Machine Translation from Japanese into English

MAKOTO NAGAO, JUN-ICHI TSUJII, AND JUN-ICHI NAKAMURA

Invited Paper

This paper describes the outline of our Japanese to English machine translation system, which is supported by the Agency of Science and Technology of the Japanese Government. Many new methodologies are introduced to obtain high-quality translation results. The analysis is based on case grammar, which is suitable for a word-order-free language such as Japanese. The dictionary is rich enough to handle many specific expressions. It contains not only case frame information, but also semantic information, idiomatic expressions, and many others. In the transfer phase, the system applies many structural transformations, so that the structural difference of the same contents in Japanese and English can be relieved. In the generation phase, many structural transformations are again applied so that the ellipsis problems can be avoided, and that better stylistic expressions can be obtained. The system is running mainly for the abstracts of scientific and technical papers. The evaluation method of the translated results is also discussed, with many example translations.

I. Introduction

There is a theory of language translation which was advocated by Y. Nida, that the translation is a mapping of the sentential meaning in one language to another. That is, the utterance must be interpreted not only in the context of utterance and situation, but also by the culture of the language and the society. Therefore, the surface sentential structure may not necessarily be preserved, but the translated (or created) sentence must give the same mental effects to the receivers of the translated language as to the receivers of the source language. That would be a wonderful translation. This theory, however, presupposes the existence of the same cultural background in different language societies. If such does not exist, in country pairs such as Japan and the United States, we have to have some other translation theories.

Many western literary works translated into Japanese may not stir the Japanese readers with the same mental effects as those affecting the U.S. readers, but the fundamental appreciation of that literature by the Japanese readers will not be hampered by such incomplete translation. We can accept literary translation and can imagine that the way of thinking of western people in such and such. Foreign concepts and customs in different countries can be appreciated through imperfect translation. Language has the ability of explaining totally complex concepts foreign to the reader so that he can understand them as precisely as he wishes. We can get pseudo-experience of foreign culture through language, which basically must be obtained by real experience.

There are many kinds of research on language comprehension and some trials to apply it to machine translation, even though there is no precise definition and consensus about what language understanding by computer means. We think this kind of language understanding framework is quite doubtful and inapplicable to a practical machine translation system because we have no way at present of memorizing in the computer the cultural background of different languages and nations which are essential by the Nida's translation theory. We think that the standpoint of literal translation is inevitable between languages of different families. The language understanding ability of a human being will cover the gap of incomplete translation. When the documents to be translated are limited to a specific domain of science and technology, there is no difference of cultural background between the languages. The problems can be purely limited to the syntactic, semantic, and contextual information plus culture-independent, domain-specific knowledge. Within this scope we can escape the serious problems which Nida pointed out for the translation. The machine translation system we are developing is limited to this scope of the same knowledge domain between languages.

The levels of translation may be classified into the following four:

1) Free and creative translation which aims at the same mental reactions by the readers of source and target languages.

2) Sentence-by-sentence translation in the free sentential style. Language particularities are fully considered in translation.

3) Literal translation. The sentential structures of the source language remain strongly in the target language.

Manuscript received August 27, 1985; revised January 27, 1986.

The authors are with the Department of Electrical Engineering, Kyoto University, Sakyo, Kyoto 606, Japan.

0018-9219/86/0700-0993$01.00

However, the selection of translation words is correct from the standpoint of semantics.

4) Mechanical translation. The sentential structure of the target language is a crude mapping of the one in the source language, and the selection of translation word is almost one to one.

Translation at the fourth level may be possible between languages with similar structures, like Japanese and Korean, or French and Italian. But this translation level is quite unsatisfactory between languages of quite different characteristics such as Japanese and English. Many of the current commercial machine translation systems are at the fourth level. Systems of this level usually require heavy revision of the translated text by a qualified translator (often called revisor, or post-editor). This is often unbearable for the revisors, because they are irritated by the double tasks of understanding the content from bad translation, and finding a way to revise the translation to increase understandability. They feel that it would be better to carry out the human translation themselves than to suffer the revision task. Another element affecting the possible usage of the level four system is the possibility of controlling the input sentences. Sentential structures and the meanings of words are to be strictly limited. The writer must obey this framework, and the original text must be pre-edited according to this framework.

Our target for machine translation from Japanese into English is of the second level translation, but our present achievement is estimated as that between the second and the third levels. We do not aim at translation of literary works, but at as accurate as possible information transfer of scientific and technical papers. Even with such a limited purpose we have been forced to introduce quite sophisticated linguistic processes in our machine translation system between Japanese and English. This is due to the great difference of linguistic structures between those two languages. Some of the typical characteristics of the Japanese language compared to English are given in the following.

1) A Japanese sentence is composed of a continuous string of three kinds of characters; namely, Chinese characters, Hirakana characters, and Katakana characters. Automatic segmentation of a character string into words is quite difficult. Dictionary consultation of character strings by the longest match principle is generally used with the additional consideration of the inflectional variations.

2) The word order is comparatively free in Japanese compared to English. Postpositions attached to nouns indicate some limited case usages of noun phrases, but the usage is not always unambiguous. A typical postposition "GA" is used to indicate subject and object positions. The disambiguation is usually done by semantics and contextual information.

3) When an animate noun appears as an agent in a sentence it must also appear as the subject in the sentence. Varieties of nouns can be a subject in English, but not in Japanese.

4) Frequent omission of the subject noun occurs, especially when it is a human. Very often the subject pronoun "I" is omitted in Japanese. Sometimes the object noun is also omitted when the context allows it.

5) Words of the Japanese language have no plural form. The distinction of the number is to be done by semantic/contextual information of the sentence.

II. Design Principles of a Machine Translation System

A. General Principles of Our Machine Translation System

The basic principles of our system are as follows:

1) Use every piece of surface and syntactic information as much as possible. Write as detailed syntactic rules as possible by dividing the set of grammatical rules into subsets. Each subset is called a subgrammar and corresponds to a linguistic phenomenon, such as noun phrase conjunction, embedded sentence determination, and so on.

2) Develop a grammar writing system (software system) which can accept any present or future sophisticated linguistic theory.

3) Semantic information is introduced to help syntactic analysis, transfer, and synthesis to be as accurate as possible. We aim at a well-balanced usage of syntax and semantics in the whole process of machine translation. Machine translation based on the use of domain-dependent semantic network will be fascinating and effective for sentences from narrowly limited areas. However, a system of this kind cannot cope with the complicated situations of a wider world, whose semantic description is almost impossible in reality. Therefore, we must primarily rely on syntax.

4) Detailed linguistic phenomena are more likely word-specific than explainable in the framework of general linguistic theory. Therefore, word-specific rules are to be accepted in the system. In our system, such grammatical rules are written as past of the entry of lexical items, and are given priority over the general grammatical rules in all the phases of analysis, transfer, and synthesis. This mechanism allows the quality of the system to be improved step by step by the accumulation and correction of linguistic facts and word-specific rules in the dictionary, and we can avoid the deadlock of the system's improvement.

5) The system must have the ability not to fail because of imperfect analysis and unknown words, but to produce translation result even with imperfect sentential structure and unknown original words. This imperfect output will be much better for the post-editor than the system's failure.

B. Computational Techniques

To realize the above conditions we developed a powerful grammar writing system. A typical feature of this system is that a linguist can write a grammar divided into a set of subgrammars, each of which corresponds to a particular grammatical phenomenon, such as conjunctive noun phrase, embedded sentential parsing, and so on. A linguist can write a control structure for the analysis (transfer and synthesis) process to a certain extent, such that a certain subgrammar is followed by another subgrammar under a certain condition, and so on. The rewriting rules of the system can specify tree to tree transformations. Each node of a tree has an arbitrary number of attribute–value pairs. The system has a powerful matching mechanism of trees with the attributes. This rewriting rule system is quite powerful for many complicated linguistic phenomena. Further strength is given to the system by adding the capability of writing a LISP program associated with each rule to check and change any part of a tree in any way the user wants. It has also the ability to call the rewriting rules which are specific to lexical items and which are stored in

the dictionary entries of these lexical items. This grammar writing system is called GRADE, and is implemented by LISP. GRADE has such expressive power that it is used at all the stages of analysis, transfer, and synthesis.

C. Linguistic Framework

The machine translation process can be largely divided into analysis, transfer, and synthesis, but we have three additional stages to cope with the difficulties in language pairs such as Japanese and English, which are shown in Fig. 1. One is the pre-transfer loop, which converts the analyzed structure of a Japanese sentence into a much more neutral

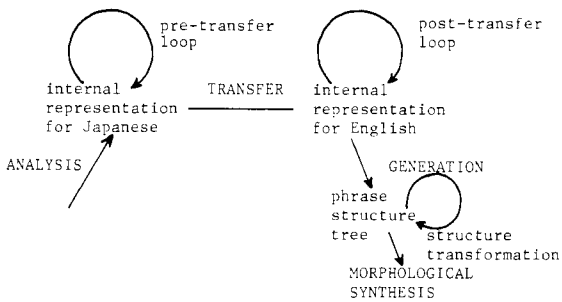


Fig. 1. Processing flow for the transfer and generation stages.

deep structure representation. Another is the post-transfer loop, which converts the deep structure obtained by the transfer stage to a much more appropriate internal representation of the target language. The third one is the structural transformation in the target language during the generation process to obtain a better stylistic expression in such cases as, for example, that of a noun appearing in the "tool" case slot, which can never appear in a subject position in Japanese, while it occurs very often as subject in English. Similarly, a top-heavy sentence in English will be replaced by the anticipatory subject "it." All these three additional stages are optional, and the system can produce target language sentences without them. But for machine translation between languages from completely different language families, like Japanese and English, structural transformations during these stages are essential.

