

機械翻訳に対するロングマン辞書データベースの応用

長尾 真 中村 順一 畑崎 香一郎 藤田 公一
(京大・工)

1. はじめに 自然言語処理を行う場合、処理のために必要な情報を貯えておく処理用辞書をどのようにして作成するかは、重要な問題である。自然言語の解析アルゴリズムや生成アルゴリズムの研究が盛んに行われている。これらのアルゴリズムを多量の文に対して適用しようとした場合、最大のネックは、これらのアルゴリズムが使用する、個々の単語についての情報、即ち、処理用辞書の見出し語が少なすぎるという点にある。見出し語の個数という点では、通常の中型辞書程度で十分であろう(専門用語、固有名詞を除く)。計算機による辞書引きが可能な通常辞書として、新コンサイス英和辞典¹⁾、Longman Dictionary of Contemporary English (LDOCE)²⁾ などがある。本来、これらの辞書は、人間のために作られたものである。このため、計算機による自然言語処理のために必要な情報を十分に含んでいるわけではない。しかし、LDOCEには、比較的多くの情報が含まれている。そこで、LDOCEをオン・ラインで英文の解析、日英機械翻訳における訳語選択、英文生成に利用する実験を行った。本報告では、LDOCEに含まれている、自然言語処理のための情報について紹介し、これを用いた実験結果について述べる。

2. Longman Dictionary of Contemporary English (LDOCE) LDOCEは、イギリスのLongman社から発行されている英和辞書である。これには、自然言語処理にとって、次の4つの便利な点がある。

(1) この辞書は、計算機を利用して作成されたものである。このため、印刷物としてだけでなく、磁気テープの形でも入手可能である。この磁気テープには、固もしくは絵の部分を除く、印刷に必要なすべての情報が校正済みのデータとして含まれている。そこで、オン・ラインによる利用が容易である。

(2) 単語(Head Word (HW))に対して、そのHWが用いられる統語的な環境の記述がコード化されて与えられている。このコードは、Grammatical Code (GC) と呼ばれている。GCは、従来の「他動詞」や「自動詞」の区別を細分化したものである。GCは、各HWの各定義(語義)毎に与えられている。

(3) 単語に対して Box Code (BC) が与えられている。このコードには、名詞に対する意味的制限が意味マークとして含まれている。なお、このBCは、印刷物中には表わされていない。³⁾

(4) HWの定義が、約2000単語に制限された単語(Controlled Word)により記述されている。このため、定義自体を計算機により処理し、利用できる可能性がある。⁴⁾

以上の4点の内、(2)と(3)との特徴が計算機による自然言語処理にとって重要である。

2.1. Grammatical Code (GC) GCは、HWがある定義で用いられる場合の統語的な環境の記述をコード化したものである。GCは、動詞、形容詞、副詞と名詞に対して与えられている。GCの種類の数を表1に示す。GCは、アル

77 バットと数字の組み合わせで表現されている。たとえば、

They considered him their enemy.

という文中で用いられる動詞 "consider" の定義に対しては、"X1" という GC が与えられている。ここで "X" は、Vt + O + C の形で用いられることを表わしている。また "1" は、動詞の後に 1 個以上 (この場合は 2 個) の名詞句 (NP) が現われることを表わしている。即ち、"X1" は、Vt + O + C の形で用いられ、O と C は共に NP であることを示している。動詞に関する GC の一覧を表 2 に示す。

表 1. GC の種類

動詞	31種
形容詞	15種
副詞	7種
名詞	23種
計	76種

表 2 Grammatical Code 一覧 (動詞)

動詞コード	文型例	例文	GC
I	Vi	He <u>lived</u> to be 90.	I3
L	Vi C	She <u>became</u> famous.	L7
T	Vt O	I <u>hoped</u> clean the window.	T2
D	Vt IO DO	He <u>warned</u> her (that) he would come.	D5
X	Vt O C	They <u>considered</u> him their enemy.	X1
V	Vt O NP	to <u>have</u> a house built.	V8

