Hitachi's Research and Development in Knowledge Engineering

Koichiro Ishihara, Tadashi Hirose,
Atsushi Okajima and Hiroyuki Kaji
Systems Development Laboratory, Hitachi Ltd.,
Kawasaki, Japan

Abstract

Knowledge Engineering is still in its young phase, but is growing steadily. At Hitachi, a Japanese firm, a lot of R&D activities are being carried out in this fast growing field as many other major industries in the world. In this paper we first oberview some of these activities and then discuss three representative systems in some detail.

LONLI is an extension of Prolog and has interfaces to other languages such as Fortran and PL/1. It is being used widely within the firm as a system description language.

Machine Translation is one of the most important applications of Knowledge Engineering and is more so in Japan because of its language barrier. We are developing English to Japanese and Japanese to English systems. Their translation processes will be described with the emphasis on the analysis of English and Japanese texts.

1. R&D Activities at Hitachi

1.1 Architecture

Research projects are being carried out to build Prolog and LISP machines on both microprocessors and large scale computers. Both sequential and parallel types are being investigated.

1.2 Software Systems for Knowledge Engineering

(1) LONLI (Logic OrieNted Language Inferencer)

LONLI is an extended Prolog system and will be described later.

(2) S-LONLI (Super-LONLI)

S-LONLI is an end user language for knowledge representation based on LONLI and is an amalgam of object oriented language, logic programming language and frame type language. The emphasis is put on "writability", the ease with which a user can define his knowledge. The S-LONLI processor is written in LONLI and converts a source program into LONLI.

(3) KBMS (Knowledge Base Management System)

KBMS is a knowledge base manager that manipulates three types of knowledge, logical, generalization-hierarchy, and dictionary type knowledge. The system's features are summarized as follows:

a) Insertion and retrieval of knowledge in the knowledge base are easily performed with a simple knowledge manipulation language.

b) The self-associative retrieval function makes it possible to locate desired information even when the knowledge available for query is very limited.

The system is written in LONLI and uses a relational database for knowledge storage.

1.3 Application Software

A wide variety of application systems are being developed. See Table 1.

2. LONLI-A Language for Knowledge Engineering

The word "LONLI" is coined from "Logic OrieNted Language Inferencer" and its pronunciation quite resembles that of the Japanese word for "logic".

Language specification of LONLI is based on DEC-10 PROLOG[8], but our aim to develop another type of PROLOG is to build a practical language, a language with which large application programs can be developed. With this aim in mind, we added the following features.

(1) Handling of katakana and kanji (Chinese characters)

(2) Handling of real numbers and arrays

(3) Interface to other languages (Fortran, PL/1) and to relational data base (RDB1)

(4) debugging facilities

Let us take the interface to Fortran as an example (See Fig.1). Two methods are provided.

(a) Argument copy interface

Arguments of a predicate are copied (1)and handed down to subroutine to Fortran (fsub 1) (2). After processing by fsub 1,the value is unified with other arguments (3). Relations between Predicate names and subroutine names and the arguments types are predefined (A).

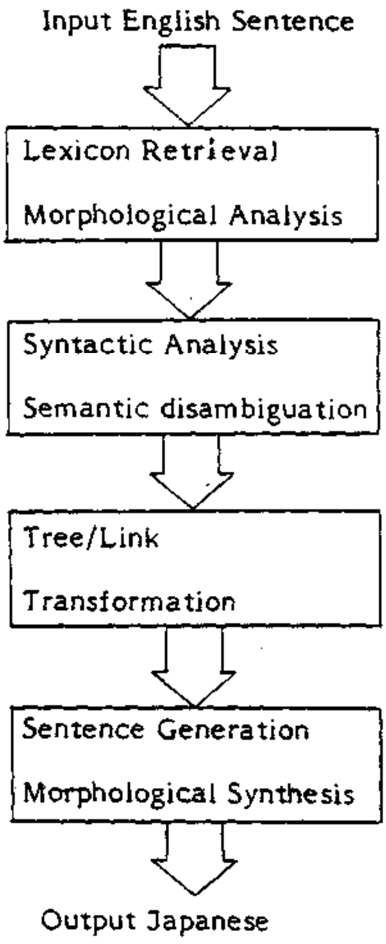
(b) Arguments share interface

Global area is defined in LONLI (B) and the address of the area is handed down to subroutine to Fortran (fsub 2) (4). LONLI supplies special predicates to refer and up-date variables in the area. When performance is important, the second method is used.

An interpreter of LONLI has been implemented on VOS3 and its processing speed is currently about 10 KLIPS on HITAC M-200H. Our image of problem solving using LONLI is shown in Fig. 2.