At the moment there are about sixty subgrammars for the analysis, and about 900 rewriting rules in total. The sentence generation process is also composed of a subgrammar network. There are about 800 rewriting rules for the transfer and generation processes, and the number will be increased further in the near future.

At present, the dictionary contains about 16 000 items, and it will be increased up to 70 000 items in the near future. Among these 16 000 items, verbs and adjectives number about 2000, adverbs about 400, and the remainder consists mainly of special terminology words. The information to be written in the dictionary entry is different according to each part of speech, but generally contains the following kinds of information:

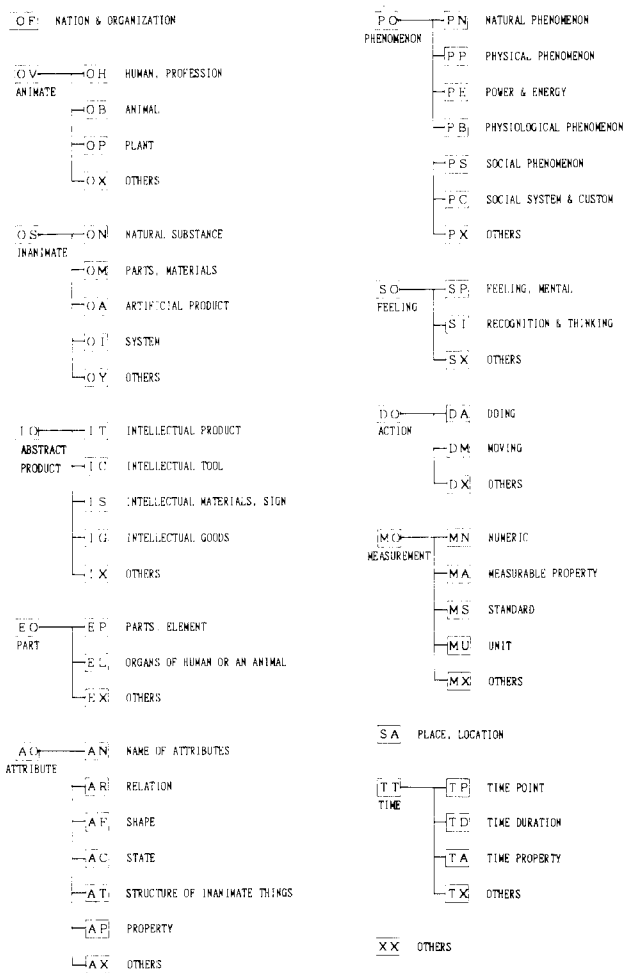
head word, number of characters of the word ending, Chinese character part, reading in Kana, variant, derivational words, related words, morphological part of speech, conjugation, prefixal information, area code, syntactic part of speech, subcategorization of part of speech, case patterns, aspect, modal, volition, semantic primitives, thesaurus code, co-occurrence information (adverb, predicative modifier), idiomatic expressions, degrees, degrees of nominality, etc.

Here one of the important pieces of information is the case pattern for verbs and nouns. We have distinguished more than 30 cases as shown in Table 1. Each case slot in a

Table 1 Case Relations Used in Japanese Dependency Structures

(1) SUBject	(17) ATTribute
(2) OBJect	(18) CAUse
(3) RECipient	(19) TOOl
(4) ORIgin	(20) MATerial
(5) PARtner	(21) COMponent
(6) OPPonent	(22) MANner
(7) TIMe	(23) CONdition
(8) Time-FRom	(24) PURpose
(9) Time-TO	(25) ROLe
(10) DURation	(26) COnTent
(11) SPAce	(27) RANge
(12) Space-FRom	(28) TOPic
(13) Space-TO	(29) VIEwpoint
(14) Space-THrough	(30) COmpaRison
(15) SOUrce	(31) ACOmpany
(16) GOAl	(32) DEGree
	(33) PREdicative

Table 2 System of Semantic Primitives for Nouns

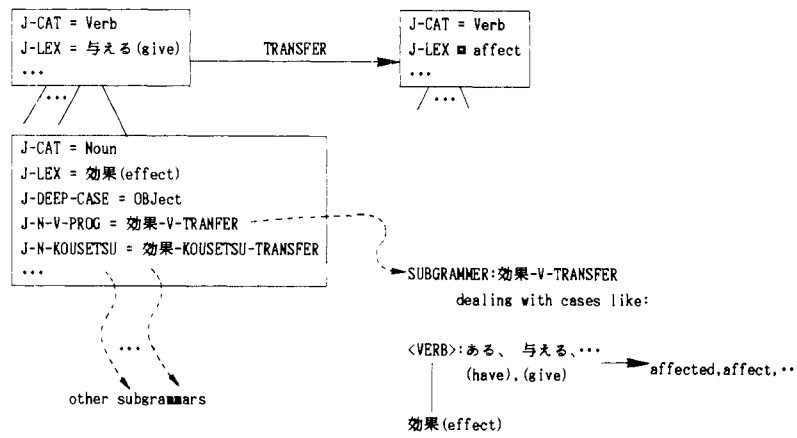


case pattern of a verb usage has semantic information about the nouns which can appear in that slot. The noun has the corresponding semantic codes in its entry. We have distinguished more than 50 semantic primitives (codes) as shown in Table 2. The noun dictionary also contains information about specific verbs which co-occur with the noun as shown in Fig. 2. The specific information of these categories is checked prior to the standard rule applications, and the default rules are applied last, to get some output which is better than nothing.

III. Lexicon-Driven Sentential Analysis

A. Case Grammar

As mentioned in Section II, the Japanese language is nearly word-order free, and many omissions of phrases



(a)

[0] Default Translation Equivalent

Japanese Noun	Translation Equivalent in English	Subject Field
<効果>	<effect>	<GENERAL>
Other Choices &their Subject Fields		

[A] Corresponding between Derivatives

Succeeding expression	Deep Case Relation	Surface English Preposition	Translation Equivalent
<のある>	<OTH>	<null>	<effective>
<のない>	<OTH>	<null>	<ineffective>

[B] Collocation with Particles

Succeeding Particle	Deep Case Relation	Surface English Preposition	Translation Equivalent
<への>	<REC>	<on>	<effect>

[C] Collocation with Miscelaneous Modifiers

Modifier	English Translation Equivalent of the Modifier

NOTE	<<arbitrary comments can be written>>

[D] Collocation with Verbs and Adjective

Japanese Predicate		ARG1	ARG2	ARG3	ARG4
<及ぼす>	J-SCM	<が>	<を>	<に>	
	J-DCR	<SUB>	<OBJ>	<REC>	
	POS		<#>		
English Predicate		ARG1	ARG2	ARG3	ARG4
<have>	E-SCM	<SUBJ>	<OBJ1>	<=on>	
	E-DCR	<OBJ>	<PAR>	<REC>	
	J-DCR	<SUB>	<OBJ>	<REC>	

Japanese Predicate		ARG1	ARG2	ARG3	ARG4
<表す>	J-SCM	<が>	<を>		
	J-DCR	<SUB>	<OBJ>		
	POS		<#>		
English Predicate		ARG1	ARG2	ARG3	ARG4
<take>	E-SCM	<SUBJ>	<OBJ1>		
	E-DCR	<AGT>	<OBJ>		
	J-DCR	<SUB>	<OBJ>		

Japanese Predicate		ARG1	ARG2	ARG3	ARG4
<ある>	J-SCM	<が>	<に>		
	J-DCR	<SUB>	<SPA>		
	POS	<DEL>			
English Predicate		ARG1	ARG2	ARG3	ARG4
<effective>	E-SCM		<SUB>		
	E-DCR		<OBJ>		
	J-DCR		<SPA>		

(*) J-SCM : Japanese Surface Case Marker
J-DCR : Japanese Deep Case Relation
E-SCM : English Surface Case Marker
E-DCR : English Deep Case Relation
POS : The position where the entry noun occurs (marked by '#')

(b)

Fig. 2. Lexicon-oriented invocation of grammar rules. (a) Activating a lexical rule for a noun "effect" from a governing verb "to give." (b) Form-oriented description of a transfer rule for a noun "effect."

Entry Word	
Number of Characters in the Inflectional Part	
Number of Characters in the Stem	
Component Words (if the Entry Word is a compound)	
Pronounciation	
Variants	
Derivatives	
Antonyms, Synonyms, and other Related Words	

Information for Morphological Processing	
Classification of Words for Morphological Processing	(mark one of the following) 名 副名 動 形 形動 副 連体 接 助 助動 接頭 接尾
Inflection Type for Verbs	(mark one of the following) 五 上一 下一 サ変 ザ変 カ変 ア カ サ タ ナ ハ マ ヤ ラ ワ
Inflection Type for Adjectives	(mark one of the following) ダナ ダナノ
Inflection Type for Auxiliaries	(mark one of the following) 形 形動 動 特
Type for Post-positional Particles	(mark one of the following) 格 接 副 並 終 準
Inflection Type for Suffixes	(mark one of the following) 体 動 形 形動
Number Code for Restriction on the Precededing Words	

Syntactic & Semantic Information		
Subject Field		
Part of Speech		
Case Frame ID		
Entry Word for Transfer Dic.		
Semantic Class based on Aspect	(mark one of the following) 瞬間/状態/継続/準状態/推移	(mark one of the following) 瞬間/状態/継続/準状態/推移
Semantic Class based on Intentionality	(mark one of the following) 意志/準意志/無意志/記述	(mark one of the following) 意志/準意志/無意志/記述
Other Semantic Classification		
Code for the thesaurus		
NOTE		

	Case Frame				Case Frame			
	SCM	DCR	SER	OBL	SCM	DCR	SER	OBL
1								
2								
3								
4								
5								
6								
7								
8								
	Collocation with adverbs				Collocation with adverbs			

NOTE

Fig. 3. Coding sheet of verbs (Japanese dictionary).

occur in a sentence. To cope with this difficulty we have to employ syntactic and semantic information simultaneously for the analysis of a sentential structure. The best grammatical framework for this is the case grammar which was first discussed by C. Fillmore in his paper in 1968. A similar idea existed in Japan long before Fillmore, and had been used to some extent. After his paper, there were many improvements and changes in the theory for the purpose of applying the idea actually to the computer analysis of the Japanese language.

