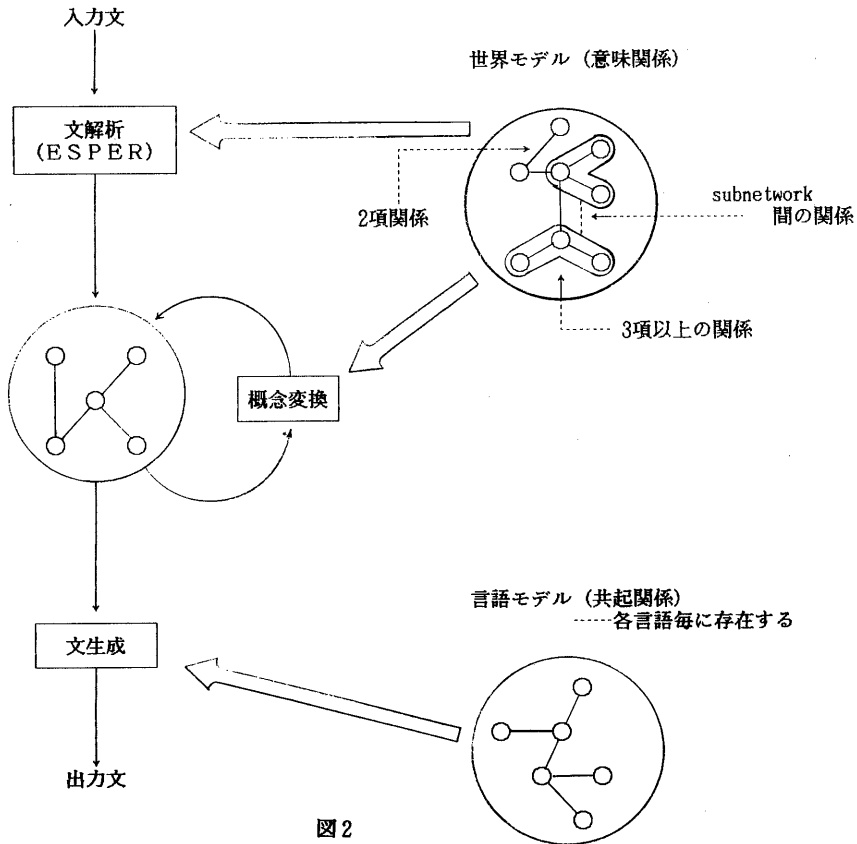
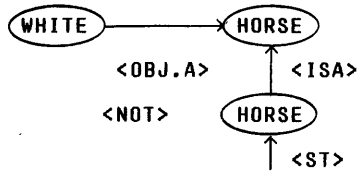


3. 概念変換

取り扱可能な変換規則の数が一定であるとする、一般には中間表現の入力文に対する抽象度をあげるほど、質の良い翻訳結果が得られることになる。入力文を解析するという行為はまさに、この抽象度を上げるために行なわれるのである。

本論文で紹介する変換方式は機械翻訳システム ATLAS / II の変換部に採用されている変換方式である。ATLAS / II の処理概要を図2に示す。





(A)

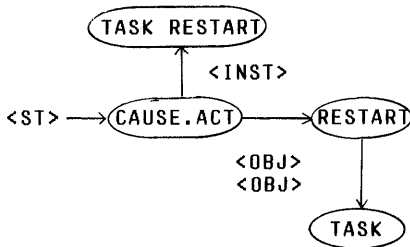
- (HORSE-1, WHITE, <OBJ.A>)
- (HORSE-1, HORSE-2, <ISA>)
- (*, HORSE-2, <NOT>)
- (HORSE-2, *, <ST>)

(B)

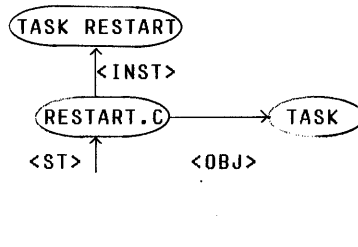
また、1つの変換規則の適用範囲を広げるために、シソーラスにおける上位の概念記号や変数を記述するときのようになっている。変数は*と数字で表わす。relationは部分ネットの関係を表わす。conditionは変数の値域を指定する。ル中に存在するかを調べる。

次に、日本語から英語への概念変換の例を紹介する。

英語では主語 - 他動詞 - 目的語という構文パターンはきわめて基本的である。[2]のため、他動詞を中心して生成した方が、生成しやすく、かつ自然な文が生成される。日本語では自動詞+で表現されているものも、もし英語に他動詞があるならば、再実行させるという文の概念構造は下図(A)のようになる。



(A)



(B)

日本語においては再実行という単語は自動詞であるが、英単語のrestartは他動詞にもなる。自動詞と他動詞では、当然、表わす意味が異なっている。英単語restartの自動詞の表わす概念をRESTART、他動詞の表わす概念をRESTART.Cとする。このとき、上図のCAUSE.ACTとRESTARTはまとめて他動詞概念であるRESTART.Cにした方がよい。これによれば、上図(A)の概念構造は上図(B)の概念構造に変換される。このときの変換規則は次のようなものである。

