1 Machine translation systems MTo and MTr

Let us assume that the MT system is constructed based on a transfer system. As shown in Fig.1, the original machine translation system MTo parses a block of source language sentences So, constructs the intermediate expression or form SIo, transforms it to the target language intermediate form TIo1 and TIo2 by word and structure transfer and finally generates a block of the target sentences To.

In order to identify the kind of corrections given in a target language expression, PECOF needs a simple syntactic and lexical analyzer of the target language at least. If a part of the target sentences is corrected and yet TIo2 remains unchanged except for some syntactic term expressions, a syntactic rule corresponding to some specific word usages of the target language is checked, the inappropriate part of syntactic rules is identified and modified according to the correction given by posteditors.

If the correction spreads all over the parts of the output sentences or is made by entire replacement, a pars-

ing system of the target language is needed. Furthermore, if some wrong transfers are made over several transfer stages, an MT system which works in the reverse direction is much required to identify them. It is called the reverse MT system briefly and denoted with MTr.

In similar to the original MT system, the reverse MT system parses the corrected target language expression Tr, constructs the intermediate form TIr2 and TIr1, transforms TIr1 to the intermediate form SIr of the source language and generates the source language expression Sr which should be almost the same as the original source language expression So. The order of word transfer and structure transfer in the reverse MT system is reverse to that of the original MT system. The relation between both the MT systems is illustrated in Fig.1. The case system, the semantic category system and the intermediate form in the reverse MT system are the same as those of the original MT system. The intermediate form in both the systems consists of several pairs of a case label and a term having the semantic and the syntactic category name in option.

PECOF analyzes post-edited output by using the reverse MT system and searches for the first transfer stage in which the intermediate form of the corrected output obtained by the reverse MT system differs from that of the original MT system.

After identifying the word block to be corrected and the kind of correction PECOF tries to correct the corresponding part of the word dictionary and the transfer rules. Various methods from documentation to automatic correction can be considered. From a practical point of view, it will be efficient to give various error patterns through the intermediate forms and correspondingly to provide the correcting procedure of the database of the original MT system. If unknown error patterns occur, PECOF only classifies the patterns to ask the posteditors about the correcting method.

After processing the wrong parts in the current transfer stage, PECOF updates the intermediate forms of the following stages of the original MT system output based on the corrections performed in the current stage. When some discrepancies still remain between the intermediate forms of the original output and those of the postedited output, PECOF applies the same correcting procedures to the following transfer stages repeatedly.

The typical error patterns are classified into three classes. They are related to syntactic structure of target language, structure transfer and word transfer. In section 4, some types of the error patterns and the corresponding correcting procedures are described.

3. System construction and correcting information feedback

In this section, the construction and the function of PECOF as well as those of both the MT systems are described in some details.

3.1. Overview of the MT systems

Figure 2 shows a schematic construction of our MT systems and the database. The systems belong to a kinds of transfer systems. [6] The body part of the processing, MAPTRAN, is divided into parsing, transfer

and generation. Three kinds of dictionaries are implemented for word transfer, rewriting rules and structure transfer.

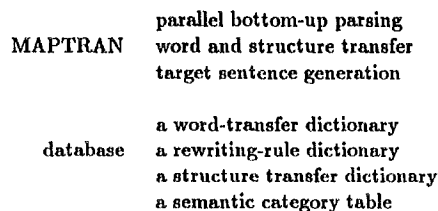|	
MAPTRAN	parallel bottom-up parsing
	word and structure transfer
	target sentence generation
database	a word-transfer dictionary
	a rewriting-rule dictionary
	a structure transfer dictionary
	a semantic category table

Fig.2 Construction of the MT systems and the database

Both the MT system from Japanese to English and that of the reverse direction use the same processing system MAPTRAN. [7] It is constructed on a hierarchical module basis and can be expanded into some computer languages such as C and LISP. [8] It parses the input sentences in a parallel bottom-up manner.

The word transfer dictionary used here is constructed by combining an original word transfer dictionary with a source language word dictionary for analysis and a target language word dictionary for generation. The data structure of the dictionary can be semi-automatically transformed to an appropriate form corresponding to the change of the programming language.