The semantic primitives shown in Table 2 are used to specify what kind of nouns can be in what case slot. For example, a verb "eat" demands a noun associated with one of the semantic primitives "animal" as the agent of the verb, and a noun of the semantic code "eatable material" as the object. These case slot specifications are given for each usage of every verb in the dictionary. The analysis dictionary format for a verb is shown in Fig. 3. We have two kinds of information. One is the slots for the deep case frame, and the other for the surface cases. The former corresponds to Fillmore's case concept, while the latter corresponds to the usage of postpositions in Japanese. For example, the postposition "GA" is a surface case which corresponds to either agent or object in the deep case according to the noun before "GA." The distinction can be done by the difference of semantic primitives given to the noun as shown below:

	Surface Case	Semantic Primitives of *N*	Deep Case
eat	N—GA	eatable material	object
	N—GA	animal	agent
	N—O	thing	object

When both "-GA" and "-O" appear in a sentence, the decision can be made unambiguously, that "GA" is agent and "O" is object, without any semantic checking. Examples of the usage of "GA" in the above sense are given below:

- Boku-wa sakana-ga tabe-tai.
  I fish eat want
  (I want to eat fish.)
- Boku-ga (sakana-o) tabe-tai.
  I fish eat want
  (I want to eat fish.)
  ↓
  (emphasis)

B. Embedded Sentences

The analysis of a sentence is done by finding predicative verbs in a sentence. We look for a verb in a sentence from the left. The main verb usually comes in the final position (right end), but in the case of an embedded sentence, a verb can be in the middle of a sentence, just before a noun (N), which is the noun modified by the embedded sentence. An example is shown in Fig. 4. In this case, we have two

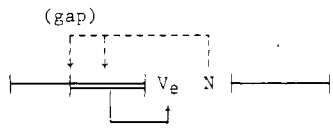


Fig. 4. Embedded sentence.

problems to solve. One is the determination of the left boundary of the embedded sentence. This is generally not so difficult except in the case of some complicated expressions. Another is the determination of the gap in the embedded sentence, that is the determination of the case of the noun N to the verb V_e in Fig. 4. This is quite difficult because the postposition of N to the verb V_e is eliminated by the embedding transformation, and we have no surface case to judge the deep case of N to V_e.

We have classified four types of embedding as follows. Distinguishing between them generally requires the semantic relationship of N to V_e, or to some other nouns in the embedded sentence.

1) A Gap in a Case: N occupies a case slot of V_e in this case. First, the case slots of V_e are semantically checked with the noun phrases in the embedded sentence. After this operation, the semantic checking of N with the remaining vacant case slots of V_e is performed to determine the modification relationship of the embedded sentence to N. Examples are given below.

•Ashi-ga -no utsukushii onna-wa toku-da.

leg beautiful woman advantageous

(legs of a woman are beautiful.)

A woman whose legs are beautiful is advantageous.

•Nedan-o shira-nai hōseki-o kat-ta.

price know not jewel buy(past)

(I) bought a jewel whose price I did not know.

3) Certain Categories of Nouns which Entail Embedded Sentences: Such structures as "at the time when...," and "in the case which...," are classified in this category, which can be regarded as an apposition. We can list the words of this category easily.

4) Particular Exceptional Expressions: Very few, but very special expressions exist which cannot be handled by either syntactic or semantic information. They must be handled by much wider contextual information.

C. Conjunctive Noun Phrase

The determination of a conjunctive structure of long noun phrases is quite difficult. There are two typically different cases. One is caused by the ambiguity of a word.

• Yosokuchi to kotonaru atai

predicted value (and / with) different value

1) predicted value and different value
2) value which is different from the predicted value.

• Watashi-wa kinou it-ta mise-ni -e mata it-ta.

I yesterday go (past) shop again go (past)

embedded sentence
agent omission
N

(I went to the shop where (I) went yesterday once more.)

• Watashi-ga kinou it-ta mise-ni kare-ga it-ta.

I yesterday go (past) shop he go (past)

embedded sentence
N

(He went to the shop where I went yesterday.)

2) A Gap in a Case Modifier: N occupies the modifier position of a noun which occupies a case slot of V_e. This sentential structure is difficult to recognize because the relation between N and the noun which is modified by it cannot be specified clearly. Frequently the relation is "part and whole," or "object and its property." Examples are given below.

The ambiguity comes from the different interpretation of "TO," which could be two different parts of speechs, a conjunction and a postposition.

Another difficulty is to know the pair of phrases in a sentence which are conjoined. For example:

• high speed electronic circuit and switching device

has many interpretations such as

(high speed electronic circuit) and (switching device)
high speed ((electronic circuit) and (switching device))

Such ambiguous structures exist both in English and Japanese. The conjunctive structures of long noun phrases are particularly difficult to analyze. A key for the analysis of this kind of noun phrase conjunction is to check the properties of head nouns in the conjunctive phrases. If the nouns have the same property, such as coincidence in number, in semantic primitives, word formation (the same postfixes, etc.), they are judged as the head nouns of conjunction. When the nouns are in a strong relation in the thesaurus system, such that "circuit" is a narrower term (NT) of "device," the reliability of the conjunction of this phrase is very high. A similar algorithm can be applied to the conjunction of prepositional phrases, predicate phrases, and so on. In the case of the conjunction of sentences, difficulties arise in the segmentation of the component sentential boundaries. Anyway, by using as much syntactic and semantic information as possible, we can almost resolve the ambiguity problem of conjunctive phrases. We have about one hundred grammatical rules of complex structures for the conjunctive phrase analysis. The conjunctive phrases which could not be analyzed by our system usually had inherent ambiguity which could not be solved, even by humans, without specific knowledge of the field discussed in the sentences.

D. Disambiguation of Multiple Parts of Speech

It is seldom that a word has multiple parts of speech in the Japanese language except for auxiliary verbs and postpositions. A noun sometimes can serve as an adverb, but the distinction can be made by syntactic environment.

In English a noun is very often used as a verb, and the disambiguation of these two cases is a very complex task. When many nouns are used in a sentence, a check should be done for every alternative interpretation for the disambiguation. This is a tremendous task. We have two approaches for the disambiguation of noun and verb. One is to check the syntactic environment of a few words' range. The other is quite sophisticated and is realized by the strong control ability of GRADE.

When the analysis process encounters an ambiguous word with two syntactic possibilities, it assumes one as a candidate, and continues the analysis of a phrase which contains the word. If the analysis is successful for the local syntactic structure of a word sequence, then the possibility for the other part of speech is neglected. If the analysis fails for the local word sequence including the ambiguous word, then the system discards all the recent partial analysis results which led to the failure, and comes back to the initial point of analysis of the ambiguous word, and restarts the analysis of the phrase with the word as an alternative part of speech. For example, in a sentence:

- He went to school after breakfast which he had at his friend's home.
- He went to school after he had breakfast at his friend's home

we have an ambiguous word "after," which can be both a conjunction and a preposition. The system performs the analysis of the word sequence to the right of "after," and gets the result as a noun/sentential phrase. Then "after" is determined as "preposition/conjunction." In the case of noun/verb distinction, the feedback process is rather complicated because it is difficult to determine what range of word sequence the system must analyze to get a success/failure information. If the range is set wrong, the result becomes wrong. The boundary of the range must be determined carefully. The worst case will be the whole sentence as the retrial range of word sequence. In the case of "radio signals analysis," we have two interpretations. "signals" can be a plural noun or a verb. This is an inherent ambiguity. In such a case the system produces only one result.

E. Other Problems

There are many other problems in the analysis of a Japanese sentence. Tense structure is quite different in English and Japanese, especially in such sentences as those having embedded sentences and subordinate sentences. Another element is the difference of modality including aspect in English and Japanese. We have detailed contrastive studies of these linguistic phenomena between English and Japanes, and the major algorithms are implemented in the program.

The analysis by case is not an easy task to do. The case frame information is described by the assumption of normal sentential structure. But the actual sentence has lots of variations, such as word order exchange, passivation, and the sentential structure changes by causative, possibility, and feasibility expressions, and so on, which are distinguished by the postpositional phrases attached to the predicates. Therefore, before applying the case frame information in a pattern matching process, the predicative part is analyzed first to know which sentential transformation must be applied. After the transformation of an input sentence to a standard form, the case frame is matched to the obtained structure with the word-order-free mode of GRADE. All these operations can be represented by GRADE's rewriting rules and the interpretation system. The whole analysis process of a Japanese sentence is shown in Fig. 5.

IV. Transfer Stage from Japanese to English

A. Annotated Dependency Structure

The intermediate representation that we adopted as a result of the analysis in our machine translation system is an annotated dependency structure. Each node has an arbitrary number of features, such as part of speech, surface case, deep case, number, tense, semantic codes, and so on. This tree representation is powerful and flexible for sophisticated grammatical and semantic checking, especially when the completeness of semantic analysis is not assured and trial-and-error improvements are required at the transfer and generation stages.

We have three conceptual levels for grammar rules in the transfer and generation phases as well as in the analysis phase.