3. English to Japanese Traslation System

Our English to Japanese (E-J) system is based on a syntax directed parser 10, 11 and a syntactic transfer mechanism. The translation process is as follows. (See Fig. 3 and Fig. 4)



The lexicon is consulted with the help of morphological analysis. Idiomatic collocations are also entered in the lexicon.

Syntactic analysis is important and will be described later.

Clausal elements and phrasal elements are taken apart to word elements and their word order is changed so as to fit in a Japanese sentence.

Post positions that are peculiar in Japanese text are added. Proper Japanese equivalents are selected from the lexicon and predicates are conjugated.

Fig. 3 E-J system

Now let us look into the syntactic analysis and semantic disambiguation. (See Fig. 4).

(1) After the lexicon is consulted, part of speech for each word is decided.

When a word has several candidates, the words before and after it will be checked.

(2) Phrasal elements (PE) are constructed from word elements (WE) and Clausal elements (CE) from PE and WE. During this process, syntactic roles are assigned to each element. This process is repeated until finally a sentence is constructed.

(3) In deciding a syntactic role, we classify verbs into a dependency pattern which is a simplified version of Hornby's [12]. See Table 2 and 3.

(4) Semantic Disambiguation. To select proper meaning of a word, we are trying several method. we classify nouns into several semantic categories such as Time, Place, Abstract Concept, Instrument and so on. We have written rules for selecting proper meanings by the combination of a verb pattern of a transitive verb and the category of its object. We have also developed rules for the combination of prepositions and its object. Although this approach is primitive and its effect is limited, it works well in some cases. For more powerful semantic analysis, we need more extensive research. E-J system is written in PL/1. And the dictionary has about 70,000 entries. Test translation is under way using articles on economics such as Wall Street Journal, and Business Week and augmentation of rules and the dictionary is in progress. See Fig. 5 for sample translation.

4. Japanese to English Translation System

Japanese language is rather difficult to analize by a syntax oriented parser because of its ambiguous usage of function words and word order flexibility. To overcome this difficulty, we have decided to adopt semantics directed approach and to use Conceptual Dependency Diagram (CDD) as an intermediate representation[13]. The translation process is shown in Figure 6.

Japanese sentence is not distinctly separated into words as can be seen in sample translation (Fig. 8). So, the first step is to identify words with the help of morphological analysis.

Then after syntactic and some semantic analysis, a CDD is formed. The idea of CDD is similar to the dependency grammar defined by Hays [14] and Robinson [15]. Predicate phrase nodes (PPN) play important roles in the CDD and nominal phrase nodes (NPN) are connected to PPN according to their "case". Post positions in Japanese sentences usually represent the case of their predecessor and are important clues to the analysis.

A CDD is a complete representation of a Japanese sentense. However, because of the structural difference between English and Japanese texts, direct translation from a CDD (Japanese style) does not usually produce good English. One reason is that Japanese sentences are usually passive or static while English is an active or dynamic language. So, a CDD of Japanese style is transformed to one of English style. See Fig. 7.

Phrase structure is generated from an English style CDD and after morphological synthesis, an English sentence is output.

J-E system is also written in PL/1. The dictionary and the rules are being augmented. Technical documents such as patents and manuals are main objects of J-E system. Sample transltation is shown in Fig. 8.

One of the major problems of J-E system is the frequent ellipsis and functional multiplicity of Japanese languages. So we feel that controlled Japanese or sub-Japanese language should be constructed and this is the target of further study.

5. Concluding Remarks and Acknowledgement.

I have reviewed Hitachi's R&D activities in Knowledge Engineering and I am afraid they are not always systematic. But in this young and fast growing field, the approaches from both sides, i.e. basic technologies and application systems are necessary and important. Also international collaboration will be important, and I hope this paper will be stimulating and of some help to this end.

The authors would like to acknowledge the continuing guidance and encouragement of Dr. Kawasaki and Mr. Ihara of SDL. Thanks are also due to Mr. Nitta of ARL who has lead the research on machine translation.