このように、LDOCE の GC は、表層格フレームを記述したものであると言える。即ち、浅いレベルの解析と生成に必要な情報、言い換えれば、C FG 程度の情報を含んでいる。また、表 2 以外に形態素処理レベルのコー

数字コード

0	V (動詞が何と伴う必要がない)
1	V + NP (+ NP)
2	V + Vinf
3	x + TO-V
4	V + Ving
5	x + THAT-clause
6	x + WH-clause
7	V + Adj
8	V + Ved
9	V + Adv, N + 叙述的語句, Adj + 叙述的語句

ドも含まれている。そこで、GC を用いて形態素処理を行うこともできる。なお、GC は、格フレームについての意味レベルの記述を含んでいない。そこで、GC だけを用いて表層レベルと深層レベルの記述 (たとえば深層格フレーム) との間の変換を一般的に行うことはできない。

2.2. Box Code (BC)

LDOCE

の磁気テープ版には、Box Code (BC) と呼ばれるコードが単語の各語義に対して与えられ

ている。BC は 10 個のフィールドから成り立っている。名詞・形容詞・動詞の BC では、このうち 3 つのフィールドがそれぞれ名詞に対する意味的制限を与えている。3 つのフィールドは、box1, box2, illus⁺⁾ と呼ばれる。これらのフィールドの意味マーカの意味は品詞によって異なり、次のようになっている。

(a) 名詞

名詞の場合は、box1 のみが与えられる。box1 は単語がその語義の意味で使われた場合の意味マーカを表わしている。

(b) 形容詞

形容詞の場合も、box1 のみが与えられる。box1 はその形容詞によって修飾され得る名詞の意味マーカを表わしている。

(c) 動詞

box1 は、動詞の主語となり得るもの意味マーカを表わしている。

+) illus が意味的制限を表わすのは、その単語が D または X の GC を持つときのみで、その他の場合は辞書中のマシ絵に関する情報を表わしている。

box 2 は、その動詞の GC が T, V, X のいずれかの場合、その動詞の目的語、GC が D の場合は直接目的語、GC が L の場合は名詞補語に、それとれなり得るものの意味マーカを表わしている。

illus は、その動詞の GC が D の場合はその動詞の間接目的語、GC が X の場合は名詞補語に、それとれなり得るものの意味マーカを表わしている。その他の GC を持つ動詞に対しては、box 3 は与えられない。

次に意味マーカについて述べる。意味マーカはアルファベット 26 文字と 1桁の数字によって表わされている(表3)。これらの意味マーカは階層構造を持ち、それは木の形になるが、その木は唯一ではない。つまり、木のあるノード(ある意味マーカ)に注目すると、そのノードが根ノードでなければ親ノードは1つ以上存在する。例えば、C (concrete) を根とする意味マーカの木は図1のようになる。すなわち、各意味マーカは木の終端ノードに対応するいくつかの意味マーカ(基本意味マーカと呼ぶ)そのものが、その組み合わせを表わしている。その基本意味マーカは表3の記号で言えば、B, D, F, G, J, L, M, N, P, T である。

表3 LDOCE 意味マーカ

A	Animal (not human) - sex unspecified
B	Female animal
C	Concrete (not abstract) - animate or inanimate
D	Male animal
E	Solid or liquid (not gas)
F	Female human
G	Gas
H	Human - sex unspecified
I	Inanimate (not abstract) - solid, liquid or gas unspecified
J	Movable (solid)
K	Male (animal or human)
L	Liquid
M	Male human
N	Not movable
O	Animal or human - sex unspecified
P	Plant
Q	Animate (plant, animal or human)
R	Female (animate or human)
S	Solid - movable or not movable unspecified
T	Abstract
U	Collective and animal or human
V	Plant or animal (not human)
W	Not animate (abstract or inanimate)
X	Not concrete or animal (abstract or human)
Y	Abstract or animate
Z	UNMARKED (no semantic restriction)
1	Human or solid
2	Abstract or solid