The lowest level: default grammar rules which guarantee the output of the translation

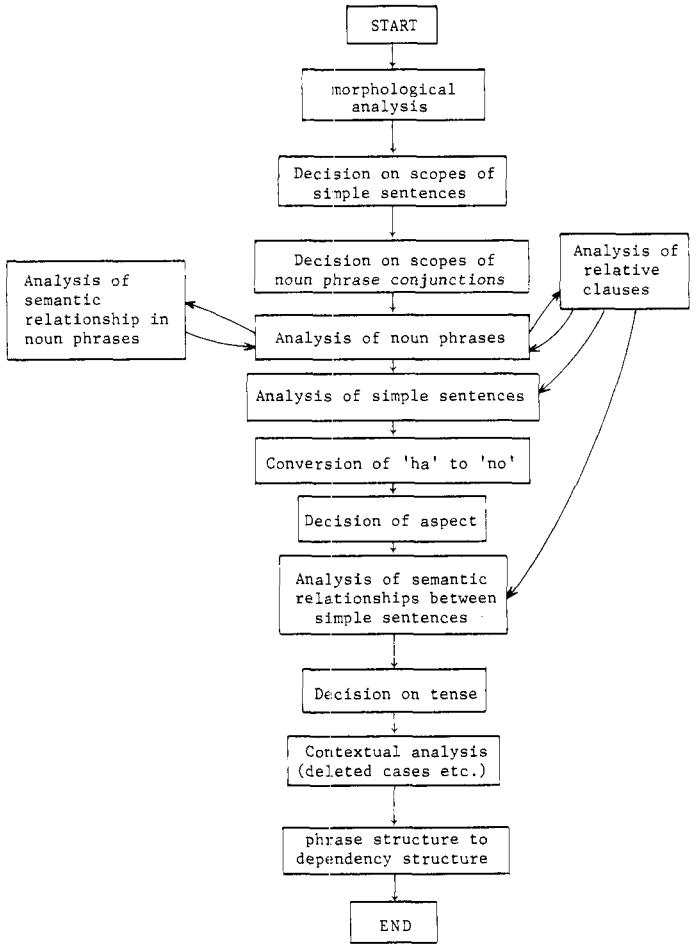


Fig. 5. Basic flow of the analysis of a Japanese sentence.

process. The quality of the translation is not assured. Rules of this level apply to those inputs for which no higher layer grammar rules are applicable.

The kernel level: main grammar rules which choose and generate a target language structure according to semantic relations among constituents which are determined in the analysis stage.

The topmost level: heuristic grammar which attempts to get elegant translation for the input. Each rule has a heuristic nature, in the sense that it is word-specific and it is applicable only to some restricted expressions which are found by a strong pattern matching function of GRADE.

The order of application of these rules is organized along the principle that "a better rule has priority."

We use the deep case dependency structure as a semantic representation. Theoretically, we can assign a unique case dependency structure to each input sentence. In practice, however, the analysis phase may find out several alternative structures because of the syntactic and semantic ambiguities. Therefore, we use, as an intermediate representation, a structure which makes it possible to annotate multiple possibilities as well as a multiple level representation. An example is shown in Fig. 6. The properties at a node are represented as a vector, so that this complex dependency structure is flexible in the sense that different interpretation rules can be applied to the structure.

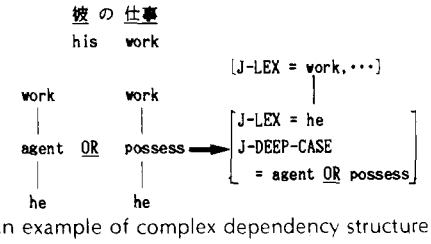


Fig. 6. An example of complex dependency structure.

Beside the ordinary grammatical rules which involve semantic checking functions, the grammar allows the reference to a lexical item in the dictionary. A lexical item contains lexical rules corresponding to its special grammatical usages and idiomatic expressions. During the transfer and generation stages, these rules are activated with the highest priority. This feature of using the lexical rules makes the system very strong and flexible for dealing with exceptional expressions. The improvement of translation quality can be achieved progressively by adding lexical rules as

well as linquistic information and word usages in the dictionary entries.

Some heuristic rules are activated just after the standard analysis of a Japanese sentence is finished, to obtain a more neutral (or target language oriented) analyzed structure by the pre-transfer loop. Semantic and pragmatic interpretations are done in the pre-transfer loop. The more the heuristic rules are applied at this stage, the better will be the result. Figs. 7 and 8 are some examples.

B. Word Selection in the Target Language

Word selection in the target language is a big problem in machine translation. There are varieties of choices of translation for a word in the source language. The main principles adopted in our system are, as follows:

1) Field restriction by using a field code, such as electrical engineering, nuclear science, medicine, and so on.

2) A semantic code attached to a word in the analysis phase is used for the selection of a proper target language or phrase.

3) The sentential structure in the vicinity of a word to be translated is effective for the determination of a proper word or phrase in the target language.

Table 3 shows examples of a part of the verb transfer dictionary. The selection of an English verb is done by the semantic categories on nouns related to the verb. In the table, the number *i* attached to verbs like form-1, produce-2, is the *i*th usage of the verb. When the semantic information attached to nouns is not available, the column indicated by ϕ is applied to produce a default translation.

In most cases, we can use a fixed format for describing a word selection rule for lexical items. We developed a number of dictionary formats specially designed for the ease of dictionary input by computer-naive expert translators.

The expressive power of format-oriented description is, however, insufficient for a number of common verbs such as "SURU" (make, do, perform, ...), and "NARU" (become, consist of, provide, ...) etc. In such cases, we can represent transfer rules directly by tree structures. An example is

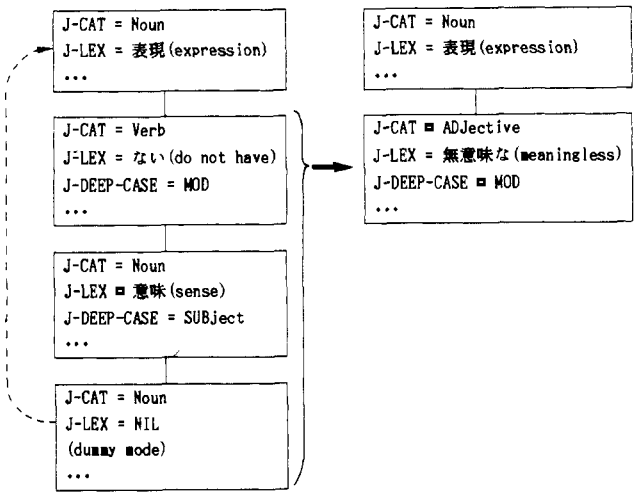


Fig. 7. An example of heuristic rule used in the pre-transfer loop.

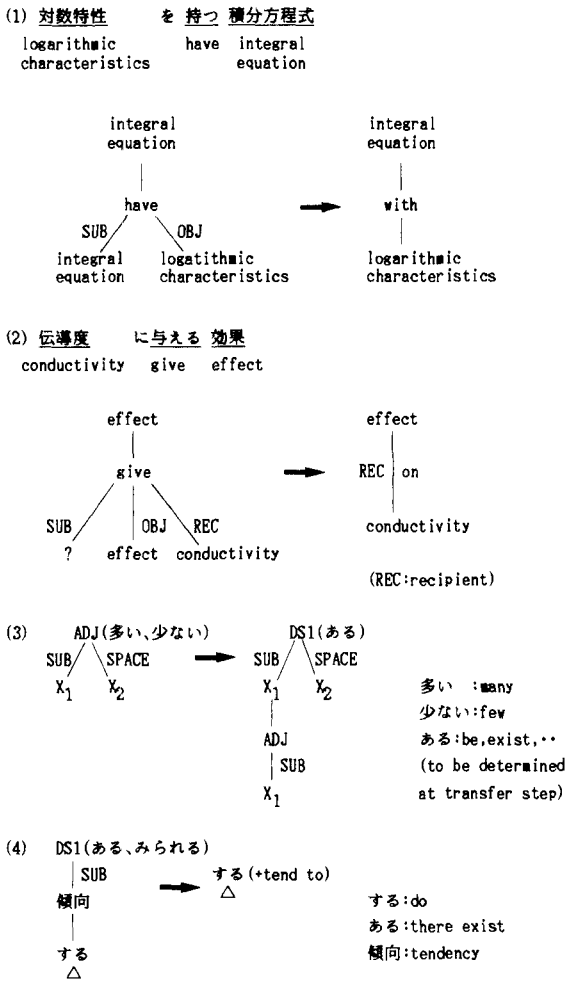


Fig. 8. Examples of pre-transfer rules.

Table 3 Word Selection in Target Language by Using Semantic Markers

生ずる SYO ZURU	Xが生ずる X GA SYO ZURU		non-living substance structure	form-1	form X (obj)
			social phenomena	take place	X take place
		X	action, deed, movement reaction	occur-1	X occur
			standard, property state, condition relation	arise-1	X arise
			ϕ	produce-2	produce X
	XがYを生ずる X GA Y O SYO ZURU		non-living substance structure	form-1	X form Y
		Y	phenomena, action	cause-1	X cause Y
			ϕ	produce-2	X produce Y
上げる AGERU	XがYを上げる X GA Y O AGERU		property	improve-1	X improve Y
		Y	measure	increase-2	X increase Y
			ϕ	raise-1	X raise Y

Semantic marker for X/Y

shown in Fig. 9. Every usage of a verb is listed with their corresponding English sentential structures and semantic conditions.