References

[1] "1984 Hitachi Technology Exhibition", (Hitachi Ltd., Nov. 1984)

[2] Niwa K., Sasaki K., "A new Project Management System Approach : The 'know-how' Based Project Management System" Project Management Quarterly,Vol. XIV, No. 1 (March 1983)

[3] Masui S., "Decision-Making in Time-Critical Situations", Proc. IJCAI (August 1983)

[4] Niwa K. et al. "An Experimental Comparison of Knowledge Representation Schmes." AI Magagine Vol.5, No. 2 (1984)

[5] Watanabe T. etal "An Expert System for Computer Room Facility Layout," 5th International Conference of Expert Systems and their Application (May, 1985)

[6] Yasunobu S., Miyamoto S. et al, "Fuzzy Control for Automatic Train Operation System," Proc. 4th IFAC/IFIP/IFORS International Conference on Control in Transportation Systems (April 1983)

[7] Mori F., Hamaguchi T. et al, "AREQS : An Integration of Data Base, Knowledge Base and Simulator for Business Decision Support System", Proc. 14th Hawaian International Conference on System Science (June 1981)

[8] Clocksin W. F., Mellish C. S. "Programming in Prolog," (Springer-Verlag 1981)

[9] Kinukawa H., Matsuoka H., "Japanese Sentence Analysis for Automatic Indexing", Proc. Coling 80 (Oct. 1980)

[10] Nitta Y., Okajima A. et al, "A Heuristic Approach to English-into-Japanese Machine Trnalation," Proc. Coling 82 (1982)

[11] Okajima A., Nitta Y. et al, "Lexicon Structure for Machine Translation" Proc. ICTP-83 (Oct. 1983)

[12] Hornby A.S. "Guide to Patterns and Usages in English", Second edition (Oxford University Press, 1975)

[13] Nitta Y., Okajima A. et al, "A proper Treatment of Syntax and Semantics in Machine Translation", Proc. Coling 84 (1984)

[14] Hays D.G., "Dependency Theory : A Formalism and Some Observations", Language Vol. 40, No. 4 (1964)

[15] Robinson J.J., "Case, Category and Configuration", Journal of Linguistics vol. 6 no. 1 (1970)

[16] "Machine Translation", Hitachi Technology '85 (Hitachi Review Special Issue, May 1985)

Table 1. Application Systems

Field	System	Comments	Lab.
Natural Language	ATHENE(Machine Translation)	PL/1	SDL
	Japanese to English	Technical documents	
	English to Japanese	News about economics	
	NL Interface to DBMS	PL/1 & LONLI	SDL
OA, DA	Layout of computer rooms	LISP, LONLI	SDL
	TSS Operastion Guidance		HRL
	Logic circuit conversion		SDL
Software Engineering	Re-usable software engineering	LISP/PROLOG	SDL
Industry	WPAC(Project Management)	LISP	SDL
	Diagnosis of Nuclear Reactor		ERL
	Fuzzy control of subway trains	Fortran & PL/1	SDL
Pattern Recognition	Speech understanding		CRL
	Vision understanding		CRL, SDL
Robotics	Intelligent Vehicle		MERL, PERL

CRL : Central Research Laboratory

HRL : Hitachi Research Laboratory

SDL : Systems Development Laboratory

MERL : Mechanical Engineering Research Laboratory

ERL : Energy Research Laboratory

PERL : Production Engineering Research Laboratory

Table 2 Syntactic Roles

Code	Role
SUBJ	Subject
OBJ	Object
TOOBJ	Object in To-infinitive Form
NAPP	Noun in Apposition
GOV	Governing Verb
TOGOV	Governing Verb in To-infinitive Form
ENGOV	Governing Verb in Past Participle Form
ADJV	Adjectival
ENADJ	Adjectival in Past Participle Form
ADVL	Adverbial
SENT	Sentence

Table 3 Dependency Pattern of Verb

Code	Verb Pattern	Examples
V1	Be + …	be
V2	Vi (≠ Be) + Complement, It/There + Vi + …	get, look
V3	Vi 〔 + Adverbial Modifier 〕	rise, walk
V6	Vt + To-infinitive	intend
V7	Vt + Object	begin, yield
V8	Vt + that + …	agree, think
V14	Vt + Object 〔 + not 〕 + To-infinitive	know, bring