3.2. Designation of correction

In order to designate the location to be corrected, a number is attached to each word in the output sentences. A word and a word sequence can be designated by a number or by a pair of the first word number $n1$ and the last word number $n2$ like $n1$-$n2$. Replacement of words is designated as follows:

replace $\underline{n1}$ (or $\underline{n1}$-$\underline{n2}$) by " $\underline{a\ new\ word\ sequence}$ "   (1)

Insertion and deletion can be indicated in a similar manner.

Movement of a word group ranging from the word number $n1$ to $n2$ to the front of the word of the number $n3$ is designated as follows:

move $\underline{n1}$-$\underline{n2}$ to $\underline{n3}$.   (2)

For a given correction, it will be not sometimes easy for PECOF to identify the key item to be corrected in the dictionary when the original MT system does not have the sufficient linguistic knowledge and has a lot of possible reasons for correction. In such cases, it will be effective that PECOF is informed of the key information of correction by posteditors.

However, detailed descriptions are laborious for posteditors and hard for PECOF to comprehend. One way to solve this problem is to indicate one or two words that conflict with the words to be corrected in the output sentences or phrases. Besides, it is sometimes desirable to add some words that stand for the kind of correction. The designation is given in a form where the reason is added to correction in option. For example, the reasoning for (1) is written as follows: where it ($n1$ or $n1$-$n2$) conflicts with $\underline{n3}$ in terms of $\underline{R1}$. In the above the underlined parts are words to be given by a posteditor. R1 stands for a kind of correction like

TENSE conflict and SEMantic CATegory conflict.

Similarly, the posteditor can add a syntactic symbol such as a part of speech a technical term a compound and an idiom to a word group involved in the given new word sequence if necessary.

3.3. Correction by PECOF

Figure 3 shows the main functions of PECOF.

PECOF
- (1) Analysis of the corrected parts by using the reverse MT system
- (2) Identification of the part to be corrected in the database
- (3) Correction of the part of database or documentation of the corrected patterns which cannot be completely identified

Fig.3 The main functions of PECOF

3.3.1. Installation

In order to keep lexical information of the words appearing in the source language sentences till the end of correcting, PECOF needs some record type database. The current dictionaries are constructed there by copying the parts of the dictionary in the file and modification of the dictionary are performed on the database.

The reverse MT system needs the fundamental rewriting rules to be implemented based on the case grammar to parse and comprehend the target language output modified by posteditors. Most of parsing systems based on a case grammar will be available with some modification.

3.3.2. Performances

First, PECOF makes the reverse MT system to analyze the postedited output and construct the intermediate forms. When the postedited output involves some words which are not contained in the word dictionary and also in the syntactic information given by the posteditor, PECOF tries to identify the syntactic information and the equivalents of the unknown words by referring to the information of the related words in both the target and the source language expressions or by asking the posteditor about them later.

After constructing the intermediate form, PECOF compares it with that of the original MT system. If there are some differences between them, PECOF makes the reverse MT system to further transfer the intermediate form in the reverse direction. If the reverse MT system has sufficient capability of translation, it will be able to yield almost the same intermediate form as the original MT system at a certain transfer stage though the same form might be able to be obtained only in the given source language sentential expressions. Furthermore, if the original MT system can parse and normalize the source language sentences correctly, both the intermediate forms coincides with each other by the end of the word transfer stage of the reverse MT system at the latest.

If the same intermediate form is obtained, PECOF stops the transfer by the reverse MT system and begins to backtrack. Then PECOF tries to remove the difference between the next-stage intermediate form of the original MT system and that of the reverse MT system.

More precisely, PECOF identifies the irrelevant part of the intermediate form of the original MT system by comparing it with that of the corrected results given by the reverse MT system and corrects the data and the applied conditions in the database according to the procedures determined from the difference patterns.

If the reasoning of corrections is given in a form of the conflicting words and the associated information as mentioned in section 3.2, PECOF examines the data to be corrected and the irrelevant applied conditions by referring to the syntactic and semantic attributes of the conflicting words, and corrects the data and the applied condition to be more relevant by refining unifying or replacing the old conditions.

Some databases such as word dictionaries omits detailed items like the applied conditions of a word if they are generally held. These deficit items are implicitly designated by a general condition table implemented for each category of words. In such cases, if a specific equivalent is designated together with the reasoning by a posteditor, the applied conditions of the equivalent derived from the reasoning are written in the corresponding part of the record of the current word dictionary directly or through a pointer.