This mechanism of the transfer stage bridges the gap between Japanese and English expressions. There are, however, still many odd structures after this stage, and we have to adjust the English internal representations into more natural ones by the post-transfer loop. An example is given

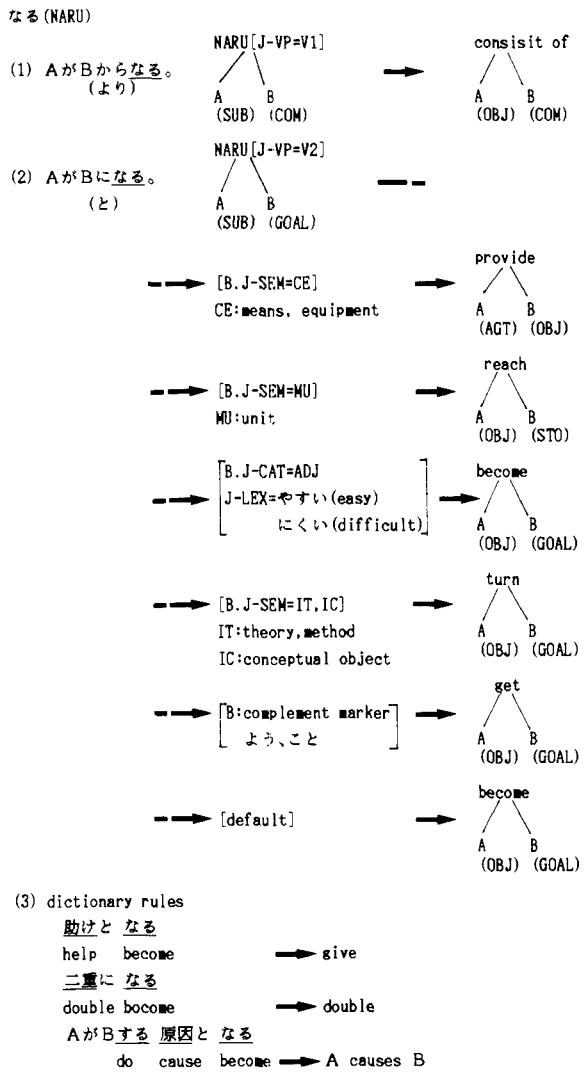


Fig. 9. An example of dictionary transfer rules of popular verbs.

in Fig. 10, where a Japanese factitive verb, SASERU, is first transferred to an English word "make," and then a structural change is made to eliminate it, to have a simpler and more direct expression. Another example is shown in Fig. 11, where a term corresponding to "the number of" is inserted in between "increase" and "car," because "the number" can "increase," but "a car" does not "increase."

Postpositions in Japanese generally express the case slots

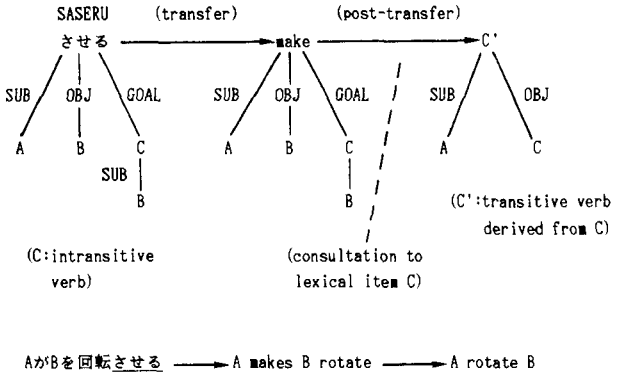


Fig. 10. An example of post-transfer rule application.

for verbs. A postposition, however, has different usages, and the determination of English prepositions for each postposition is quite difficult. It also depends on the verb which governs the noun phrase having that postposition.

Table 4 illustrates a part of a default table for determining deep and surface case labels when no higher level rule

Table 4 Default Rule for Assigning a Case Label of English to a Japanese Preposition "ni" (The particular selection of a preposition is described in the entries of English verbs which govern the whole structure.)

J-SURFACE-CASE	J-DEEP-CASE	E-DEEP-CASE	Default Preposition
に(ni)	RECipient	REC, BENeficiary	to (REC-to, BEN-for)
	ORIgin	ORI	from
	PARticipant	PAR	with
	TIMe	Time-AT	in
	ROLe	ROL	as
	GOAl	GOA	to
	...	...	...

applies. Tables of this sort are defined for all case combinations. In this way, we guarantee at least a literal translation to be given for an input. A better choice of a preposition depends on the usage of a verb, so that every usage of a preposition for a particular English verb is written in the lexical entry of the verb, and is used in the selection of the preposition.

V. Determination of Global Sentential Structures in the Target Language

Global sentential structures of Japanese and English are quite different, and correspondingly the deep structure of a Japanese sentence is not the same as that of English. The

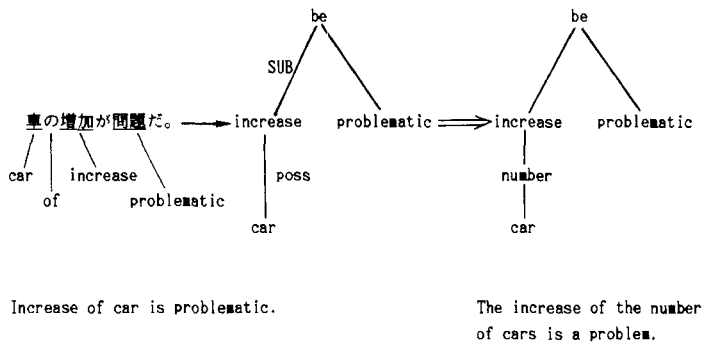


Fig. 11. An example of structural change at the post-transfer loop.

fundamental differences between the Japanese internal representation and those of English are absorbed at the (pre-, post-) transfer stages. But as the stage of English generation, some structural transformations are still required in the case of such expressions as embedded sentential structures and complex sentential structures.

An example of a structural transformation is shown in Fig. 12. The relative clause "why..." is generated after the structural transformation.

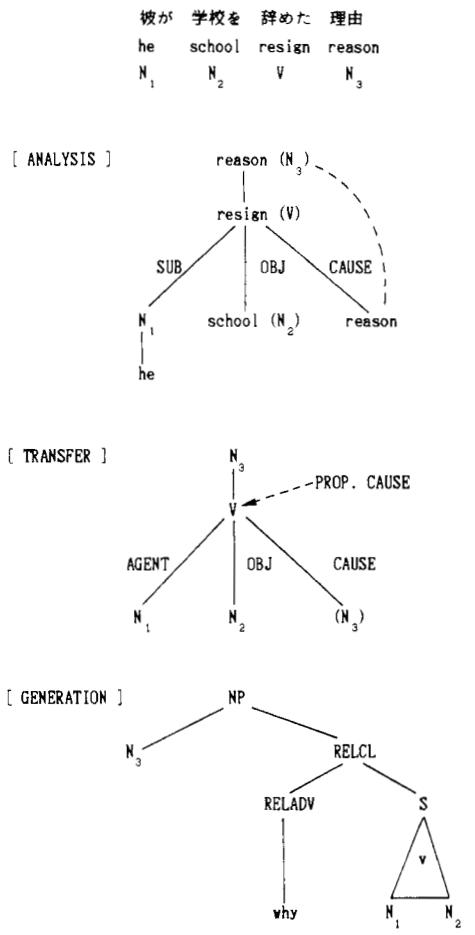


Fig. 12. Structural transformation of an embedded sentence of type 3.

Connection of two sentences in compound and complex sentences is done according to the information shown in Table 5. An example is given in Fig. 13.

The process of sentence generation in English is as follows. After the transfer is done from the Japanese deep dependency structure to the English one, conversion is done from the English deep dependency structure to a phrase structure tree with all the surface words attached to the tree. The processes explained above are involved at this generation stage. The conversion is performed step by step from the root node of the dependency tree to the leaf as a top-down process. Therefore, when a governing verb demands a noun phrase expression or a to-infinitive expression in its dependent phrase, which may be a verbal phrase or a noun phrase, a proper structural change of the phrase must be performed. Noun to verb transformation, and noun to adjective transformation are often required due to the difference of expressions in Japanese and English. When we

Table 5 Correspondence of Sentential Connectives

Japanese Sentential Connective	Deep Case	English Sentential Connective
Renyo	tool	by -ing...
(-shi) te	tool	by -ing...
Renyo	cause	because...
(-shi) te	〃	〃
-tame	〃	〃
-node	〃	〃
-kara	〃	〃
-to	time	when...
-toki	〃	〃
-te	〃	〃
-tame	purpose	so-that-may
-noni	〃	〃
-you	〃	〃
-you	manner	as if
-kotonaku	〃	without -ing...
-nagara	accompany	while -ing...
-ba	circumstance	when...
...	...	...

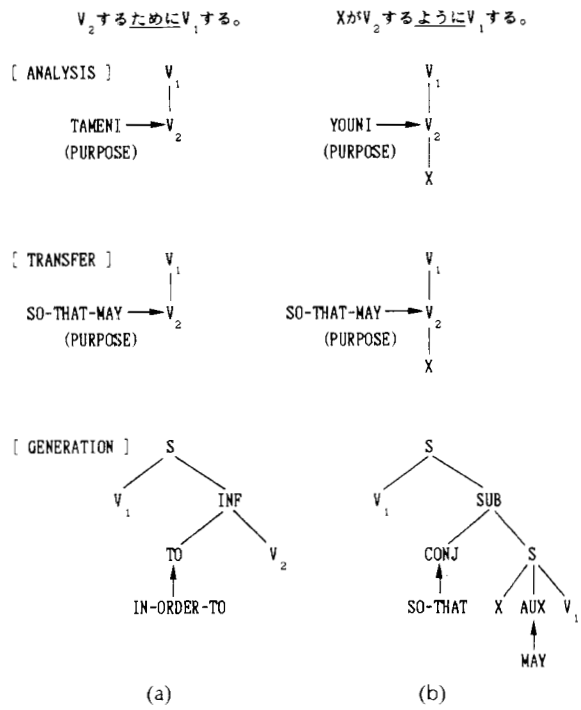


Fig. 13. Structural transformation of an embedded sentence.

cannot find a verb or a noun corresponding to a noun or a verb, respectively, in the dictionary, we make reference to a synonym word and see if it has a verb or a noun derived from the word. An example is shown in Fig. 14. The generation goes down from the root node until all the leaf nodes are converted to a phrase structure tree.