- ((*0, CAUSE.ACT, <OBJ>) (*1, CAUSE.ACT, <INST>) (*2, CAUSE.ACT, *3))
- (CAUSE.ACT, *4, *5) ,
- (*1, *6, <INST>) (*2, *6, *3) (*6, *4, *5) ,
- (J->E, J<-E) ,
- (C.V, *0, <>) (*6, *0, <CAUSE TO>) (ENG.C, *6, <>))

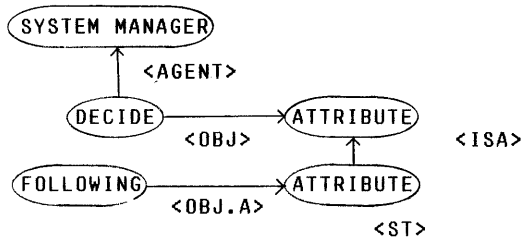
condition部で、(*6, *0, <CAUSE TO>)という関係は*6が*0を他動詞化した概念であることを示す。この例文の場合、*6がRESTART.Cに、*0がRESTARTに当たる。次の(ENG.C, *6, <>)は*6が英語の概念に存在することを示す。

この変換された概念構造から次のような英文が生成される。
Task restart restarts a task.

日本語の叙述表現についても同様の変換が必要になる。
次の例文について考えてみる。

"システム管理者が定める属性は次の属性である。"

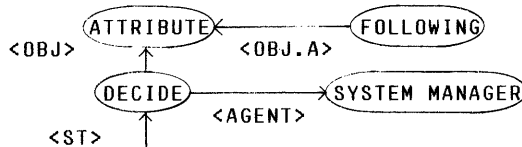
この文の概念構造は下図のようになる。



この概念構造を他動詞中心の構文に変換する変換規則は次のようなものである。

((*0-1,*1,*2) (*0-1,*0-2,<ISA>) (*0-2,*,<ST>) (*0-2,*3,*4)
 (*5,*0-2,*6),
 (*0-1,*1,*2) (*1,*,<ST>) (*0-1,*3,*4) (*5,*0-1,*6),
 (J->E,J<-E),
 (C.V,*1,<>) *2=<AGENT> | <INST> | <OBJ>
)

この規則によって下のように変換される。



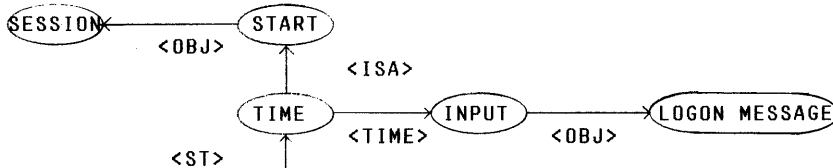
この変換規則は<ISA> というアークによって同じ概念を持つノードが繋がれているときに適用される。condition中の *2 = ... という表現は*2のとりうる値を数え上げて表わしているものである。
 この概念構造によれば、次のような英語が生成される。

The system manager decides the following attributes.

この概念変換に似た例として形式名詞が述語になった場合がある。たとえば、次のような文の場合である。

"セッションの開始はログオンメッセージを入力したときである。"

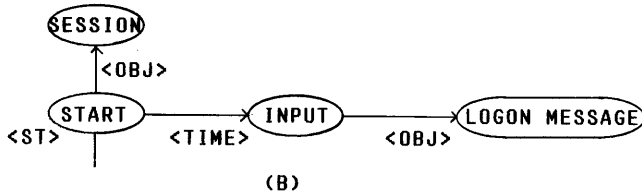
この概念構造は下図のようになる。



この概念構造を下図(A)のルールにより下図(B)のような概念構造に変換する。

((*0,*1,<ISA>) (*2,*1,*3) (*1,*4,*5),
 (*2,*0,*3) (*0,*4,*5),
 (J->E,J<-E),
 (C.V,*0,<>) (C.FN,*1,<>) *3=<TIME> | <PLACE> | <COND>)

(A)



生成される英文は、

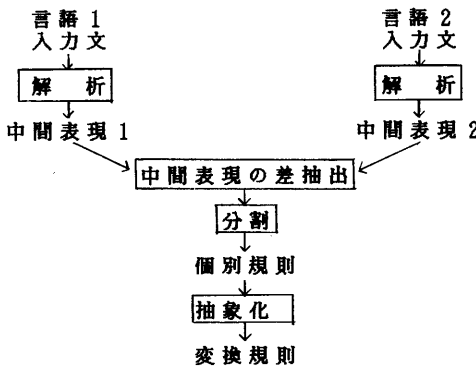
The session starts when a logon message is input.

