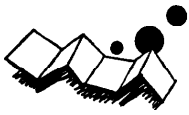


解説

計算機からみた意味表現†



辻井潤一††

1. はじめに

計算機による自然言語処理研究の性格は、1970年前後で大きく変化する。それ以前の研究が、言語の規則性を統語的な側面（文法）からのみ扱えようとしていたのに対して、それ以後の研究では、統語的な規則性に加え、言語が持つと考えられるさまざまな規則性を、どのように計算機処理にうまく反映するかが、研究の主要課題となった。文法以外の規則性の多くは、言語の「意味」と深く関連しているため、研究の焦点が統語処理から意味処理へと変化したといっても良い¹⁾。

本解説では、計算機による自然言語処理の研究において、意味処理がどのように行われてきたか、とくに処理を行うにあたって、自然言語の意味がどのように把握され、記述されてきたかを、いくつかの基本的な見方から整理することにする²⁾。

2. 意味記述への視点

2.1 論理学からの接近——MGの立場

最近、自然言語の意味論全般に強い影響を与え始めている理論に、Montague Grammar (MG) の理論がある。MGは、論理学における「多世界モデルによる意味論」を、ほぼ忠実に自然言語に適用したもので、論理学的意味論の流れの中ではそれほど新しい試みではない。MGの意義は、従来定式化が不可能と考えられていた自然言語においても、適当な統語規則と変換規則とを用意すれば、モデル理論的解釈が可能な論理式へ、かなり直接的に変換できることを示したことである。MGに関しては、(誤植が非常に多いことさえがまんすれば) わかり易い入門書(3)も出ているので、詳細は譲るとして、ここでは以下の議論との関連で問題となるMGにおける意味へのアプローチを、

簡単に整理しておく。

MGの枠組を図-1に示す。図の中で変換規則としたのは、変形文法における変換規則とは異なり、各統語規則によってまとめあげられる各単語(あるいは句)に対応する論理式から、まとめあげられた結果の句に対応する論理式を合成する規則である。自然言語の各単語には、それぞれ論理式が対応しており、文中の単語の論理式から、文全体の論理式が、この規則によって計算されることになる。文の意味は、この論理式とモデルとの対応によって扱えられる。この過程では、述語BOYは、 $BOY(x)$ を満足するモデル中の要素すべての集合と対応づけられ、結局、自然言語中の単語boyは、述語BOYを媒介にして、「述語BOYを満足するモデル中の要素」の集合と対応することになる。この対応関係は、MGにおいては、先験的に与えられていることを想定している。

意味論を、記号とその記号が指示する対象物との間の関係を把える分野であるとする、上の過程は、記号の意味を、指示される対象物そのものと考えたことになる。これを意味の外延による定義という。しかし、この素朴な意味論を、自然言語の意味論とするには、さまざまな不都合がある。例えば、unicorn(一角獣)や「現在のフランス国王」のように指示対象が全くない、したがって外延による定義では全く意味のない語が、自然言語にはあること、「よいの明星」と

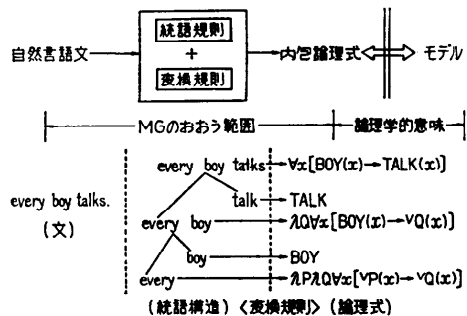


図-1 MGの枠組

† Semantic Description—From View Point of Computer Processing by Junichi TSUJII (Department of Electrical Engineering, Kyoto University).

†† 京都大学工学部電気工学第2教室

「明の明星」のように、外延だけでは全く同じであるにもかかわらず、「意味」的相異のある語が存在すること等である（この他、よく知られた例に、believeのように opaque context をひき起す動詞の存在がある。文献 3) 参照）

このようなことから MG では、モデルと自然言語との対応を考えると、単に外延的な関係だけでなく、単語と指示対象物との間にあって、両者の関係をつける機構を想定する。この機構のことを内包と呼ぶ。したがって、自然言語の表現とモデルとの間にあって、この両者を結ぶ役割を果たす論理式は、MGにおいては、通常の一階述語論理のような論理式ではなく、内包論理式になっている。