After this process of phrase structure generation, some sentential transformations are performed such as the following:

i) When an agent is absent, the passivation transformation is applied.

ii) When the agent and object are both missing, the predicative verb is nominalized and placed as the subject, and such verb phrases as "is made," and "is performed" are supplemented. An example is shown in Fig. 15.

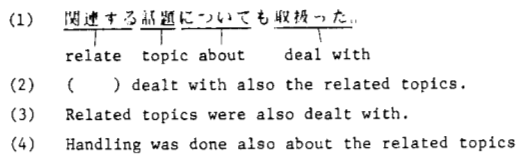


Fig. 14. An example of transformations in the related topics. Original Japanese (1) has no subject. (2) is a literal translation. (3) is a passive form of (2). Another direction is the nominalization of the verb phrase "deal with." Since the dictionary entry of "deal with" has no nominal form, the system sees the entry of synonym of "deal with" and finds out the word "handle," and the nominalization "handling." This noun is then put on at the subject position and a sentence (4) is generated with a nominal verb "be done."

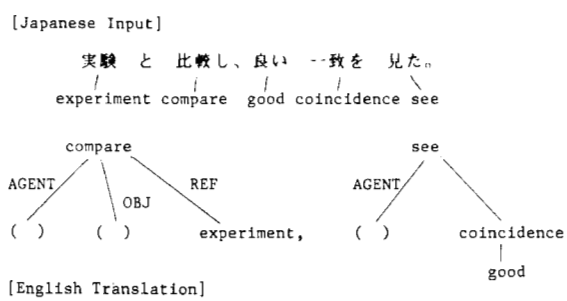


Fig. 15. An example of a sentence without subject and object. Nominalization of "compare" is performed; "see" is changed to "obtain" by the consultation of the lexical entry of "coincidence."

iii) When a subject phrase is a big tree, the anticipatory subject "it" is introduced to avoid the top-heavy structure.

iv) Pronominalization of the same subject nouns and the change of pronouns including deletion are performed. An example is shown in Fig. 16.

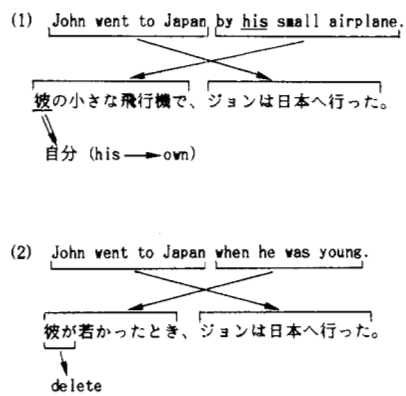


Fig. 16. Translation of pronouns.

v) Duplication of a head noun in the conjunctive noun phrase is eliminated, such as "uniform component and nonuniform component" → "uniform and nonuniform components."

There are many such structural transformations.

Another big structural transformation required comes from the essential difference between a DO-language (such as English) and a BE-language (such as Japanese). In English, the case slots such as tools, cause/reason, and some others come to the subject position very often, while in Japanese such expressions are never used. The transformation of this kind is incorporated in the generation grammar such as that shown in Fig. 17, and it produces more natural English expressions. This stylistic transformation is very important in machine translation between Japanese and English.

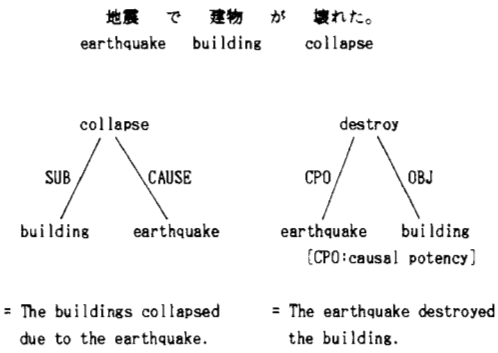


Fig. 17. An example of structural transformation in the generation phase.

VI. Some Problems which Require Contextual Information and Knowledge

There are many problems which require contextual information and common sense knowledge for better machine translation. Some of them, which are not yet incorporated into our machine translation system, are pointed out in the following.

1) Anaphoric and cataphoric expressions are often used in ordinary texts, although they did not appear in the texts of abstracts of scientific and technical papers which were to be translated in our system. For the anaphoric usage we have a method which determines the referents reasonably well. We stack up in what we call the anaphora stack the head nouns of the obligatory case slots of the sentences being analyzed. When a pronoun appears we take out the noun at the top of the stack and make semantic checking of the whole sentential meaning by replacing the pronoun with the noun. That is, we check whether the noun can occupy the slot of the pronoun in the sentence. If not, we take out the noun next to the top of the anaphora stack, and check the replaceability of the pronoun in the same way. This operation is performed until nouns of a small number of previous sentences are referenced. The use of the stack is to check in sequence the nouns nearer to the pronoun.

In the case of the pronoun "you," a distinction should be made between singular and plural by the analysis of the sentences where "you" appears. In the case of "they" we have to distinguish the semantic codes, human and nonhuman, according to the meaning of the sentence in which "they" is contained. We can make reference to the verb which governs "they." If the verb has a semantic code of human action, such as, think, write, and say, "they" must be human. In the case of "their X" we look for objects whose parts are X, such as "cars, and their wheels." Or, we look for an object which has some relation to X, such as "students, and their answers to a teacher" (students and teacher are related terms).

If a proper noun is not found by these processes, we will

check the possibility of a cataphoric usage. But before going on to the cataphoric usage check, we have to see several other possibilities. One is the possibility that a part or the whole of a sentence can be the encapsulation of a concept, which is referred to by a pronoun. There is sometimes a case when a concept is explained by a series of sentences. In such a case, we have to have a semantic network built up by the information contained in the set of sentences, and whose top node is the concept explained by the set. In this case, the topical theme of the set of sentences is generally the concept which is referred to by the pronouns in these sentences, and it must be registered in the top of the anaphora stack. This means that a topical concept of a sentence must be pushed down as late as possible, or the concept should come up to the top of the stack as a bubble comes up the surface of a glass of water.

A pronoun in the cataphoric usage often appears in a sentence which starts a paragraph. In that case, our stack for the anaphoric reference is vacant, and we have inevitably to go to check a cataphoric reference.

Anticipatory "it" can be detected by checking the syntactic structure such as "it is ... that ..." in a sentence.

2) The determination of the function of articles are quite difficult. We have no reliable way of attaching "a," "the," or nothing to a noun in a sentence. As for the particular usage of "the" which refers to an object in the previous phrase like,

$$\ldots \text{a } X \ldots . \qquad \ldots \text{the } Y \ldots .$$

we can check whether $X = Y$, or X and Y have the relation of synonymity. If there exists such a relation between X and Y, then we can translate this particular usage of "the" as "SONO" in Japanese. Thesaurus information is very useful in this respect. In the other usages of the definite article, we can just neglect them in translation into Japanese. In translation from Japanese into English, however, we have great difficulty in inserting definite and indefinite articles in their proper places in the generated English.

3) Ellipsis appears very often in Japanese sentences. In our machine translation system, omitted words are not recovered, this is avoided by using some particular English sentential styles which do not require these omitted words. Some of these mechanisms are explained in the previous section. We can use the case slot information to infer proper words for the omission. A case slot has a set of semantic codes, and a noun with any of the semantic codes is searched for in the anaphora stack. The noun with this condition is regarded as the reference for the ellipsis. However, this process is not very reliable, and we do not like to use this mechanism very often.

VII. The Japanese Government Machine Translation Project

A. Outline of the Project

The Japanese Government decided to start the machine translation project to realize a quick information transfer of scientific and technical documents written in Japanese and English. The project is funded under a grant from the Agency of Science and Technology through the Special Coordination Funds for the Promotion of Science and Technology, and was started in the fiscal year 1982. The formal title of the project is "Research on Fast Information Services between Japanese and English for Scientific and Engineering Literature." The purpose of the project is to demonstrate the feasibility of machine translation of abstracts of scientific and engineering papers written in these languages, and as a result, to establish a fast information exchange system for these papers. The project term was four years ending in 1986.

The project was conducted in close cooperation between four organizations. At Kyoto University, we had the responsibility of developing the software system for the core part of the machine translation process (grammar writing system and execution system), grammar systems for analysis, transfer, and synthesis, detailed specification of what information is to be written in the word dictionaries (all the parts of speech in the analysis, transfer, and generation dictionaries), and the working manuals for constructing these dictionaries. The details have been given in the previous sections and in some other documents. The Electrotechnical Laboratories (ETL) were responsible for the machine translation text input and output, morphological analysis and synthesis, and the construction of the verb and adjective dictionaries based on the working manuals prepared at Kyoto. The Japan Information Center for Science and Technology (JICST) was in charge of the noun dictionary and the compiling of special technical terms in scientific and technical fields. The Research Information Processing System (RIPS) under the Agency of Engineering Technology was responsible for completing the machine translation system including the man–machine interfaces to the system developed at Kyoto, which allow pre- and post-editing, access to grammar rules, and dictionary maintenance.

The project is not primarily concerned with the development of a final practical system; this will be developed by private industry using the results of the project.

The Japanese source texts being used are abstracts of scientific and technical papers published in the monthly journal, *A Current Bibliography of Science and Technology* published by JICST. At present, the project is only processing texts in the electronics, electrical engineering, and computer science fields. English source texts will be abstracts from INSPEC in these fields. The sentence structures used in abstracts tend to be complex compared to ordinary sentences, with long nominal compounds, noun-phrase conjunctions, mathematical and physical formulas, long embedded sentences, and so on. The analysis and translation of this type of sentence structure is far more difficult than that of ordinary sentence patterns. However, we have not included a pre-editing stage because we wanted to find the ultimate limitations on handling this type of complex sentence structure.