If the correction in postediting lacks detailed information about wrong translation and confident reasons necessary for correcting the database, PECOF arranges the related parts of corrections of the corresponding original target and source expressions, classifies them by some attributes of the error patterns and adds them to a document of error patterns. Then PECOF urges the users to formulate the correcting procedures of the part of database corresponding to the error patterns.

4. Miscellaneous correction information in postediting

4.1. Syntactic structure correction

Every part of the translated output is required to meet syntactic patterns of the target language even by modifying and complementing the given source language sentences. The occurrence of some syntactic errors and their corrections in the target language expressions can be detected when some parts of the target language expressions are corrected though the intermediate forms are the same as those of the corrected expressions except for some syntactic term expressions.

4.1.1. Word expressions

Let us describe the rewriting rules of a noun phrase as follows:

RR:ADJP (Rewriting Rule: ADJective Phrases)
$$<NP(n1\{v\}(PRED:*, OBJ:n2,..))>$$
$$::=<NP1(n1)><PREP><NP(n2)> \qquad 1$$
$$<NP(n1(PRED:v, K1:*,K2:n2,...))>$$
$$::=<NP1(n1)><INF(PRED:v,K1:n1,K2:n2,...)> \qquad 2$$
$$\qquad (3)$$

The above expressions are useful for transformation between sentences and the intermediate expressions in parsing and generation of sentences.

In the rewriting rule expressions, $ni(i=1,2)$ and v are a noun term and a verb term respectively. $n1\{v\}$ means that $n1$ is a noun term derived from a verb term. $<NP(t)>$ and $<INF(t)>$ denote the non-terminal sym-

bols of a noun phrase and an infinitive phrase corresponding to a term expression or an intermediate form t. The symbol * denotes the term prefixed to a frame which includes and modifies the symbol *. In this case, it stands for n1{v}. PRED and OBJ denote a PREDicate case label and an OBJect case label respectively. K1 and K2 stand for some case labels. The term expressions of the left side hand of the rewriting rule (3)-1 describe that n1{v} is modified by n2 which depends on a predicative noun n1 as the objective term.

In usual cases, the preposition used for modifying a noun by a noun is the preposition "of" and "of" is taken for the deficit value of the preposition. If a specific preposition "prep" is indicated to the noun term n1{vt} by a postediter, PECOF records "prep" in the preposition item of the noun word n1 of the word dictionary together with the applied rewriting rule as follows:

MODifier:<PREP>,<PREP>:prep,RR:ADJP1   (4)

Example 1

(a)..inquiry 'of ← into' the question..
(b)..discussion 'of ← on' the question..

where 'a ← b' means the replacement of 'a' by 'b'.

Corresponding to error corrections in postediting in the above, the recorded items of "prep" in the word dictionary are "into" and "on" in a form of (4) respectively.

The correcting information of the other kinds of syntactic errors can be fed back to the original MT system in a similar way.

4.1.2. Rewriting rule construction

An MT system sometimes lacks some rewriting rules. For example, technical papers often devise and use a concise phrase expression instead of a long complicated expression under the condition that no ambiguity is brought.

Let us consider the following rewriting rules of a noun phrase:

<NP(n1(PRED:v,K1:*,K2:n2,..))>
::=<NP1(n1)><RLC(PRED:v,K1:*,K2:n2,...)> (5.1)
<NP1(n(OBJ:*,DET:det))>
::=<DET(det)><NP2(n)>   (5.2)

where RLC and DET stand for RELative Clause and DETerminer respectively.
When an intermediate form

car(OBJ:*(OBJ:*,DET:a),PRED:be drive-en,
    AG:motor)   (6.1)

is given, a relative clause

" a car that is driven by motor"   (6.2)

is generated by the rewriting rule (5.1) and others. Now , suppose that the relative clause is replaced by a phrase
"a motor driven car"   (7.1)

in postediting. PECOF infers the rewriting rule which rewrites (6.1) to the above corrected phrase and adds to (5.1) an alternative rewriting rule

::=<DET(det)><NP(n2)><PastParticiple(v)>
    <NP(n1)>   (7.2)

Some words can be used only by a specific syntactic

rule which belongs to a general syntactic rule. The sentences that include these words can be parsed and interpreted by the general rules which are not conditioned by various syntactic patterns in details. On the contrary, the generation of the sentences needs the information of the specific usages of the words. For example, the verb "doubt" conventionally takes the OBJect term through WHETHER-CLAUSE in an affirmative sentence and THAT-CLAUSE in a negative sentence. The information is needed to generate the target sentence from the intermediate form TIo2 and is given by some rewriting rules or by complementing the rewriting rules by means of adding the necessary information to the word dictionary as shown in Expr.(4).