すこし、話題が内包と外延という論理学上の問題に入るようになったが、MGにおける意味へのアプローチを整理すると、

(1) 自然言語文からそれに対応した内包論理式に変換する過程、と

(2) 変換された内包論理式をモデルとの対応のものと評価する過程

の二つの過程があることになる。計算機における意味処理においては、この二つの過程を「計算可能な過程」として実際に計算機プログラムとして実現しなければならない。このためには、Montague が先験的に与えられているとした語の内包や外延を計算機的にいかかに実現するか、また与えられた文から MG における内包論理式に相当する表現をいかに抽出するかを計算過程として明らかにしなければならない。

2.2 計算機からの接近

図-2 に計算機における意味処理の枠組を示す。MG では、「モデル」は予め与えられたものとして、自然言語の記述の外側におかれていたが、計算機処理においては、モデルに相当するものもまた何らかの形式で

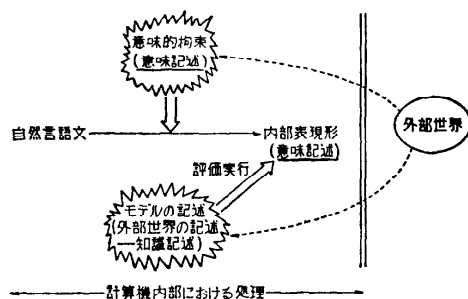


図-2 計算機における意味記述の枠組

計算機的に実現されていなければならない。例えば、自然言語理解システムの先駆となった質問応答システムにおいては、人間からの質問文を、モデルと照合し処理することによって、何らかの答を返す必要があった。このためには、まず質問文をモデルによって解釈できる表現形に変換すること、次にその表現形を使って、モデル上での処理を行うことが必要である。Bobrow の STUDENT⁴⁾ では、英文で与えられた代数の応用問題を、ひとまず方程式に変換し、この方程式を解くことで質問に答えることができた。方程式が MG の内包論理式に、方程式を解くことが、モデルとの対応による論理式の評価に対応している。このように文の意味を、自然言語とは別個の、独立した規則性を持つ別の体系での記述を使って表わし、その体系上で計算・処理を行うという枠組は、MG 以前から計算機分野では、一般的なものであった。MG のモデルに対応する「自然言語とは別個の、独立した規則を持つ体系」のことを「外部世界」と呼び、この世界での規則性を計算機的に記述することを、「知識記述」と呼ぶ。自然言語の意味記述の一つの側面は、文の意味をこの「知識記述」と関連づけて把えることである。このことについては、3. で詳しく論じることにする。

MG においては、自然言語文から論理式への変換は、カテゴリ文法にもとづく統語規則とそれに付随した変換規則とによって、直接的に行われることが仮定されていた。しかしながら、実際の自然言語文の解析においては、意味的な情報を付加しなければ、非常に多くの曖昧さが生じることはよく知られている。

MG では、例えば「テーブル」と「家具」という二つの単語は、現実世界を反映したモデルとの対応において、たまたま部分集合の関係にある二つの外延集合に対応づけられるにすぎない。しかしながら、心理学者の Miller⁵⁾ が指摘するように、このような外部世界での集合間の関係を反映したような心的構造がすでに人間の中に形成されており、これが人間の持つ文法知識と有機的に結合されて、文解析の能力を形成している。Miller は、この心的構造を心的辞書 (Mental Lexicon) と呼んでいるが、この存在が文の文法的に可能な解析の中から、最も自然な解析を選び出すのに重要な役割を果たしている。計算機から見た意味記述のもう一つの側面は、文解析の際の意味的拘束条件として、この心的辞書に相当する外部世界の近似表現を計算機的に実現することである。この問題については、4. において詳しく論じることにする。

3. 文の意味—モデルとの対応

文の意味を、計算可能な「別の体系」で表現するという方法には、別の体系として何をえらぶかによってさまざまなものを考え得る。STUDENT のとった代数方程式という表現形式は、計算機で計算可能という意味では理想的であったが、反面、自然言語文の「意味」を把えるという見方からすると、極くかぎられた表現になっていた。

一般的で、しかも計算可能な枠組で文の意味を記述しようとする試みには、述語論理式を用いる方法がある。この方向の最も初期のものは、Coles 等による一階述語論理の採用である⁶⁾（日本語に対する同様な試みを西田等が行っている⁷⁾）。これらの試みには、一階述語論理を採用したことによって、表現力そのものが自然言語の意味を記述するには十分ではないという欠点があった。ただし、一階に限定したことによって、モデルとの対応による真理判定は、導出原理 (resolution principle) に基づく定理証明プログラムによって実行できることになり、計算可能という点からは、原理的な資格を備えていたことになる。