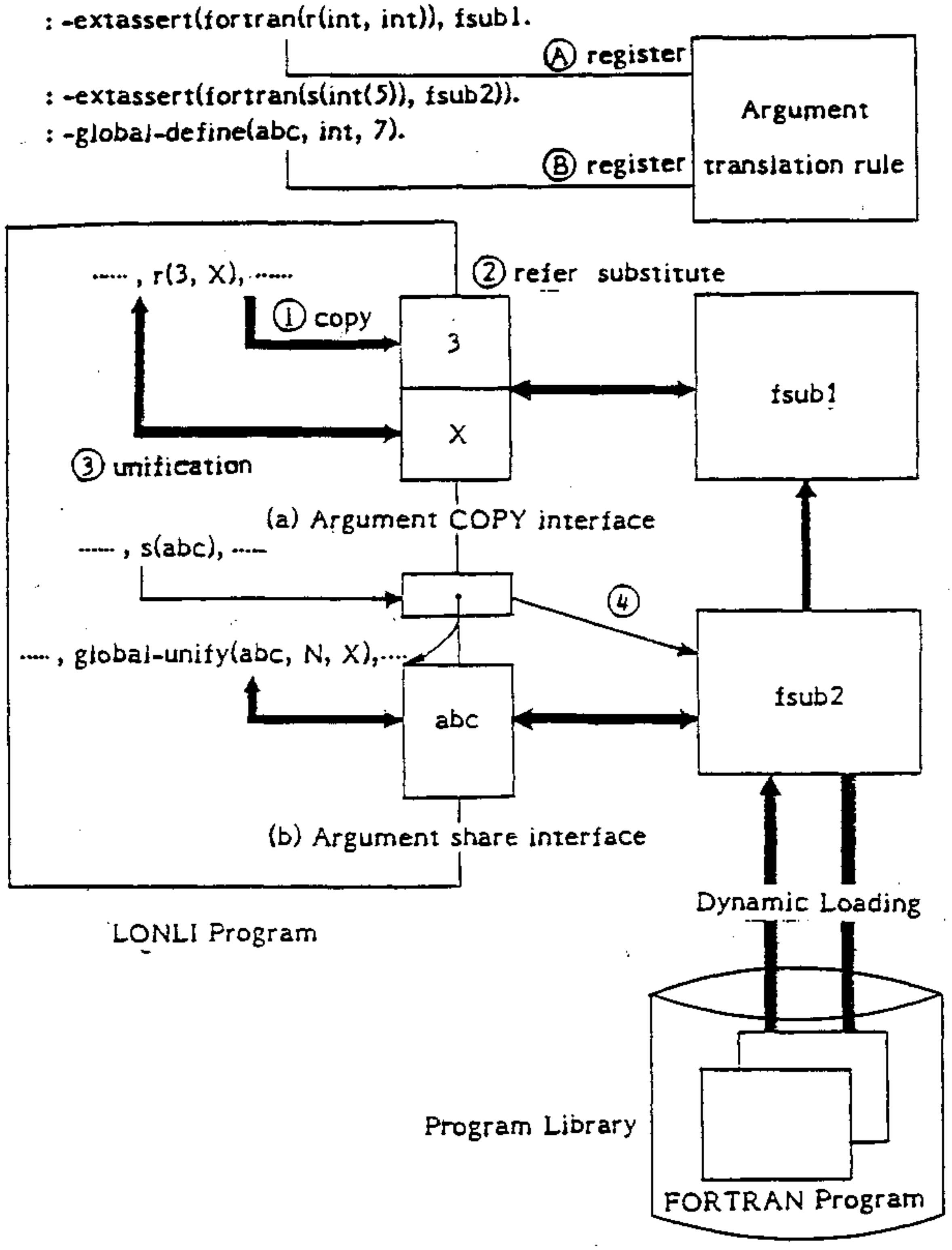

: -extassert(fortran(r(int, int)), fsub1.
Ⓐ register
: -extassert(fortran(s(int(5)), fsub2)).
: -global-define(abc, int, 7).
Ⓑ register
Argument translation rule
②refer substitute
....., r(3, X), ...
①copy
3
X
fsub1
③unification
(a) Argument COPY interface
....., s(abc), ...
④
....., global-unify(abc, N, X),.....
abc
fsub2
(b) Argument share interface
LONLI Program
Dynamic Loading
Program Library
FORTRAN Program