B. Evaluation of Translation Quality

The following two aspects have been adopted to evaluate the quality of the sentences translated by the project system. They are to some extent independent indicators.

1) Intelligibility: An evaluation of the extent to which the translated text can be understood by a native speaker of the target language. In Japanese to English translation, we evaluate the extent to which an average British or American

reader can understand the output without any reference made to the Japanese original.

2) Accuracy: The degree to which the translated text conveys the meaning of the original text as well as a measure of the amount of difference between the input and output sentences are evaluated. The evaluation is done by Japanese translators specializing in Japanese to English translation.

Evaluation of intelligibility is done based on a scale of five; the categories are described below:

1) The meaning of the sentence is clear, and there are no questions needed. Grammar, word usage, and style are all appropriate, and no rewriting is needed.

2) The meaning of the sentence is clear, but there are some problems in grammar, word usage, and/or style, making the overall quality less than in 1.

3) The basic thrust of the sentence is clear, but the evaluator is not sure of some parts because of grammar and word usage problems. The problems cannot be resolved by any set procedure; the evaluator needs the assistance of a Japanese evaluator to clarify the meaning of those parts in the Japanese original.

4) The sentence contains many grammatical and word usage problems, and the evaluator can only guess at the meaning after careful study, if at all. The quickest solution will be a retranslation of the Japanese sentence because too many revisions would be needed.

5) The sentence cannot be understood at all. No amount of effort will produce any meaning.

As the evaluation number increases on the above scale from 1 to 5, intelligibility decreases. The evaluator uses the above scale to evaluate the output sentence without any reference to the Japanese original in the first place. When the output sentence contains untranslated words in Japanese, the English translation of those words is provided by a Japanese rewriter before the evaluation. This evaluation work has been carried out to date by one British and one American evaluator, neither of whom has the ability to read or evaluate Japanese. Both evaluators have one year's experience in proof reading and checking translations of general scientific and technical literature, but neither have specialized knowledge in the field of electrical engineering, which has been used for all the input material up to now.

Accuracy is evaluated on a scale of 0 to 6; that is, seven categories.

0) The context of the input sentence is faithfully conveyed to the output sentence. The translated sentence is clear to a native speaker and no rewriting is needed.

1) The context of the input sentence is faithfully conveyed to the output sentence, and can be clearly understood by a native speaker, but some rewriting is needed. The sentence can be corrected by a native speaking rewriter without referring to the original text. No Japanese language assistance is needed.

2) The context of the input sentence is faithfully conveyed to the output sentence, but some changes are needed in word order.

3) While the content of the input sentence is generally conveyed faithfully to the output sentence, there are some problems with things like relationships between phrases and expressions, and with tense, voice, plurals, and the positions of adverbs.

4) The content of the input sentence is not adequately conveyed to the output sentence. Some expressions are missing, and there are problems with the relationships between clauses, phrases and clauses, or sentence elements.

5) The content of the input sentence is not conveyed to the output sentence. Clauses and phrases are missing.

6) The content of the input sentence is not conveyed at all. The output is not a proper sentence; subjects and predicates are missing. In noun phrases, the head noun is missing, or a clause or phrase acting as a verb and modifying a noun is missing.

As the evaluation number increases on the above scale 0 to 6, the accuracy decreases. This part of the evaluation has been done by four Japanese translators, each of whom have one or two years experience in Japanese to English translation. The whole evaluation process was monitored by a Japanese translation specialist with extensive experience in translation work.

We describe here the results of the evaluation of the translation of 1682 sentences taken from the monthly journal, *A Current Bibliography of Science and Technology* published by JICST. Of these, 791 were the ones which are often referred to for the development of the analysis grammar, and the remaining 891 were added as test material this time. All the sentences were given to the machine translation system with no pre-editing. The 791 sentences forming the first group were originally selected out of a 1000 after eliminating 120 which contained ungrammatical Japanese expression, and a further 90 which contained long mathematical or chemical formulas. The deletion of the latter was because, in the early stages, the analysis grammar to deal with formulas had not been completed. The second group of sentences were all these which were in the abstract without any such selection.

Tables 6 and 7 present the evaluation results for intelligi-

Table 6 Evaluation Results for the First Group of 791 Abstracts

	Accuracy									
Intelligibility	0	1	2	3	4	5	6	Defective	Total	Percentage of Total (%)
1	98	0	9	4	1	2	2	0	116	14.7
2	0	186	8	28	22	11	4	0	259	32.7
3	0	1	135	45	42	17	16	3	259	32.7
4	0	0	20	19	25	10	24	1	99	12.6
5	0	0	5	2	8	6	36	1	58	7.3
Total	98	187	177	98	98	46	82	5	791	
Percentage of total (%)	12.4	23.6	22.4	12.4	12.4	5.8	10.4	0.6		

Table 7 Evaluation Results for the Second Group of 891 Abstracts

Intelligibility	Accuracy 0	1	2	3	4	5	6	Defective	Total	Percentage of Total (%)
1	61	0	7	5	0	3	7	0	83	9.3
2	0	142	22	27	8	13	9	0	221	24.8
3	0	0	138	68	44	26	17	4	297	33.3
4	0	0	10	24	35	16	37	4	126	14.1
5	0	0	0	1	6	7	149	1	164	18.4
Total	61	142	177	125	93	65	219	9	891	
Percentage of total (%)	6.8	15.8	19.9	14.0	10.4	7.3	24.6	1.0		

bility and accuracy for the two groups of abstracts. Tables 8 and 9 give a comparison of the two groups. As the system was not tuned to the sentences in the second group, there were many unknown grammatical structures and missing words in the dictionary, which made the evaluation result worse than that in the first group.

Table 8 Comparison Between First and Second Groups for Intelligibility

Intelligibility	First Group 791 (%)	Second Group 891 (%)
1	14.7	9.3
2	32.7	24.8
3	32.7	33.3
4	12.6	14.1
5	7.3	18.4

Table 9 Comparison Between First and Second Groups for Accuracy

Accuracy	First Group 791 (%)	Second Group 891 (%)
0	12.4	6.8
1	23.6	15.8
2	22.4	19.9
3	12.4	14.0
4	12.4	10.4
5	5.8	7.3
6	10.4	24.6
Defective	0.6	1.0

As these tables show, when the accuracy of translation goes down, so, too, does the intelligibility. We did not find any examples of intelligibility being low when accuracy was high, but we did find a reasonable number of cases where the translation accuracy was evaluated as low, but the intelligibility was rated high. Table 10 lists typical sample sentences for each evaluation type.

Just as there are no clear and objective criteria for evaluating the quality of Japanese to English translations done by humans, standard criteria for judging the results of machine translations have yet to be established. The evaluation methods proposed here are still in the trial stage, and much more refining and improving is still needed.

The translation quality and the amount of post-editing needed is closely related to the quality and nature of the original text. It is quite natural to expect that simple sentences can be translated accurately and intelligibly. We need to develop some way to evaluate the degree of difficulty of the original text along with the translation evaluation. Only within the wider context can accuracy and intelligibility be meaningfully discussed.

The JICST abstracts used in this project were written primarily with the aim of condensing as much information as possible into a few sentences. This means that there are many long sentences, many of which are not very correct Japanese from a linguistic point of view. This is one obvious factor contributing to the poor evaluation results shown in Tables 6–9.

Evaluation of the quality of machine-translated sentences is closely linked to the way in which the machine translation output is to be used; hence, to the ease with which post-editing can be done. Only a minimum of post-editing will suffice to transfer the technical meaning in the original to the specialist in a particular field for the purpose of information service. However, when the translated text is for publication in a wider circle (for example, technical manuals), style and naturalness of sentential expressions become more important as does the exact meaning. Depending on these situations, the yardstick for intelligibility will change as well.

VIII. Conclusion

Machine translation does not necessarily require very deep understanding of the sentential meaning as is believed in the AI circles. To achieve deep understanding we have to have very sophisticated mechanisms of inference with a huge amount of general knowledge of the real world. Still, we cannot guarantee correct inferencing by machine. For example, we have often a phrase like

SHOGAIKOKU	TONO	BŌEKIMASATSU	NI	JYORYOKU	SURU
foreign countries	with	conflict in trade	help		do

which is a kind of abbreviated expression. We have to recover a much more precise Japanese expression to get a proper English translation. If the mental attitude of the speaker of the sentence is in favor of the improvement of foreign relations, the insertion of KAISHŌ (resolving) should be done as

> (We) help resolve the conflict in trade with foreign countries.

But if the speaker's attitude is quite contrary, the insertion

Table 10 Translations and Their Evaluation Results

Intelligibility = 1

大気中の屈折率については慣性区間内の全規模のゆらぎに対し成立する矛盾を説明した。

the contradiction is explained which is established to 全規模だ fluctuations in inertial sections about refractive indexes in the atmosphere.

等価電流分布と励起表面波振幅により任意点での電磁場を決定できる。

equivalent current distribution and excitation surface wave amplitude can determine electromagnetic fields at given points.

この厳密なアプローチは，導電材料中の渦電流分布の正確な決定に導く。

this strict approach results in accurate determination of the eddy current distribution in conductive materials.

Intelligibility = 2

熱放射強度をこの表面による平面波散乱問題の解から決定する。

solutions for the plane wave scattering problem by this surface determine the thermal radiation intensity.

中性点接地は人体に対する許容接触電圧値と安全作用時間が保証されるときには《間接》接触電圧の確実な保護手段となる。

the neutral point grounding provides reliable means of protection of indirect contact voltages when permissible contact voltage value and safe operating time to human bodies are guaranteed.

半径の異なる二つの円柱導体による平面電磁波散乱（HおよびE波）の解を無限次元連立方程式を用いて厳密に求めた。

solutions of the plane scattering of electromagnetic wave (the h⁻/e⁻ wave) by two cylindrical conductors with different radii were strictly found by using the infinite order simultaneous equations.