MG にみられるように、自然言語文の意味を表現するためには、一階よりもさらに高階の論理体系が必要となる。このような高階の論理体系に対しては、モデルを規定する公理系がつかれないという意味で、原理的な計算可能性すらなくなってしまう。そこで、このような論理体系を採用するときには、直接モデル理論的な解釈を実行するか、あるいは表現力をなるべく落さずに、何らかの工夫を入れ込むことによって、一見計算可能になるような枠組をつくる必要がある。STUDENT と同時期につくられた Raphael の SIR⁸⁾ や、最近の Hendrix⁹⁾、Schubert¹⁰⁾、辻井¹¹⁾等の意味ネットワーク、Bobrow、田中等の KRL-SRL^{12,13)}、あるいは Minsky の Frame 理論以来発展してきたさまざまな「知識表現」は、すべてこのような枠組を目標してきたといえる。

3.1 Lexical Decomposition による意味記述

‘John sends Tom a letter’ は、論理式

$$(\exists x)[\text{SEND}(\text{John}, \text{Tom}, x) \& \text{LETTER}(x)] \quad (1)$$

に変換されよう。ここで、英文中の単語 send は、論理式中の述語 SEND に変換されている。これと同様に、‘Tom receives a letter from John’ は、

$$(\exists x)[\text{RECEIVE}(\text{Tom}, \text{John}, x) \& \text{LETTER}(x)] \quad (2)$$

となろう。(1),(2) の論理式の間には、このままでは何らの関係もない。したがって、二つの英文を聞いた時に、我々がほぼ同じ「意味」をうけとることを説明するためには、何らかの別の記述が必要となる。

もし、計算機内部に現実のモデルが何らかの形で実現されていれば、(1)・(2) のそのモデルでの評価値は常に同じになるであろう。しかし、このようなモデル理論による解釈を実行したときに、(1)・(2) が常に同じ真理値を持つという間接的な同義性だけでは、計算機による処理には不十分である。John, Tom が何に変化しようと、また、letter 以外のものをやりとりしようと、send と receive の間には上と同じような規則性があることがいえなければならない。

これを解決する一つの手段は、自然言語の体系とは独立に存在する体系の規則性として、すなわち、外部世界に関する知識の中に、この同義性を説明する機構があると考えることである。この立場では、次のような公理が、外部世界に関する知識としてあると考える。

$$(\forall x)(\forall y)(\forall z)[\text{SEND}(x, y, z) \leftrightarrow \text{RECEIVE}(x, y, z)]$$

しかしながら、公理によって同義性を説明する方法は、同義性を示すためには常に「定理証明」のプログラムを起動しなければならず、単に非効率であるだけでなく、前に述べたように高階の述語論理を採用した場合には不可能になってしまう。

この同義性を、外部世界での規則性とはみずに、自然言語表現に内在する同義性とも見ることできる。外部世界で同じ事象が、自然言語で表現される時に二通りに表現されると考える立場で、この立場からは、文から文の意味を表わす表現形に変換する際に、上の二つの文を同じ表現形に縮退することになる。この方法では、「物の移動」関係を示す、より抽象的な述語を導入し、この述語を使って英語の単語 send, receive の意味を表現しておく。この方式をとった例として、Schunk の CD 理論の例を図-3 に示す¹⁴⁾。Schunk は同義性を説明するために、primitive 述語の考え方を使ったが、さらにここで設定した各 primitive 述語に対して、(テーブルやプログラムによって) モデル理

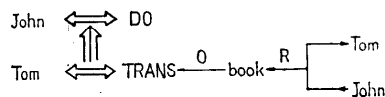


図-3 Schank の CD 理論による表現

論的解釈を与えておけば、証明手順を使うことなく、文の意味のモデル理論的解釈も実行できることになる。これが計算機からみた lexical decomposition の利点でもある。MG の枠組を忠実に計算機で実現する立場の Hobbs も、モデル理論的解釈を「計算可能」な枠組にするために、この lexical decomposition の手法を用いている¹⁵⁾。

この primitive による単語の意味記述には、どのレベルの述語を primitive として選択するか任意性があるために、一般的な手法とはなりにくい。例えば、Miller は抽象度の異なる primitive を使うと、英語の動詞 'hand' が、いくつものちがった記述を持ち得ることを示している⁵⁾。

しかしながら、計算機によって行う仕事の範囲をうまく限定すれば、その範囲内での基本的な概念や操作を設定できる場合がある。その場合には、それらの概念や操作に対応して primitive 述語を設定すれば良い。九大では何枚かの線画を連続して提供し、この線画によってあらわされた物語を日本語で表現するシステムを開発したが、この場合には、線画中にあらわれ

る位置の変化と時間的流れに関する述語を primitive 述語に選び、これらの primitive 述語の組合せで日本語の各動詞の意味を記述しておくだけで十分であった¹⁶⁾。図-4 にその例を示す。このような考え方を、より工学的な場面に適用したものが、次に述べるデータベースに対する自然言語によるアクセス・システムである。