Fig.1 Interface to Fortran

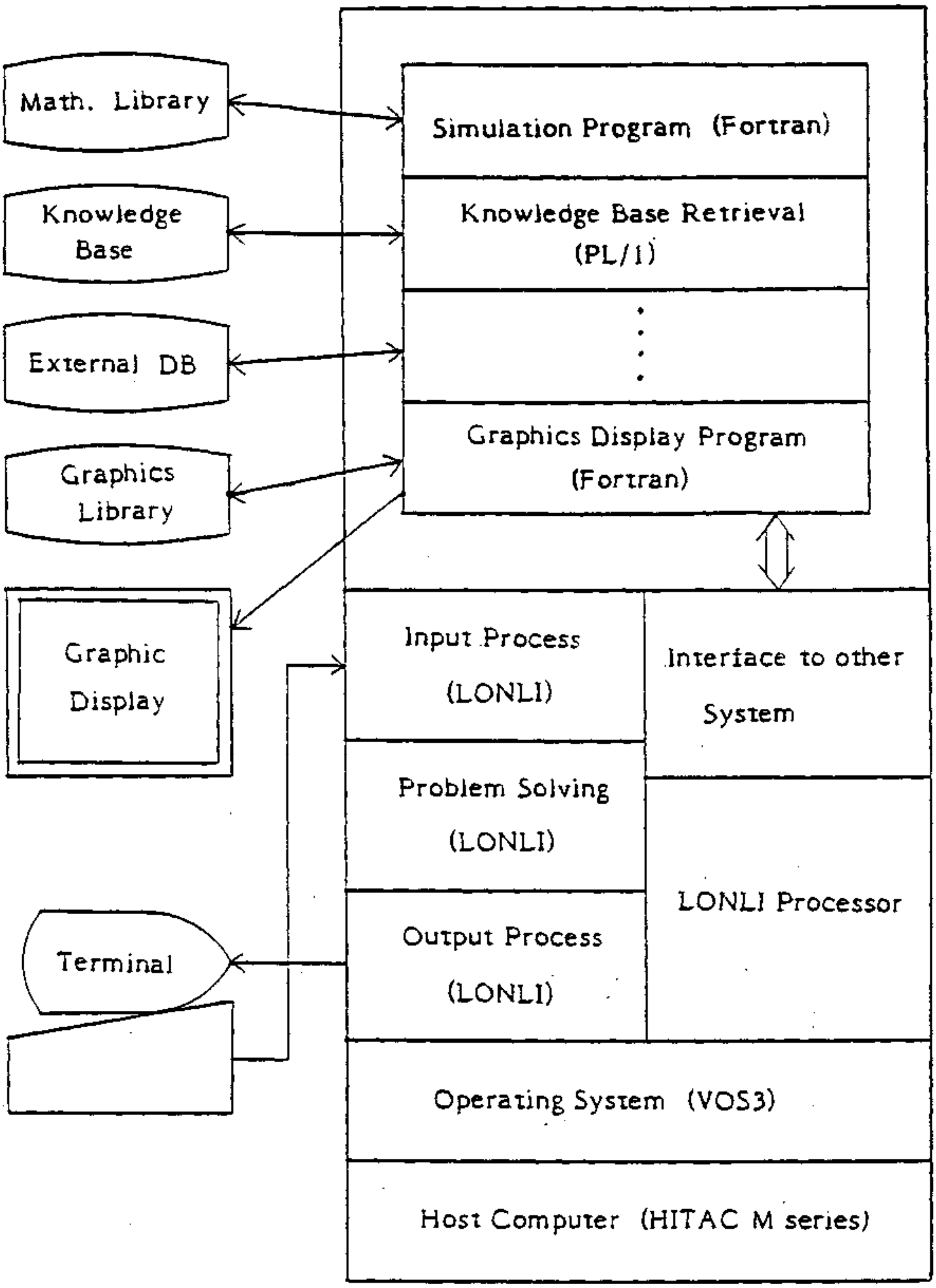


Fig.2 Knowledge Processing System

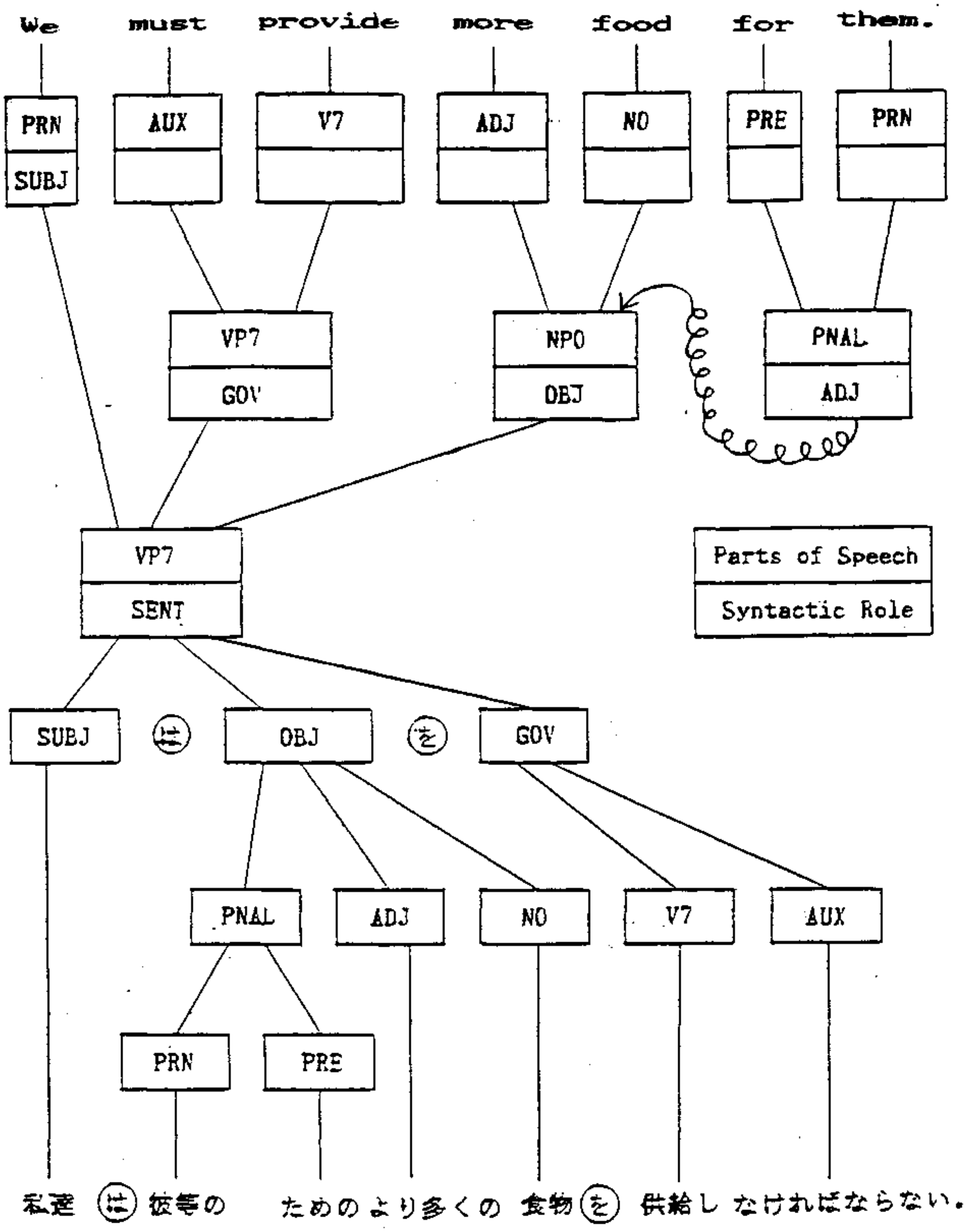


Fig. 4 E-J Translation

(1) PRESIDENT REAGAN ANNUAL ECONOMIC REPORT FEARS THAT THE FEDERAL RESERVE BOARD MAY IMPEDE THE STRONG GROWTH PROJECTED FOR THE REST OF THE 1980S.

　レーガン大統領の年次経済報告は、連邦準備制度理事会がその１９８０年代の残りのために計画された堅調な成長を妨げるかもしれないということを懸念する。

(2) THE PRESIDENT BLAMED THE FED FOR THE LENGTH AND SEVERITY OF THE 1981-82 RECESSION.

　大統領は、連邦準備制度理事会を１９８１－８１景気後退の長さと激しさについて非難した。

(3) THE PRESIDENT'S COUNCIL OF ECONOMIC ADVISERS RECOMMENDED SOME TECHNICAL CHANGES IN FED PROCEDURES FOR SOMEWHAT GREATER MONEYY SUPPLY GROWTH.

　幾分より大きなマネー・サプライの成長のために大統領の経済諮問委員会は、連邦準備制度理事会の手順のいくつか技術変化を勧告した。

(4) THE DOLLAR CONTINUED HIGHER AGAINST MAJOR FOREIGN CURRENCIES AFTER THE WEST GERMAN CENTRAL BANK FAILED TO INTERVENE IN FOREIGN-CURRENCY TRANSACTIONS AS TRADERS HAD EXPECTED.

　トレーダーが予期したように西ドイツの中央銀行が外国為替の取引に介入することを失敗した後ドルは、主要な外貨に対してより高く続いた。

(5) THE DOLLAR ROSE TO A 12-YEAR HIGH AGAINST THE WEST GERMAN MARK AND A RECORD AGAINST THE FRENCH FRANC.

　ドルは、西ドイツのマルクに対して１２年高値とフランスフランに対して記録値へ上がった。

(6) IT EASED SLIGHTLY, HOWEVER, AGAINST THE SWISS FRANC AND JAPANESE YEN.