4.2. Transfer rule correction

4.2.1. Structure transfer

If the intermediate form TIo1 is the same as that of the postedited output TIr1 though TIo2 is different from TIr2, PECOF tries to search for some structure transfer rules to meet the corrections. Structure transfer is needed so that natural and conventional target language expressions can be generated corresponding to given source language expressions.

As well known, Japanese tends to avoid the use of non-animate subjects and also interpret an event as a change of a state due to some causes rather than action on some objects by an agent. A general structure transfer rule between state change and action is given as follows:

STR:CR1(Structure Transfer Rule :Cause Result 1)
    (PRED:v1, OBJ:n1, CAUSE:n2)
→(PRED:v2, AG:n2, OBJ:n1)   (8.1)

where v2 and v1 form a pair of verbs of 'cause and result' or 'action and state change'. The typical concrete instances of them are "cause to do" and "do", "lead" and"reach", "show" and "be seen" and others.

The above fundamental rule (8.1) is implemented in the database of the original MT system as well as the reverse MT system in the structure transfer database. The specific information of cause verb "v2" for a given result verb "v1" is sometimes recorded in the word dictionary together with the name of the applied transfer rule as follows:
    CAUSE-V:v2 STR:CR1   (8.2)

When PECOF finds that an original output is postedited by the structure transfer of this type and by using a cause verb v2, PECOF records v2 in the v1 item of the word dictionary according to the form (8.2).

4.2.2. Word transfer

The word to be corrected in word transfer can be identified by using the information about replacement of full words such as verbs adjectives and nouns. More precisely, it can be identified from a term of the target language intermediate form TIr1 modified by a postediter. Some patterns to be corrected are shown in the following.

In a case structure of a language, the semantic categories of a governor and the dependants are, more or less, bounded. Furthermore, the conventional category boundary set of terms in a frame of the target language is not always the same as that of the source

language and sometimes structure transfer is needed.

The discrepancy between the semantic category boundary of a word and that of the equivalent is revealed when the word is linked with some dependants or the governor. Hence the equivalent must be chosen so that no conflict occurs on the semantic category conditions of the linked words in the target language .

Let us assume that a word t1 of a source language can be represented by an equivalent t11$'$ or t12$'$ of the target language in the following condition:

HEADWORD: t1

$$\text{EQUIV:t11}' \quad \text{COND:(K1-C1:t1, K2-C21,...)}, \qquad (9.1)$$
$$\text{EQUIV:t12}' \quad \text{COND:(K1-C1:t1, K2-C22,...)}, \qquad (9.2)$$

where K_i and C_i ($i=1,2,...$) denote a case label and a semantic category name respectively and both t11$'$ and t12$'$ are the equivalents of t1 under the condition that the term t1 appears in the context of the case-category label of K1-C1 accompanied with a term of case-category label K2-C21 or K2-C22.

Furthermore, suppose that a posteditor replaces the equivalent t11$'$ by t13$'$ under a condition that the word t1 is accompanied with a word t23 of a case-category label K2-C23 and C23 is a subcategory name of C21.

Then, the equivalent applied condition (9.1) of the word t1 of the word dictionary are replaced as follows:

HEADWORD:t1

$$\text{EQUIV:t11}', \quad \text{COND:(K1-C1:t1,K2-\{C21-C23\},...)}, \quad (9.3)$$
$$\text{EQUIV:t13}', \quad \text{COND:(K1-C1:t1,K2-C23,...)}, \qquad (9.4)$$

If t13$'$ is the same as t12$'$, (9.4) and (9.2) are unified as follows:

$$\text{EQUIV:t12}', \quad \text{COND:(K1-C1:t1,K2-\{C22+C23\},...)}, \quad (9.5)$$

where $\{C_i \pm C_j\}$ denotes the union or the difference set of the sets expressed by the category names C_i and C_j.