3.2 自然言語によるデータベース・アクセス

データベースに対する自然言語によるアクセスは、自然言語による質問文の意味をデータベース検索用言語に翻訳することによって実現できる。しかも、STUDENT の代数方程式と同じように、一旦この表現形に変換できれば、あとは自然言語処理とは独立にその表現を計算すること（つまり、データにアクセスすること）ができる。

この種の先駆的なシステムとしては、BBN の LUNAR、あるいは最近では Illinois の PLANES¹⁷⁾ 等の試みがあるが、ここでは SRI の LADDER を例にとって考える¹⁸⁾。LADDER は、ARPA net で分散的に蓄積された海軍のデータベースを対象に、自然言語でアクセスしようとするシステムで、(図-5 参照) のような構成をしている。ここで自然言語処理と関係するのは、INLAND と呼ばれるプログラムで、これは質問文

'What is the length of the Kennedy?'

を、
(IDA' ((NAME EQ JOHN # F. KENNEDY) (? LENGTH)))

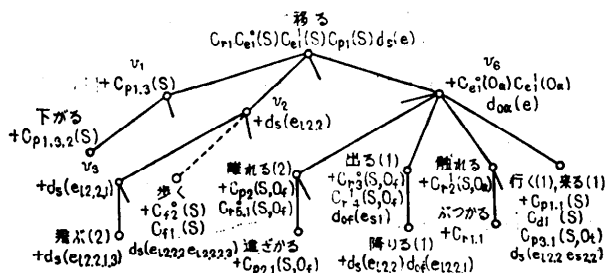
すなわち「名前が JOHN # F. KENNEDY である船の長さを求めよ」という検索式に変換する。この変換は、MG の場合と同様に、統語規則と変換規則とを対にした規則によって実行される。(図-6 参照) に INLAND で使われた規則の例と、それによる変換過程の例を示す。

LADDER のような「自然言語によるデータベース・アクセス」システムが対象とする分野には、

1. 語彙数がかなり少ない。
2. 自然言語文を検索式に変換する部分と、データベースの検索を行う部分(LADDER では IDA, FAM) とが明確に分離できる。といった特徴がある。このために、対象とするデータベースとその上での操作を決めれば、primitive 述語の選定とそれによる単語の意味

意味成分	構造的特徴	分析法	備考
$C_{p1}^1(s)$	EXST ¹ (s)	辞書	存在
C_{r1}	$t_1 - t_0 > 0$	数値	時間経過
C_{p1-1}	$0 < t_1 - t_0 < \Delta t$	"	"
$C_{p1}(s)$	$ I(s) \neq 0$	"	変位
$C_{p1-1}(s)$	$u^1(s) \cdot I(s) > 0$	"	前進
$C_{p1-1}(s)$	$ I_p(s) \neq 0$	"	} 垂直方向の変位
$C_{p1-1-1}(s)$	$I_p(s) < 0$	"	
$C_{p1}(s, O_a)$	$ K^1(s, O_a) > K^0(s, O_a) $	"	離れ
$C_{p2-1}(s, O_a)$	$ K^1(s, O_a) - K^0(s, O_a) > L$	"	"
$C_{p2-1}(s, O_a)$	$\{ I(s) < K^1(s, O_a) \} \cap \{\theta(I(s), K^1(s, O_a)) < \Delta\theta\}$	"	s の O_a 方向の接近
⋮	⋮	⋮	⋮

(a) primitive 述語*



(b) primitive による動詞の意味記述

* primitive 述語は、その構造的特徴が線図形上で満足されているかどうかを直接調べることによって、その真偽が判定される。

図-4 日本語動詞の意味記述

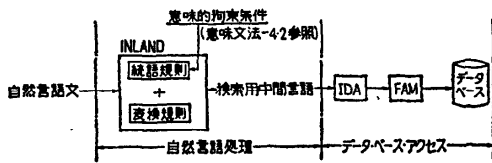
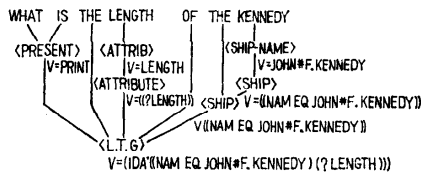


図-5 LADDER システムの構成

< L.T.G > ⇒ <PRESENT> THE <ATTRIBUTE> OF <SHIP> |
(IDA (APPEND <SHIP><ATTRIBUTE>))

(注) < >で囲まれた記号は、書き換え規則の非終端記号(4.2参照)。(IDA (APPEND -----))は、この統語規則につけられた変換規則で<SHIP>、<ATTRIBUTE>に対応して作られた表現式から、<L.T.G>に対応する表現式を作り出す。