Intelligibility = 3

米国ボストンのＮｅｗ　Ｅｎｇｌａｎｄ電話会社において，マイクロコンピュータをベースとしたエネルギー管理制御システムを設備し，エネルギー原価の低減に成功した事例を紹介。

the cases are introduced in which the success is made in the reduction of energy costs by equipping the energy management and control systems based on microcomputers in new england telephone company of the united states boston.

システム工学，情報工学，ソフトウエア工学など新しい技術の急速な成長により，ＩＥＥおよび電気工学教育の全パターンでの重点が移動した。

rapid growth of new technology such as systems engineering, information engineering and software engineering moved the emphasis of iee and electrical engineering education in all pattern.

電子産業ではプラスチックの利用が増大し，効率的かつ高信頼度の組立技術が重要となっているが，その一つである超音波溶接の装置と工程について解説。

the utilization of plastics has been increasing by electronic industries, and efficient and reliable assembly techniques become important, and the explanation is made regarding , as an example, equipment and processes of the ultrasonic welding.

Table 10 (*Continued*)

Intelligibility = 4

線形理論が適用不能の場合の放射場方程式では新変数が現れ，変数空間の拡張により方程式系は閉じる。

by the expansion of variable spaces linear theories new variables appear by the equations of a radiation field which can not be applied case, and systems of equation are closed.

非対称でノントリビアルな振動の伝搬を検出し，金属とフェライト間に誘電体がある時には，ギャップモード型のノントリビアルな振動に当たる分岐を含む事が分った。

the propagation of nontrivial asymmetrical vibrations is detected, and it is found that the junctions are included which are corresponding to vibrations metal ferrite vibration gap mode gap mode vibration.

Ｋｕｂｏの順序指数（Ｊ．Ｐｈｙｓ．Ｓｏｃ．Ｊａｐａｎ，１９６２，１７）の自然な一般化である順序付けられたはん関数を導入。

the functionals are introduced which are natural generalization of the ordinal exponent j.phys.soc. japan by the kubo and are put in order.

Intelligibility = 5

より安定な新しい計算法を主として理論解と比較した結果ならびに安定性の検討結果を報告。

results of study of results and stability more comparing new calculation method calculation methods to theoretical solutions mainly in the stability are reported.

本方法は変換器形シンセシスのみならず，現在の変換器特性に関する摂動効果解析にも利用できる。

this method can be utilized not only for converter-type syntheses but also for analyses of the perturbation effect about present converter characteristics.

気体が放電間隔を飛行する時間に近いパルス周期のとき，気体流をパルス周期非自続放電で有効に励起できることを示した。

gas discharge spacing time time pulse period pulse period time gaseous flow non-self-sustaining discharge pulse period non-self-sustaining discharge fact.

Accuracy = 0

ランダム不均質成層媒質からの周期変調信号の反射特性を研究。

reflection characteristics of periodically modulated signals from media of random nonhomogeneous layers are studied.

小規模不均質性をもつ平面不均質媒質からの空間変調波の反射を研究。

the reflection of space modulated waves from the plane nonhomogeneous medium with small-scale heterogeneity is studied.

電気計測法，データ処理，自動化機器の進歩で自動化船が増加した。

advances in electrical measurement, data processing and automatic equipment increased the number of automated ships.

Table 10 (*Continued*)

Accuracy = 1

多層配線板製造技術に関して，多層配線板材料，積層・接着，穴加工およびパターン形成の各項目を解説。

each item of the multilayer printed circuit board material, the lamination and the adhesion, the hole drilling and the pattern formation is explained about the multilayer printed circuit board production technology.

プリント回路の自動はんだ付け機で良質のはんだ付けを行うために必要な注意事項を列記した。

the precautions are enumerated which are necessary in order to carry out 良質だ soldering by automatic soldering machines of printed circuits.

凹凸のある平面境界をもつ導電性半空間からの熱放射をＬｅｏｎｔｖｉｃｈの近似境界条件を使う平均場法で研究。

the thermal radiation from electrically conductive half spaces with rugged plane boundaries is studied by the mean field methods using approximate boundary conditions of leontvich.

Accuracy = 2

システム工学，情報工学，ソフトウエア工学など新しい技術の急速な成長により，ＩＥＥおよび電気工学教育の全パターンでの重点が移動した。

rapid growth of new technology such as systems engineering, information engineering and software engineering moved the emphasis of iee and electrical engineering education in all pattern.

米国ボストンのＮｅｗ　Ｅｎｇｌａｎｄ電話会社において，マイクロコンピュータをベースとしたエネルギー管理制御システムを設備し，エネルギー原価の低減に成功した事例を紹介。

the cases are introduced in which the success is made in the reduction of energy costs by equipping the energy management and control systems based on microcomputers in new england telephone company of the united states boston.

保護対策として電気固有抵抗の低減の重要性を指摘し，その静電電圧計での計測を述べる。

the importance of the reduction of the electrical resistivity is pointed out as protective measures, and the measurement by the electrostatic voltmeters is described.

Accuracy = 3

家電機器は省エネルギー化が進み，また輸出が増えている。

as for household electrical appliances, the energy saving has been proceeding, and the export has been increasing as for household electrical appliances.

製品から得られる各種ベネフィットの総和についての購買者の効用を測定する１つの方法にコンジョイント測定法がある。

the conjoint measurement is one piece method for measuring the utility about totals of each benefit which can be obtained from products of buyers.

散乱特性長と回折特性長の２パラメタの存在により本質的に異なる２つの散乱状態が生ずる。

two scattering states essentially different by the existence of scattering characteristic length and diffraction characteristic lengths of two parameters arise.

Table 10 (*Continued*)

Accuracy = 4

スケーリング関係と表の精度は基礎定数に対する値を用いて１０５分の１のレベルまでテストした。

the accuracy of tables was tested up to levels of 1 105 by using values to scaling relationship and basic constants.

任意の電子のエネルギー，波長，および観測の角度に対するシンクロトロン放射の密度を予測するためにスケーリング関係を用いることができる表を作成した。

the tables were made of which scaling relationship could be used in order to predict the density of the synchrotron radiation of the energy wavelength of arbitrary electrons and observations to angles.

次に双極子に対称軸まわりの大きな初期スピンを持たせ，運動方程式を導き三個の積分を見出し軌道を構成する。

next, dipoles have large initial spins of around symmetric axises, motion equations are derived, three pieces of integral are found, and orbits are formed.

Accuracy = 5

波面反転する理想反射鏡の前の不均質媒質からの多重逆散乱のときでも鏡は不均質性が反射波に与える影響を完全に補償する。

mirrors completely compensate effects of the heterogeneity in the multiple reverse scattering in front of the ideal reflection mirrors inverted on wave surface from nonhomogeneous media on reflected waves also.

ＤＣ解の一意性，自励振動回路網のある与えられた領域での漸近安定性を論ずることができる。

the asymptotic stability in the regions by which there is an unique self-excited vibration network of the dc solution and for giving the unique self-excited vibration network of the dc solution can be discussed.

解の安定性が良いこと，数値的条件が良いこと，解の二次収束過程によって繰返し計算の回数が少くてすむことがこの方法の特長である。

characteristics of this method are that the stability of solutions is good, that numerical conditions are good and 済む ing thing.

Accuracy = 6

荷電粒子の運動は準静的共振モードＨ００の周波数の近くに周波数スペクトルが集中した回折放射バーストをともなうことを明らかにした。

it was done that the motion of charged particles accompanies the diffractive radiation bursts which frequency spectra concentrated in the neighborhood of frequencies of quasi-static resonance mode h 00 be obvious.

この方法では伝搬方向を指定してそれを分離でき，一次元方程式を多次元の場合に容易に拡張できる。

it can be separated by designating directions of propagation by this method, and first-order equations can be easily expanded to multidimensional cases.

無限長軸電流を囲んでいる，有限長かつ有限厚さの導電円筒シェルを考え，この軸電流によって円筒中に渦電流が誘起されたときの磁界を求める。

conducting cylindrical shells axial currents of an infinite length are surrounded conducting cylindrical shell conducting cylindrical shell are considered, and the magnetic fields are obtained when eddy currents are induced by this axial current in cylinders.

of JYOCHŌ (accelerate) will be done as

(We) help accelerate the conflict in trade with foreign countries.

Thus the sentence acquires opposite meaning depending on the speaker's/receiver's attitude. It is quite difficult to know the speaker's attitude about whether he/she is in favor of something or not from the text segments processed by a computer.

Anyway, these sophisticated interpretations can often be left to the human being who reads the translated text. If the machine makes excessive inferencing in the wrong direction, that will be worse than doing nothing. In this sense, machine translation system should not do too much in the way of extra-linguistic inferences. We have to make an effort to find much more solid and reliable linguistic and nonlinguistic information from the text itself going on to the help of general "knowledge," which is vague and hard to make precise enough to be used for machine translation purposes.

References

[1] M. Nagao, J. Tsujii, and J. Nakamura, "The Japanese Government Project for Machine Translation," *J. Assoc. Comput. Linguistics*, vol. 11, no. 2-3, pp. 4–9, 1985.

[2] J. Tsujii, J. Nakamura, and M. Nagao, "Analysis grammar of Japanese in the Mu-Project," in *Proc. COLING 84*, 1984.7.

[3] M. Nagao, T. Nishida, and J. Tsujii, "Dealing with incompleteness of linguistic knowledge in language translation," in *Proc. COLING 84*, 1984. 7.

[4] J. Nakamura, J. Tsujii, and M. Nagao, "Grammar writing system (GRADE) of Mu-Machine Translation Project and its characteristics," in *Proc. COLING 84*, 1984. 7.