Example 2

Japanese represents emotional state expressions by using adjective verbs. They have the same form irrespective of the active or the passive type. On the other hand, English has a different expression depending on whether the object to be described is active or passive. For instance, they are "interesting" versus "interested", "enjoyable" versus "enjoy oneself", "exciting" versus "excited" and others.

Let us take an example. It is assumed that the main part of the Japanese-English word dictionary of a word "TAIKUTSU-DE ARU" or " be boring" in English is given as follows:

HEADWORD :退屈である(TAIKUTSU-DE ARU)
  Part of Speech:adjective verb, ROLE:PRED-ATTR.EMOTION
  EQUIV: be boring   (9.6)

Let us suppose that for a given Japanese sentence "WATAKUSHI-WA KARE-NO KOUEN-GA TAIKUTSU-DE ATTA" a modified sentence "I was (boring ←bored) in his lecture" is given by a posteditor. PECOF identifies the wrong expression and refines the corresponding items of the word dictionary by referring to the format (9) as follows:

EQUIV:be boring,COND:(PRED-ATTR:*,
        OBJ-THINGS,RECIPIENT-HUMAN),   (9.7)
EQUIV:be bored, COND:(PRED-ATTR:*,
        EXPER-HUMAN,OBJ-THINGS),   (9.8)

where * stands for the word "TAIKUTSU-DE ARU".

Example 3

A Japanese verb "oshieru" has several meanings, namely, "teach, tell, show, instruct and so forth". The object case categories of "teach" and "tell" of English in this context are a school subject like mathematics and simple information like a way to a station respectively.

The system MTo chooses the equivalents by examining these category conditions. If the postediting results suggest that the system needs more precise conditions for equivalent selection, PECOF tries to refine the category conditions or to add the other case conditions.

4.3. Idiomatic expression transfer

An idiomatic expression (including a compound) consists of several words and has a somewhat definite meaning when compared with a single word. Every language has various unique idiomatic expressions. It is considered that one of the most efficient and reliable methods in machine translation is to cover the main content to be translated with a frame of an appropriate idiomatic expression of the target language. Therefore, it will be necessary for a good MT system to have a large storage of various idiomatic expressions and to be capable of selecting an appropriate target idiomatic expression to cover the main part of a given source language expressions. PECOF helps gathering the idiomatic expressions given by posteditors.

At first, PECOF identifies an output part modified by a posteditor to be an idiomatic expressions. The information of identification is given by the posteditor or is inferred from a phrase form of the replaced part. PECOF selects a keyword among the words involved in the idiomatic expression and identifies the corresponding source language word by referring to the current word dictionary mentioned in section 3. Then PECOF records the corresponding source language idiomatic expressions "$\alpha 1$ t $\beta 1$" at the location labeled with the heading of the keyword "t" in the word dictionary so that the idiomatic expression can be applied with priority in the word transfer by the longest match method. If a similar heading idiomatic expression "$\alpha 2$ t $\beta 2$" which has the same target expressions is already involved in the location, these heading expressions are unified as follows:

"$\{\alpha 1,\alpha 2\}$ t $\{\beta 1,\beta 2\}$" $\leftarrow$ "$\alpha 1$ t $\beta 1$", "$\alpha 2$ t $\beta 2$"

The semantic and syntactic information necessary for the dependants and the governor are added to the idiomatic expressions.

Example 4

HEADWORD: 重点 (JYUTEN)

$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$

IDIOM1: に重点を{おく、しぼる}
     (NI-JYUTEN-WO{OKU,SHIBORU})

(PRED-THINKACT: 重点を{おく、しぼる},
LOC-THINGS:-に)

put emphasis on

(PRED-THINKACT:put emphasis ,
LOC-THINGS:on-)

5. Some illustrative examples

This section shows some illustrative examples of the modification of dictionaries by posteditors' correcting information. They are to be tested by the experimental system PECOF which is under construction.

Example 5.1

The input Japanese sentence So and the output To of the original MT system are given respectively as follows:

So: その工具には臨界切削領域から熱を除く能力がある。
To:The tool has the ability which removes heat from the critical-cut-range.