LADDERにおける変換例

(注) 変換の過程が、MGの場合とほぼ同じように、各統語規則につけられた変換規則によって実行されていることに注意

図-6 LADDER の文法規則と変換例

記述、および自然言語処理とは独立に行われる推論・演繹といった操作の問題が、ほぼ自動的に解決することになる。

この種のデータベースのアクセスシステムは、自然言語による応答システムの中で最も実用に近づいたものの一つであろう¹⁹⁾。この成功は、この分野が自然言語の処理をする部門とは独立に、独自の計算規則を持つ体系(例えば、関係代数)を持っていたこと、文の意味をこの体系に従って表現すれば、後は自然言語処理とは独立してデータベース・アクセスが実行できたことによる。データベースが知識記述のもっとも単純な形式のものだと考えると、自然言語の処理と知識記述、あるいは知識記述上での操作が独立していたことが、この分野での成功をもたらしている。

しかしながら、自然言語を処理する部門と、外部世界についての知識を処理する部門との相互関係が、このように一方的な関係になるのはまれで、多くの対話システムにおいてはこの二つの部門がかなり緊密にむすびつくことが必要となる。この場合には、外部世界についての知識も、データベースとその操作のように閉じた体系では済まなくなり、自然言語処理に向けた形で組織化されていなければならない。自然言語

理解システムと知識表現の研究とが、密接に結びついてゆくことになる。

3.3 知識の記述と応答システム

3.1では、Lexical Decomposition, すなわち単語の意味を primitive の組合せによって記述しておき、これを使って文の意味を表現することを考えた。しかしながら、人間による実際の対話においては、文の字義通りの意味からは容易に推測できない要求なり意図がある。ある文を発話することによって伝えられることがある。このような要求なり意図を把握するためには、個々の単語の意味から句のレベルの意味、さらに句の意味から文全体の意味を構成するという方法だけでは不十分であり、全体の対話の流れの中で、文の意味を把握し直す必要がある。

例えば、SHRDLU では人間からの要求が、

「red block を blue block の上から取り除け」

というように、文の字義通りの意味で表明されることを前提にしていた。しかし、同じことを

「blue block の上の red block が邪魔だ」

と表現した場合はどうであろうか。文の意味を理解したといえるためには、この文を字義通りに受け取るだけでは不十分であろう。文が字義通りの意味以外の情報を伝達できるのは、話者の意図・話題の構造・話題の対象に関する知識等が、話し手と聞き手の間にある種の了解事項として存在しているためである。したがって、この過程を実現するためには、了解事項としての知識を計算機内部にいかにも表現しておくか、あるいは、それを文の意味抽出の過程にどのように反映するかが問題になる。

Xerox の旅行計画立案システム GUS²⁰⁾ においては、意味抽出の過程を、

1. 自然言語文 → 句構造表現
2. 句構造表現 → 格構造表現
3. 格構造表現 → 中間表現

の3段階に分けて考えている。特に3段階目の変換は、旅行計画の立案システムという対象分野に固有な変換になっている。計算機システムを作る場合には、そのシステムが何を目的にして、どのような立場で人間と対話するかが予め決まっているために、これを利用して、文の字義通りの意味から「知識表現」へリンクするための表現形への変換を予めプログラム化しておくことがある程度可能であった。(図-7 参照)。

しかしながら、このような先験的知識があまりに強い形でシステム構造に反映されると、人間の応答過程

- (1) 入力文
CLIENT: I want to go to San Diego on May 28.
- (2) 句構造的表現
[S MOOD=DCL
SUBJ=[NP HEAD=[PRO CASE=NOMIN
NUMBER=SG ROOT=I]]
PVERB=[V TENSE=PRESENT ROOT=WANT]
HEAD=WANT
OBJ=[S MOOD=FOR-TO
SUBJ=I
HEAD=[V TENSE=PRESENT ROOT=GO]
MODS=(
[PP PREP=[PREP ROOT=TO]
:]
- (3) 格構造的表現
[CLIENT DECLARE
(CASE FOR WANT/E (TENSE PRESENT)
(AGENT (PATH DIALOG CLIENT PERSON))
(EVENT (CASE FOR GO (TENSE PRESENT)
(AGENT (PATH DIALOG CLIENT PERSON))
(TO-PLACE (CASE FOR CITY
:]
- (4) 中間表現
CMD: [CLIENT DECLARE
(FRAME ISA TRIP-LEG
(TRAVELLER (PATH DIALOG CLIENT PERSON))
(TO-PLACE (FRAME ISA CITY
(NAME SAN-DIEGO))
(TRAVEL-DATE (FRAME ISA DATE
(MONTH MAY)
(DAY 28)]
- (注) (2)においては、SUBJ [主語]・PVERB [動詞]・TENSE [時制]、(3)においては、AGENT [行為者]・TO-PLACE [目的地]、(4)においては、TRAVELLER [旅行者]・TRAVEL-DATE [旅行日] といったスロットが用意されている。