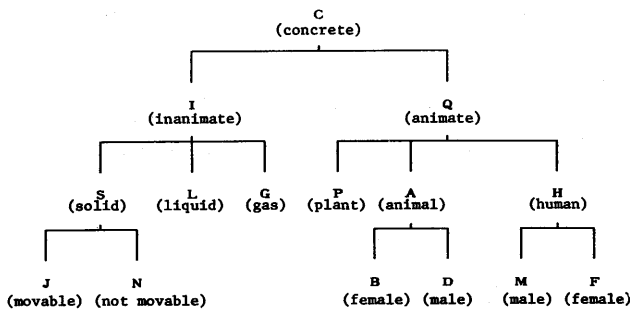


図1 意味マーカの階層構造の例

3. LDOCE データベース

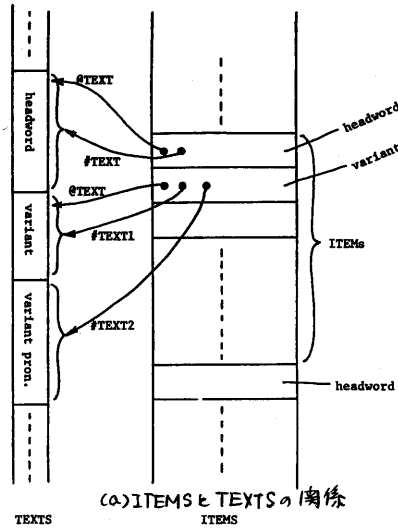
Longman 社から提供された磁気テープは基本的には出版されている印刷物を作成するためのものである。データは、各項目(たとえば、見出し語、異綴語、品詞など)毎に可変長レコードの形で貯えられている。このため、各項目を切り出す必要はない。

LDOCE のデータを効率よく、かつ柔軟に検索するため、磁気テープのデータを図2に示す形で直接アクセス装置上に貯え、利用している。2)

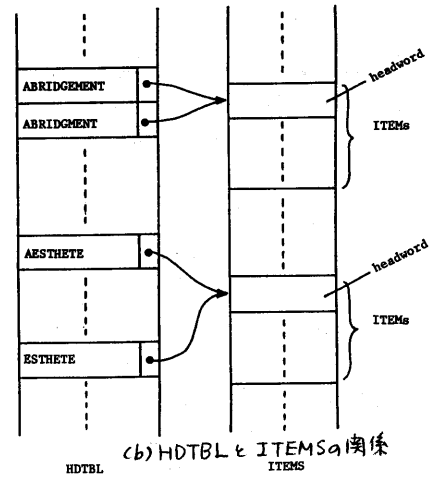
ITEMS には、辞書の各項目と TEXTS 中のエントリー(辞書のひとつの見出し語に対する項目の集合)とを対応付けるための情報が含まれている。これには、見出し語(辞書の見出し語ではなく、検索のためのキーとなる見出し語。異綴語は別のレコードとして貯えている。また字体ソフトコードなどは含まれていない。)と ITEMS へのポインタが貯えられている。LDOCE の情報を単語をキーとして検索することは、HDTBL,

ITEMS, TEXTS)をこの順序でたどることにより可能である。

LDOCEの磁気テープのデータは、2で述べたように各項目毎に区別されている。しかし、項目内のデータは、ほぼ印刷物そのままの文字列の形で表わされている。



(a) ITEMSとTEXTSの関係



(b) HDTBLとITEMSの関係

図2 LDOCEデータベースの構造

このため、原データをそのまま利用することは容易ではない。たとえば、異綴語は、"generalize, -ise"のように完全な形で書かれていない場合がある。そこで、異綴語を復元してHDTBLに登録している。また、GCは、"[L1,7]"のように、"[L1;L7]"を省略して表わされている場合がある。そこで、GCを定型化するプログラムを用意してある。

LDOCEデータベースを検査するプログラムはPL/Iで書かれている。また、LISPプログラムからLDOCEデータベースを利用するためのインターフェースプログラムも作成した。これらのプログラムは、任意の単語(語尾変化していても、不規則変化形でもLDOCE中に含まれる単語であればよい)を与えれば、そのGCやBC、活用形などの情報を返すものである。なお、これらのプログラムと4, 5で用いたプログラムとは、京大大型計算機セン