The intermediate form TIo2 of the above is

TIo2:(PRED:have,POSSESSOR:tool(OBJ:*,DET:the),
OBJ:ability(PRED:remove,
OBJ:heat, INSTR:*,...))

The output Tr corrected by a posteditor is

Tr:The tool has the ability to remove heat from the critical cutting area.

The intermediate form TIr2 of the above is the same as Tio2 except for the technical term "critical cutting area".

First, PECOF finds that the intermediate form TIo1 of the original output is the same as that of the postedited output TIr1 if the word "critical-cut-range" is replaced by "critical-cutting-area" according to posteditor's designation. PECOF corrects the word dictionary and the intermediate forms of the following transfer stages based on the replacement and finds that discrepancies remain only in the target surface expressions.

PECOF examines the rewriting rules applied to TIr2, recognizes that the form to be applied to "ability" as a modifier is an infinitive and adds the information to the MODifier item of the word dictionary as follows:

HEADWORD: 能力 (NOURYOKU)

....

EQUIV:ability,...,
MODifier:<INF>, RR:ADJP2

Example 5.2

The input Japanese sentence So and the output To of the original MT system are given respectively as follows:

So: この時分割動作によって複数の可変デューティ比の
出力が得られる。
To: A plurality of outputs of variable duty ratio is obtained with this time division operation.

The intermediate form TIo2 of the above is

TIo2:(PRED:obtain(VOICE:passive),
OBJ:output(OBJ:*,NUM:plurality,
ATTR:variable-duty-ratio),
CAUSE:time-division-
operation(OBJ:*,DET:the))

The output Tr corrected by a posteditor is

Tr:This time division operation provides a plurality of outputs of variable duty ratio.

The intermediate form TIr2 of the above is

TIr2:(PRED:provide(VOICE:active),
OBJ:output(.....),
AG:time-division-operation(....))

PECOF compares TIo2 with TIr2, recognizes that TIr2 can be obtained by applying the general Structure Transfer Rule CR1 to TIr2 with the specification of CAUSE-V being "provide" for RESULT-V of "obtain", and then writes down the information in the corresponding item in the word dictionary as follows:

HEADWORD: 得る (ERU)

....

EQUIV:obtain,....

....

CAUSE-V:provide STR:CR1

6. Conclusion

This research is still under the early stage and needs a lot of experimental investigation. This paper shows a method of modification of the database for a comparatively definite error patterns. There will be left various kinds of indefinite error patterns which should be characterized classified and corrected by some formulated procedures. However, the basic idea and system presented here will be useful for improving the translation quality of the MT system and collecting new technical words and idiomatic expressions. Furthermore, if both the original and the reverse MT systems have the posteditors feedback respectively and cooperate with each other, the integrated system will be very powerful and efficient to improve the translation quality and capability in the bilingual translation.

Reference

(1)Pigott I.M.(1981) 'The importance of feedback from translators in the development of high-quality machine translation.' In:V.Lawson(ed.): Practical Experience of Machine Translation: North-Holland,pp.61-74.

(2)Green R.(1981) 'The MT errors which causes most trouble to posteditors.' ibid. pp.101-1.304.

(3)Bernard Lavorel(1981) 'Experience in English-French post-editing.' ibid. pp.105-1.309.

(4)Rossi F.(1981) 'The impact of posteditors' feedback on the quality of MT.' ibid. pp.113-1.317.

(5)Carbonell J.G. and Tomita M.(1987) 'Knowledge-based machine translation.' In:S.Nierenburg(ed.): Machine Translation:Cambridge Univ.Press,pp.68-89.

(6)Nishida F. and Takamatsu S.,(1982) 'Japanese-English Translation through Internal Expressions.' In: Proc. of COLING 82, pp.271-276.

(7)Nishida F., Fujita Y. and Takamatsu S.(1986) 'Construction of a Modular and Portable Machine Translation System.' In: Proc. of COLING 86, pp.649-651.

(8)Nishida F., Fujita Y.and Takamatsu,S(1987) 'Refinement and Error Detection of Program Specifications by a Linking Technique of Library Modules.' IPSJ,28, 5,pp.486-498.(written in Japanese)