図-7 GUS における変換例

が本来持っている柔軟性を著しく損うこととなる。この欠点は、GUSにおいても見られると同時に、Minsky の Frame, Schanck の Script²¹⁾ 等を用いた自然言語理解システムに共通して見られるものである。入力文から得られた情報をもとにして、いかにシステムの持つ先験的知識を柔軟に変更し、応答過程を制御するかが今後の問題となる。現在のシステムが持つ「default 値」の機構程度では、明らかに不十分である²²⁾。

4. 拘束条件としての意味

計算機で自然言語を処理する場合、文の「意味」を示す表現形が必要になる。3. では、この表現形について論じた。本章では、与えられた入力文からそういった表現形を抽出する過程において使われる「意味」について考える。

4.1 意味標識と格構造

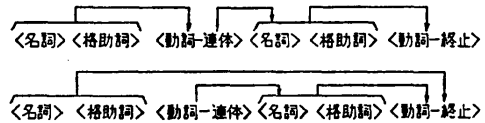
拘束条件としての「意味」の最も直観的な把握の仕方は、意味標識と格構造による方法である。すなわち、各名詞の「意味」を+ANIMATE・+HUMAN といった意味標識を使って表現し、

動詞の「意味」を、この名詞の意味標識を参照しながら、どのような種類の名詞がその動詞と結びつき得るかで示すものである。動詞は文中の複数個の名詞とそれぞれ違った関係で結びつくので、どのような関係で各名詞と結びつくかを、格(行為者格、……)を使って指定する²³⁾。日本語では、格は「格助詞」によってかなり明示的にマークされるので、動詞の意味記述に使われる意味的格と文中にあらわれる表層上の「格助詞」との対応は比較的簡単なルールで記述できる^{24), 25)}。

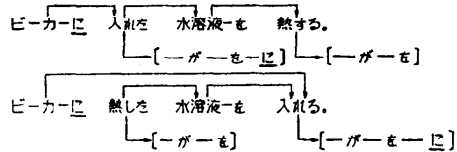
図-8 に、品詞情報のみの場合、動詞の格についての情報がつけ加わった場合、さらに格に入る名詞の意味記述が意味標識によって与えられた場合について示した。記述が詳細になればなるほど、文の解析に対する拘束条件が多くなり、曖昧さが減少してゆく。

意味標識による記述は、意味標識についての redundancy rule を組み込むことによって、推論の体系にもなる(例えば、ある辞書項目が +HUMAN の意味標識を持てば、明示的に記述されていなくても +ANIMATE が自動的に推論される etc)。この種の機構は、これまでのシステムに形をかえて繰返しあらわれている。特に、taxonomy と総称される単語間の階層構造は、名詞間(あるいは名詞的概念)の上位-下位の階層関係だけでなく、Rieger や Woods によ

(1) 品詞レベルの処理



(2) 格パターンのみを考慮した処理



(3) 名詞の意味と格パターンを考慮した処理

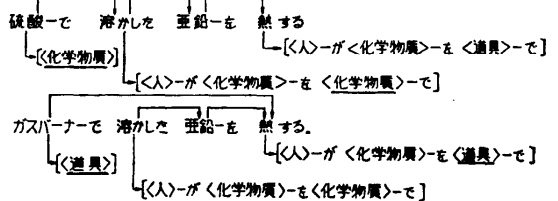
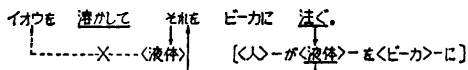


図-8 意味的拘束条件の利用

って、動詞 (predicative concept) 間の関係をも表現するように拡張されている²⁶⁾。

我々は、中学校化学の分野にあらわれる文を解析するために、このような名詞間の階層構造と動詞の格構造とを使用した²⁴⁾。しかしながら、図-9に示すような文を適切に処理するためには、単に動詞の格構造を記述しておくだけでは不十分で、その動詞の指示する「動作」が現実の世界でどのような変化をひき起すかまで記述しなければならないことが、明らかになった。また、動詞と同じように、名詞に関しても、その名詞が指示する「物」が、どのような属性を持ち、他の「物」とどういった関係でつながり得るか等の、格的な構造を持たす必要があることも、明らかになってきた^{13), 24), 27)}。