	Records	Space(KB)
HDTBL (Headword Table)	53,468	2,470
ITEMS	264,138	3,952
TEXTS	—	9,880
Total		16,302

表4 LDOCEデータベースの容量

4. LDOCEをオン・ラインで用いた英文の解析

4.1. 解析の方法

2で述べたように、LDOCEのGCは、CFG程度の情報をかなり詳しく表わしていると考えることができる。たとえば、V2("see"などについて)は、NPの目的語とINF(不定詞句)の目的補語を支配することを示している。そこで、V2に対して、

VP → V-V2 NP INF

という書き換え規則(RR)を用いることができる。

このことから、GCを直接、非終端記号として扱うCFG-RRを作り、CFGパーサを用いて英文の統語的な解析を行った。解析の流れを図3に示す。実験に用いたRRの一部を図4に示す。

このCFG-RRは、以下に示す範囲の言語現象を扱うものである。これらは、LDOCEのGCだけを用いて、CFGパーサで解析することが適当であると考えられる範囲の言

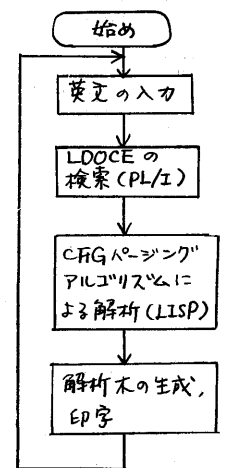


図3 解析の流れ

語現象である。

- (1) 単語の品詞のあいまいさの統語的な解消
- (2) 肯定文(疑問文, 命令文は扱わない)
- (3) 受身文, 使役文(これらは, 統語的に base-generated であるとして扱う)
- (4) 関係詞節 (gap は RR 中に直接記述)
- (5) 動詞の必須格としての前置詞句(動詞の任意格としての前置詞句は扱わない。*)
- (6) 名詞の修飾句としての前置詞句
- (7) 動詞の必須格としての副詞句(動詞の任意格としての副詞句は扱わない。)
- (8) That 節における That の省略
- (9) 関係詞節の関係詞の省略は扱わない。
- (10) 名詞句, 文等の接続詞による接続と, これによる名詞句などの省略は扱わない。

なお, RR は全部で 211 個ある。

4.2. 実験結果 解析結果の例を図5, 図6に示す。実験の結果, LDOCE の GC を用いて統語的な解析を行うと, 次の利点があることがわかった。

(1) 統語的な違いにより, 意味のあいまいさを解消できる場合がある。

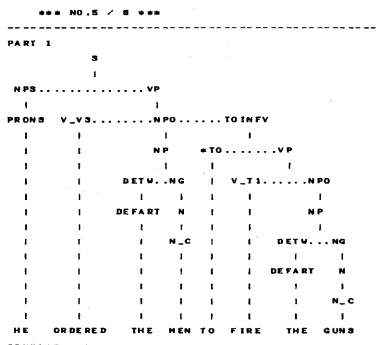


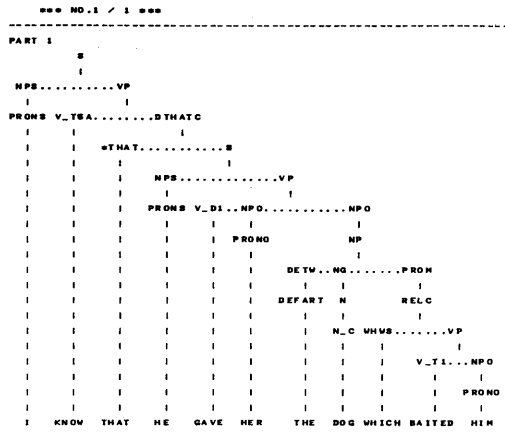
図5 解析結果の例

たとえば, "order (v)" には, 次の6個の定義が与えられている。

- (i) to give an order; command (T5c; T5b; V3)
- (ii) to give an order that (something) should be done or made (T1)