となる。

6. 自動学習の可能性について

高品質の機械翻訳を目指すには上記のように変換ルールを大量の文章に非常に詳しく調べ、収める必要がある。これを解決する方式として、自動学習による変換ルールの生成が行われている。これは、変換ルールの学習を、大量の文章から自動的に変換ルールを生成するものである。この変換ルールの生成には、変換ルールの学習を、大量の文章から自動的に変換ルールを生成する必要がある。この変換ルールの生成には、変換ルールの学習を、大量の文章から自動的に変換ルールを生成する必要がある。

その方法の一案を図に示す。



学習の目的は、対訳文の対照による変換ルールの抽出である。図に示すように、対訳文の対照による変換ルールの抽出には、対訳文の対照による変換ルールの抽出が必要である。この変換ルールの抽出には、対訳文の対照による変換ルールの抽出が必要である。この変換ルールの抽出には、対訳文の対照による変換ルールの抽出が必要である。

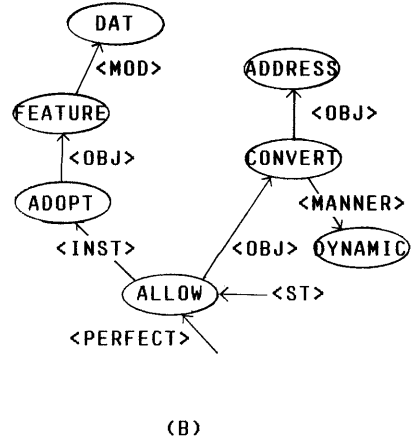
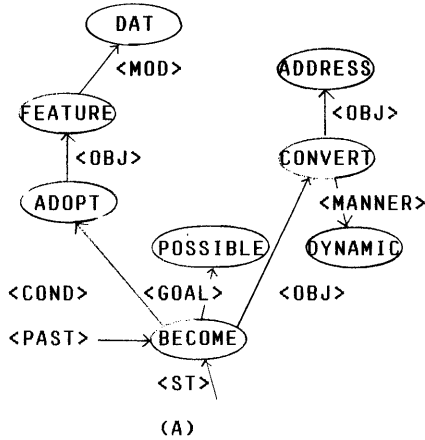
次に具体的な例にしたがって説明する。

下の対訳文が入力されたとする。

(A) DAT機構の採用によりアドレスを動的に変換することが可能になった。

(B) Adoption of the DAT feature has allowed to convert addresses dynamically.

上の2文のそれぞれの概念構造は次のようになる。



これを2項関係集合に直すと次のようになる。

- (DAT, FEATURE, <MOD>)
- (FEATURE, ADOPT, <OBJ>)
- (ADOPT, BECOME, <CAUSE>)
- (ADDRESS, CONVERT, <OBJ>)
- (DYNAMIC, CONVERT, <MANNER>)
- (CONVERT, BECOME, <OBJ>)
- (POSSIBLE, BECOME, <GOAL>)
- (*, BECOME, <PAST>)
- (BECOME, *, <ST>)

(A)

- (DAT, FEATURE, <MOD>)
- (FEATURE, ADOPT, <OBJ>)
- (ADOPT, ALLOW, <INST>)
- (ADDRESS, CONVERT, <OBJ>)
- (DYNAMIC, CONVERT, <MANNER>)
- (CONVERT, ALLOW, <OBJ>)
- (ALLOW, *, <ST>)
- (*, ALLOW, <PERFECT>)

(B)

上図の(A)、(B)で互いに等しい2項関係を取り除くと以下のようなになる。

- (ADOPT, BECOME, <CAUSE>)
- (CONVERT, BECOME, <OBJ>)
- (POSSIBLE, BECOME, <GOAL>)
- (*, BECOME, <PAST>)
- (BECOME, *, <ST>)

- (ADOPT, ALLOW, <INST>)
- (CONVERT, ALLOW, <OBJ>)
- (ALLOW, *, <ST>)
- (*, ALLOW, <PERFECT>)

これで変換規則が得られた。ただしこれだけでは全く等しい概念構造をサブネットとしてもつ解析結果が得られたときしか有効でない。これを汎用の規則とするためにはさらに抽象化の処理が必要である。抽象化は次のように行う。

1. 個別規則を集める。
2. 個別規則をネットワークの等しいものについてグループ化する。
3. 個別規則中の対応するノードで概念の異なるものをソースの上位概念で置き換える。

上記の3.の操作には人間の介入が必要であるかもしれない。

7. おわりに

概念変換はネットワークマッチングによって行われるので、規則数が増えた場合にはこのスピードが問題になってくるとおもわれる。概念変換の果たす役割は翻訳家のもっているノーハウに属するようにはなっていないものである。その速度向上が今後に残された課題である。

謝辞：日頃より、概念構造変換規則の作成に多大なる協力をしていただいているFIEの文屋俊介氏に感謝いたします。

[参考文献]

- [1] 内田裕士ほか：“日英機械翻訳システムATLAS/U”，自然言語処理29-3，1982
- [2] E. G. サイデンステッカーほか：“日本語の翻訳”，大修館書店，1983