現実世界において可能であるような、「物」と「物」の関係、あるいは「物」と「事」との関係を近似的に計算機内部に記述しておく、文法的に可能な構造の



(注) イオウは「溶かされる」ことによって、(液体)に変化していることを知っていなければ、この意味的拘束条件が使用できない。この処理に関しては、文献 33)を参照。

図-9 意味的拘束条件と知識記述

中から、この記述に合わないものを排除するのが、意味的拘束条件である。したがって、この条件をより強力にすることは、言いかえると、「物」と「物」、「事」と「物」の可能なむすびつきを、より正確・精密に記述することになり、意味的拘束条件と知識記述の区別があいまいになってくる。

しかしながら、知識表現の問題は、「ある対象分野に関して人間の持っている知識の体系」を、自然言語とは一応独立に記述し、計算機化することであり、一方意味的拘束条件の記述は、こういった外部世界での「知識」がどのように言語現象にフィードバックされ、人間や計算機の文処理能力に反映するかを記述するものである。例えば、意味標識と格構造、redundancy ruleや各種概念間の階層構造等は、知識の中でもとくに言語現象に関係する部分を切りとり、それを言語現象と関連させて記述したものである。したがって、ここでの記述には、論理式で示されるような厳密さや、完全な意味での推論の能力は必要とはならない。むしろ、人間が文を聞いた時に直観的に想起するような、単語間の連想的な関係や、それにもとづいた処理を、

どう具体化するかが問題になる。Quillianは、こういった単語間の連想関係を、semantic network を使って表わし、文や句の構造を解析するシステムを作成した²⁸⁾。その後、semantic network は Woods, Schubert, Hendrix 等によって、より論理的な性質を持つように改編されているが、これらの研究は、むしろ3.で述べた知識記述の体系として network 図式を使う立場の議論であり、Quillianの試みとは本質的に別の方向性を持っていると理解すべきである。

Wilksは、意味的拘束条件の記述が、あくまで現実世界の近似表現であり、しかも自然言語を現実世界における事象以上のことを語る能力が持っている以上、拘束条件を「制限の体系」としてではなく、「優先度の体系」として把えるべきことを主張する²⁹⁾。文法によって、文が文法的に正しい文と非文法的な文とに区別されるのとは対照的に、意味的な基準からは、意味的により妥当な解釈とそうでない解釈との間に序列がつけられるだけだと考える。したがって、文法的にそれ以外の構造が考えられないときには、たとえその構造に対する意味的解釈が妥当なものでなくとも、その解釈を受けつけるべきだとする。知識記述と意味的拘束条件の性格の相異をうまく扱った考え方である。

3. の GUS の項で、あまりに強い先験的知識を持ったシステムにおいては、柔軟性のある応答過程が得られない欠点があることを指摘した。この欠点は、これらのシステムが「知識表現」そのものと、意味的拘束条件とを混同して扱っているためである。今後、こういったシステムが実際に柔軟な応答を行えるようになるためには、知識表現からそれを意味的拘束条件に変換する過程に、より一層の柔軟性が必要となろう。この問題は、自然言語の metaphorical な用法、あるいは知識構造の学習といった、より本質的な問題と、深いところで結びついているように思える^{27), 29)}。

4.2 意味的拘束条件に対する工学的接近

意味的拘束条件を自然言語文の処理に適用してある程度の成功を収めた例として、意味文法 (semantic grammar) の考え方がある³⁰⁾。意味文法は、Burton 等が電気回路についての CAI システム Sophie を開発した時に採用したもので、計算機との応答過程のような場面にあられる文においては、埋込み・受身・使役といったような、統語的な構造変換を必要とする複雑な文体が少ないことを前提にしている。このような前提を認めれば、統語的規則性を説明するのに適した「名詞」・「動詞句」……といった記号系を採用する

必要がなくなり、「属性を表わす名詞」・「船の名前」……といった意味レベルでの単語の分類を、直接的に書換え規則の非終端記号に使用しても良いことになる。3.2で述べた LADDER, PLANES, また Sophie 等は、いずれもこのように意味的拘束条件を直接統語規則に埋め込んだ書換え規則を文法（これを意味文法と呼ぶ）として採用して成功している。

このように、あるレベルの処理に要求される柔軟性が規則によって処理しなければならないほどには大きくなければ、そのレベルでの処理に適した記号体系を省略して、より高いレベルでの記号体系と処理に直接変換してしまう手法が、対象分野を限った工学的応用を考えると有効である。日本語によるデータベース・アクセスを目指す「ヤチマタ」³¹⁾においても、形態素処理や統語処理からみるとちがった記号体系でとらえられるべきいくつかの表現を、「後置詞」的表現としてひとまとめにし、格助詞と全く同じように取扱っている。