*必須格, 任意格の区別は, LDOCE の GC による区別に従っている。即ち, GC により記述されている前置詞句, 副詞句は必須格とし, それ以外は任意格とし。

V_I4	---	V_UVI	THATC	---	*THAT S
V_I4	---	V_UVI	TOINF-0	---	*TO VP-0
V_I8	---	V_UVI	TOINFP	---	*TO VP
V_L1	---	V_UVI	VP	---	V_I8 PASTPP
V_L3	---	V_UVI	VP	---	V_I8A WHC
V_L4	---	V_UVI	VP	---	V_I8A DTHATC
V_L8A	---	V_UVI	VP	---	V_I8 THATC
V_L8A	---	V_UVI	VP	---	V_I4 PRESPP
V_L7	---	V_UVI	VP	---	V_I3 TOINFP
V_L9	---	V_UVI	VP	---	V_I2 INF
VP	---	V_I0	VP	---	V_L9 ADVP
VP-0	---	V_T1	VP	---	V_L8 PASTPP
P121 = ADJB	---	ADJ_BS THATC	VP	---	V_L7 ADV_F
ADJB	---	ADJ_BS TOINFP	VP	---	V_L7 ADJF
ADJF	---	ADJ_F8A WHC	VP	---	V_L7 ADJB
ADJF	---	ADJ_F8A WHC	VP	---	V_L8B WHP
ADJF	---	ADJ_F8A DTHATC	VP	---	V_L8A WHC
ADJF	---	ADJ_F8 THATC	VP	---	V_L8A DTHATC
ADJF	---	ADJ_F3 TOINFP	VP	---	V_L8 THATC
DTHATC	---	*THAT S	VP	---	V_L4 PRESPP
NG	---	N_N_E	VP	---	V_L3 TOINFP
NG	---	N_A N	VP	---	V_L1 NPS
NG	---	N N	VP	---	V_T6B WHP
NN	---	N_US8 WHP	VP	---	V_T6 WHP
NN	---	N_USA WHC	VP	---	V_T6 WHC
NN	---	N_US THATC	VP	---	V_T6A WHC
NN	---	N_US TOINFP	VP	---	V_T8A DTHATC
NN	---	N_P6B WHP	VP	---	V_T8 THATC
NN	---	N_P3 TOINFP	VP	---	V_T4 GERUND
NN	---	N_CS WHP	VP	---	V_T3 TOINFP
NN	---	N_CS WHC	VP	---	V_T2 INF
NN	---	N_CS WHC	VP	---	V_T1 NPO
NN	---	N_CS DTHATC	VP-0	---	V_D8B WHP
NN	---	N_CS THATC	VP-0	---	V_D8A WHC
NN	---	N_CS TOINFP	VP-0	---	V_D8A DTHATC
NP	---	DEFART ADJ_P	VP-0	---	V_D8 THATC
NP	---	DEFART ADJ_P	VP-0	---	V_D1 NPO
NP	---	N_RS THATC	VP-0	---	V_V8 PASTPP
NP	---	N_R3 TOINFP	VP-0	---	V_V4 PRESPP
NP	---	N_R PROM	VP-0	---	V_V3 TOINFP
NP	---	DEFART N_R	VP-0	---	V_V2 INF
NP	---	INDEFART N_S	VP-0	---	V_X9 ADVP
NP	---	DETW NN	VP-0	---	V_X7 ADV_F
NP	---	NG PROM	VP-0	---	V_X7 ADJ_F
NP	---	PREM NG	VP-0	---	V_X7 ADJ_B
NP	---	DETW NG	VP-0	---	V_X1 NPS
PP	---	PREP NPO	WHC	---	WHWS VP
PREM	---	ADV ADJU	WHP	---	WHW0 TOINF-0
RELC	---	WHWS VP	WHP	---	WHWS TOINFP
S	---	NPS VP	WHP	---	DEFART N_RS THATC



I know that he gave her the dog which barked him.