　スイス・フランと日本円に対して、しかしながら、それは、わずかに下げた。

Fig.5　Sample translation (E-J)

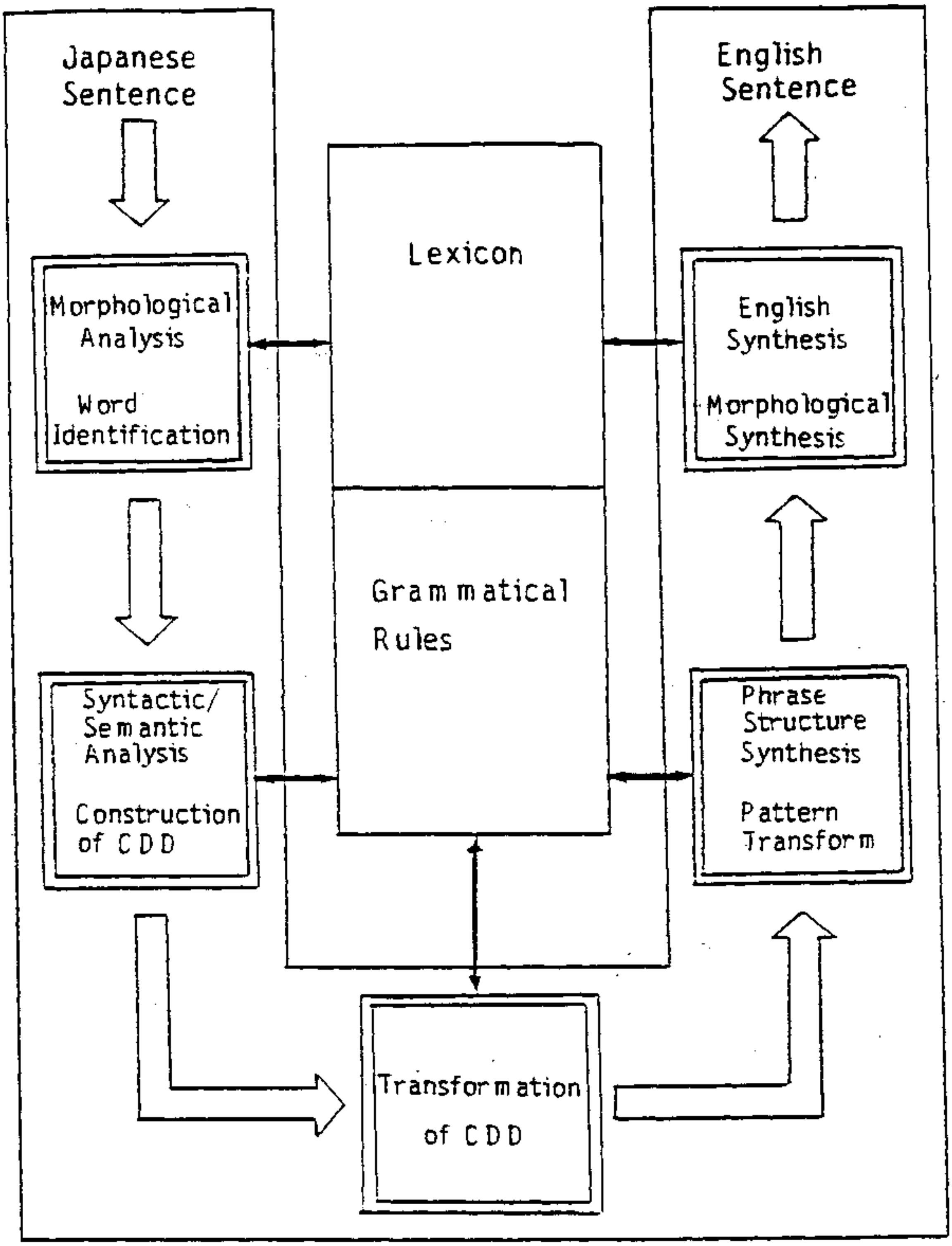


Fig. 6 J-E Trnslation Process

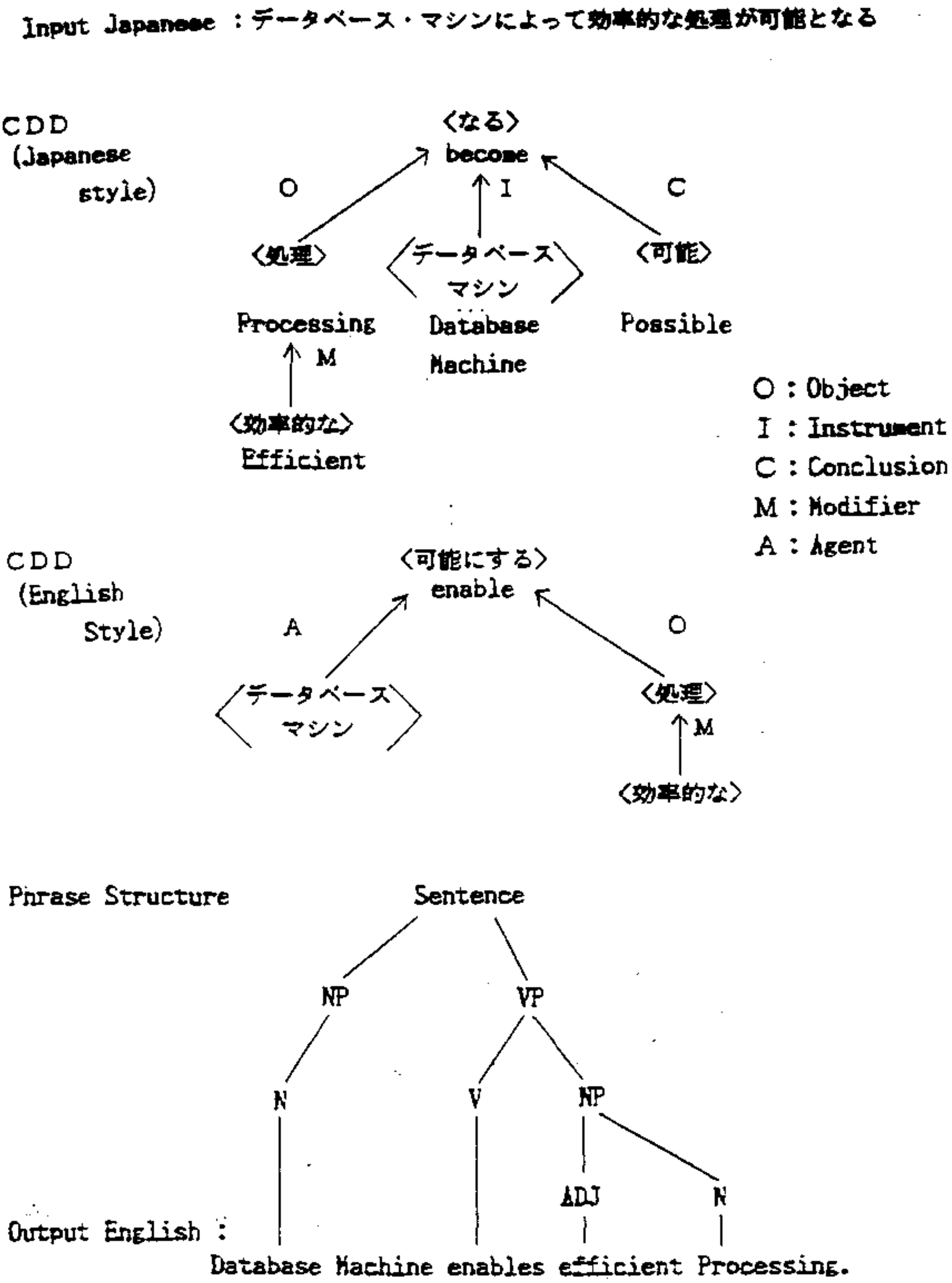


Fig.7 Conceptual Dependecy Diagram

(1) 日本文と英文の間には、表記法や語順の相違と共に、発想や観点の相違に起因する大きな構造的ギャップがあります。このギャップを解消するため、日本語と英語の意味的関係を表す新しい言語モデルを我々は開発しました。その結果、機械翻訳の実現への見通しが得られました。

Ther is a large structural gap that originates from difference in way of thinking and point of view together with difference in notation and word-order between Japanese sentences and English sentences. We developed the new language model that expresses semantic relationship between Japanese and English to resolve this gap. As a result, prospect for realization of machine translation was obtained.

(2) 多重処理システムの効率に関する故障の影響

The effect of failure about the efficiency of multiple processing system

(3) 計算機効率の閉待ち行列ネットワークモデルにおける近似的応答時間分布

Approximate response time distribution in the closed queuing network model of computer efficiency

(4) ヒルベルト空間におけるＰｓｅｕｄｏ－Ｐａｒａｂｏｌｉｃ方程式の解について初期値問題

On the solution of Ｐｓｅｕｄｏ－Ｐａｒａｂｏｌｉｃ equation in ヒルベルト space I initial value problem

(5) 仮想入出力によって、一時的データセットへの入出力処理の効率が向上する。

Virtual input/output improves the efficiency of input/output processing to temporary data set.

(6) ［目的］被試験装置内のアドレスバス、データバスを引出し、アダプタを介して試験装置に接続し、被試験装置のプログラムの走行に従って各アドレスやデータを取り込み、その状態を判定する

[purpose] To lead out address bus and data bus in device to be tested, to connect it to test device via adapter, to take each address and data in acordance with the running of the program of device to be tested and to discriminate the state.

Fig. 8 Sample Translation (J-E)