図6 解析結果の例

になる。

(2) 比較的複雑な文でも、統語的にあいまいでない場合がある。

GC を用いた RR の左辺は比較的長い(1個〜4個)。このため、一見統語的に複雑と思える文でも、あいまいさなしに解析できる場合がある(図6参照)。

(3) 語いが豊富 LDOCE の HW は約 50,000 ある。そこで、専門用語と固有名詞とを除けば、通常使用されている、ほぼすべての単語が含まれている。このため、どのような単語を用いても、一応統語的な解析が可能である。

しかし、以下の問題点があることも明らかになった。

(1) 語いが豊富すぎるため、品詞レベルのあいまいさが生じる場合がある。

現在は、特殊な使われ方をする場合の GC や品詞も、通常の使用法と同一に扱っている。このため、"Time flies like an arrow." における問題と同様の問題を生じることが多い。

(2) 敬の一致、活用型の一致を扱っていないため、誤った解析が出る場合がある。これは、CFG パーカを用いているための問題である。

5. 日英機械翻訳システムへの LDOCE 意味マークの応用

日英機械翻訳シ

ステムでは、トランスファ過程で日本語表現から英語表現への変換が行われている。その際に、日本語の単語に対して英語の適切な単語を選ぶこと、すなわち、

訳語選択が行われる。この訳語選択に LDOCE の意味マークを利用することを試みた。

トランスファの対象となる日本語中間表現(JER)は、基本的には句構造表現をとり、文法カテゴリを表わす各ノードには格マーク・意味マーク・時制・法などの情報が付加されている(図7)。特に、訳語選択をするための前提として、JER 中の名詞・名詞句にはその意味マークが解析段階で与えられると仮定する。

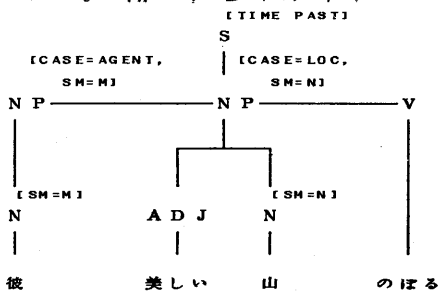


図7 日本語中間表現

かつ、その意味マークは LDOCE 意味マークの体系に従うものとする。

LDOCE 意味マークで訳語選択が行えるのは、名詞・形容詞・動詞に対してであるので、以下では図7の JIR を例にとって話を進めてゆく。

JIR 中の日本語単語には、変換辞書によっていくつかの英語表現が与えられる。例えば、「のぼる」という日本語単語を和英辞典で引くと、「rise', 'go up', 'climb', 'mount', 'ascend', 'reach', 'amount to', 'come to」といった英語表現を得ることが出来る。ところで、LDOCE では、例えば、「climb」に対しては6個の語義がある(図8)。

- climb¹
- 1 to go esp. from a lower to a higher position up, over, or through, esp. by using the hands and feet
 - 2 to go esp. from a lower to a higher position up or over (esp. mountains) as a sport
 - 3 to rise to a higher point; go higher
 - 4 to slope upwards
 - 5 (esp. of a plant) to grow upwards, esp. along a supporting surface
 - 6 to get into or out of clothing usu. with haste or some effort

図8 climb¹の語義(LDOCE)

LDOCE では GC が単語の語義に対して与えられているために、1つの単語でも語義が異なれば異なる GC を持つ場合がある。かつ、トランスファでは、この GC によって構造変換の規則が選ばれる。従って、1つの単語でも、語義が異なれば別々の英語表現と見なさねばならない。すなわち、「のぼる」に対しては計48個の英語表現が与えられることになる。ところがこの中には、「climb¹」の語義6や「amount」の「<馬に>乗る」という語義のように、明らかに「のぼる」に対応する英語表現ではないものも含まれている。理想的に言えば、このような明らかに意味の違う英語表現は意味マークを調べることにより、排除されるべきである。しかし、LDOCE の意味マークがそれほど細分化されておらず強力なチェックは出来ないことと、動詞の場合にはそれ自身に意味マークは付いておらず、動詞の格要素の意味マークのチェックにより動詞の選択を行うこととすることのために、明らかに意味の違う英語表現が選択されることがある。今回の実験では、ある日本語単語の訳語として挙げられる英語表現の集合から、最も適切なものを選択することを目的とし、そのために明らかに意味の違う英語表現はあらかじめ人間の手によって除いておく。その結果、図7の JIR の各日本語単語には図10に挙げられている英語表現が、変換辞書によって与えられることになる。

訳語選択の手続き

(a) 名詞 名詞の場合は、JIR の名詞に付けられている意味マーク(例えば、「山」には 'N' が付けられている。)を包含する意味マークを box1 に持つ訳語が選択される。

(b) 形容詞 形容詞の場合は、それが修飾する名詞(「美しい」の場合は「山」になる。)に付けられている意味マークを包含する意味マークを box1 に持つ訳語が選択される。

(c) 動詞 動詞の場合は、各々の動詞によって深層格要素(動作主・経験者・場所など)と表層格要素(主語・目的語・補語など)の対応が異なる。また、JIR では格要素に深層格マークが付けられている。従って、これらの対応(例えば、SUBJ=AGENT, DOBJ=LOC)を変換辞書の動詞の各訳語に与えておく。この対

```

(IC VERB
 <<< IS = as
 LINE HAVING
 OPTIONAL (NP : CASE IS SUBJ :
 #) = @SUBJ
 OPTIONAL (NP : CASE IS DOBJ :
 #) = @DOBJ
 OPTIONAL (NP : CASE IS IOBJ :
 #) = @IOBJ
 OPTIONAL (NP : CASE IS COMP :
 #) = @COMP
 EXIST
 *
 ANYTHINGS) >
 1 >
 (BOX-CHECK-V @SUBJ @DOBJ @IOBJ @COMP BOXCODE >>>

```

図9 訳語選択のルールの例

応に従って、JIR中より主語、目的語等を取り出し、それぞれ動詞の box 1, box 2 等との包含関係をチェックする。そしてすべての格要素について、意味マーカの包含関係が満たされる訳語が選択される。そしてこれらの手続きにおける JIR の扱いはすべて木構造のマッチング・パターンによって記述される。動詞の場合を例にとると図 9 のルールにより格要素を取り出し、意味マーカのチェック関数を起動することが指示される。

意味マーカの包含関係については、基本的には、訳語の BC 中の意味マーカ (M₁) が JIR 中の名詞・名詞句の意味マーカ (M₂) を包含すれば、包含関係が満たされたとする。包含するとは、意味マーカの階層構造の木 (例えば図 1) で M₁ が M₂ より根の方向にあるか、又は、M₁ と M₂ が同じであるということである。しかしこれでは包含するかもしれないかの 2 値しか得られないため、包含関係を満足する訳語間の優劣がつけられない。そこで、次のように 2 つの意味マーカの距離を定義し、これにより意味マーカの包含関係の評価値を算出する。ただし、M₁ が M₂ を包含するとき、

$$\text{意味マーカ } M_1 \text{ と } M_2 \text{ の距離} = \{M_1 \text{ を構成する基本意味マーカの数}\} - \{M_2 \text{ を構成する基本意味マーカの数}\}$$

$$\text{意味マーカ } M_1 \text{ と } M_2 \text{ の包含関係の評価値} = \text{Const 1} \times [-\{M_1 \text{ と } M_2 \text{ の距離}\}] + \text{Const 2}$$

以上の手続きを図 9 の JIR に対して行くと、図 10 の結果が得られる。SCORE で示されるのが上で述べた評価値であるが、ここでは Const 1=1, Const 2=31 とした。また、動詞「のぼる」の場合は、主語と目的語のそれぞれの意味マーカの包含関係の評価値の平均をその訳語の評価値とした。この結果から、例えは「のぼる」に対しては climb¹ の語義 2 が最も適切であるとされる。

6 まとめ

自然言語処理を行う場合、個々の単語についての情報を貯えている処理用辞書を十分に用意することは、重要な課題である。本報告では、処理用の辞書として、計算機による検索が可能な通常辞書 LDOCE を用いた英文の解析、日英機械翻訳における訳語選択の実験について述べた。解析の実験の結果、LDOCE の GC を用いた CFG-RR により、浅いレベルの解析が可能であることがわかった。今後は、意味マーカをあいまいさの解消に用いる実験と、LDOCE を用いた、深いレベルの解析の可能性の検討を行う予定である。また、訳語選択の実験の結果、LDOCE 意味マーカによって、訳語を 1 つだけ選ぶことはできなくても、多くの訳語の候補をより少数に絞り込むことは可能であることがわかった。これに加えて、GC により統語構造のチェック等を行えば、訳語の候補をさらに絞ることが可能である。

参考文献

- 1) 長尾真 et al., 英和辞書データベースの作成とその応用, 情報大全, 55 年度, 1031-32
- 2) Nagao, M. et al., The Data Structure of LDOCE Database, Internal report, Kyoto Univ., 1981
- 3) Procter, P., Computer Codes for the Definition Space other than Subject field, Letter to Editors of STUDOC, Longman Group Ltd., 1974.
- 4) Michiels, A., Retrieving and Using Semantic Information, Report to Longman Ltd., University de Liege, 1980.

WORD	P.O.S	BOXCODE	SCORE
(HE 1 1)	(pron)	(K NIL NIL)	30.00
(ADJOOO05 /KAREI)			
(ADJOOO05 /UTSUKUSHII)	P.O.S	BOXCODE	SCORE
(BEAUTIFUL 1 1)	(adj)	(Z NIL NIL)	21.00
(FINE 3 1)	(adj)	(Z NIL NIL)	21.00
(PICTURESQUE 1 1)	(adj)	(Z NIL NIL)	21.00
(GOOD-LOOKING 1 1)	(adj)	(H NIL NIL)	FAILURE
(LOVELY 1 1)	(adj)	(Q NIL NIL)	FAILURE
(INOOO006 /YAMAI)			
(MOUNTAIN 1 1)	(n)	(M NIL NIL)	31.00
(CROWN 1 7)	(n)	(M NIL NIL)	31.00
(SPECULATION 1 1)	(n)	(T NIL NIL)	FAILURE
(CLIMAX 1 2)	(n)	(T NIL NIL)	FAILURE
(PEAK 2 6)	(n)	(T NIL NIL)	FAILURE
(IVOOO007 /NOBORU)			
(CLIMB 1 2)	P.O.S	BOXCODE	SCORE
(CLIMB 1 1)	(v)	(H NIL N)	30.80
(MOUNT 2 2)	(v)	(O NIL S)	29.00
(ASCEND 1 1)	(v)	(O NIL S)	29.00
(AMOUNT TO 1 1)	(v)	(Z NIL Z)	21.00
(COME TO 1 2)	(v)	(Z NIL Z)	21.00
(REACH 1 4)	(v)	(Y PREP)	FAILURE
(ASCEND 1 2)	(v)	(M NIL N)	FAILURE
(CLIMB 1 6)	(v)	(J NIL N)	FAILURE
(CLIMB 1 4)	(v)	(P NIL NIL)	FAILURE
(CLIMB 1 3)	(v)	(S NIL NIL)	FAILURE
(GO UP 1 1)	(v)	(J NIL NIL)	FAILURE
(RISE 1 1)	(v)	(T NIL NIL)	FAILURE
(RISE 1 1)	(v)	(J NIL NIL)	FAILURE

WORD: 訳語の候補 P.O.S.: 品詞
 BOXCODE: BC中の意味マーカ。左から box1, illus, box2.
 SCORE: 意味マーカ包含関係をチェックした結果の評価値。
 'FAILURE'は意味マーカ包含関係が満たされなかったことを示す。

図 10 意味マーカによる訳語選択の結果