日本語においては、文中の「格」の情報は格助詞によって普通マークされるが、それ以外に「～を使って」「～のために」……といった統語的・形態素的にみれば複合的な表現を使うことも多い。我々の研究グループでは、現在日英機械翻訳システムの研究を行っているが、このために、福岡大・首藤氏の収集したこの種の表現を分類・拡張し、これらの複合表現に対して意味記述を与える作業を行っている。これまで困難と考えられてきた意味処理の問題も、こういった補助的データを集積し、それに対してキメの細かい記述を与えてゆくことにより、かなり負担が軽減されてゆくものと思われる。

5. おわりに

自然言語処理における意味の問題、とくに意味の記述に関する議論を整理してきた。統語処理が単語個々の持つ性質・特性といったものを、「名詞」・「動詞」といった数少ないカテゴリに分類して、規則化してゆくことによって行われたのに対して、意味処理においては個々の単語の持つ個別的な性質を、処理にいかんにか反映してゆくかが問題になる。統語処理からみれば、同じ品詞であり同じカテゴリに属する語であっても、意味処理の立場からすると、すべて別々の振舞いをする。これは、類義語はあっても完全な意味での同義語はないことからすると、当然の帰結である。意味処理は、すなわち単語の個性を尊重した処理である。

MITの言語学者 Bresnan, J は、「心理的実在性」の主張のもとに、従来変形規則で説明されていた種々の言語現象を、各辞書項目に書かれた辞書規則によって説明することを提唱している³²⁾。このような改編は、従来強力すぎてもすれば敬遠されてきた変形文法を、計算機向きの枠組にかえる³³⁾と同時に、単語ごとの辞書記述の重視ということから、意味処理にも非常に向いた文法体系を提供することになる。

ともあれ、自然言語処理あるいは意味処理といった問題は、単に計算機科学の一分野で抱えきれぬほど、単純ではない。今後、言語学・心理学・論理学といった境界分野との共同研究によって、自然言語と意味という、ある意味で情報科学の中心ともいべき問題が解決されることを期待したい。しかも、これらの分野が共同して成果をあげるための十分な基盤がすでにある。

本解説を書くにあたって、長尾真（京大・工）山梨正明（阪大・言語文化）田窪行則（京大・文）の各先生との議論が非常に参考となった。記して感謝する。

参考文献

- 1) 長尾 真, 辻井潤一: 自然言語処理プログラム, 情報処理, Vol. 18, No. 1 (1977).
- 2) Winograd, T.: On Primitive, Prototype, and other Semantic Anomalities, in Theoretical Issues in Natural Language Processing (TINLP) 2 (1978).
- 3) Dowty, D. R.: A Guide to Montague's PTQ, Indiana Univ., Linguistic Club, Indiana (1978).
- 4) Bobrow, D. G.: A Question-Answering System for High School Algebra Word Problems, in Proc. of FJCC (1964).
- 5) Miller, G. A.: Semantic Relations among Words, in Linguistic Theory and Psychological Reality (eds. Miller, Bresnan, Halle), MIT Press (1978).
- 6) Coles, L. S.: Talking with a Robot in English, Proc. of 1st IJCAI
- 7) 西田富士夫, 高松 忍: 「制限された日本文の論理式への変換と情報の抽出」, 信学会論文誌 D, Vol. J59-D, No. 8 (1976).
- 8) Raphael, B.: A Computer Program for Semantic Information Retrieval, in Semantic Information Processing (ed. Minsky), MIT Press (1968).
- 9) Hendrix, G. G.: Partitioned Network for the Mathematical Modeling of Natural Language Semantics, ph. D. Thesis, University of Texas (1975).

- 10) Schubert, L. K.: Extending the Expressive Power of Semantic Networks, *Artificial Intelligence*, Vol. 7 (1976).
- 11) Nagao, M. and Tsujii, J.: S-Net: A Foundation for Knowledge Representation Languages, in *Proc. of 6th IJCAI* (1979).
- 12) Bobrow, D. G. and Winograd, T.: An Overview of KRL, A knowledge Representation Language, in *Cognitive Science*, Vol. 1 (1977).
- 13) 田中穂積: 日本語の意味構造を抽出する EX-PLUS について, *信学会論文誌*, Vol-J 61-D, No. 8 (1978).
- 14) Schank: R. C. *Conceptual Information Processing*, North-Holland (1975).
- 15) Hobbs, J. R. and Rosenschein, S. J.: Making Computational Sense of Montagues Intentional Logic, in *Artificial Intelligence*, Vol. 9 (1978).
- 16) 岡田直之, 田町常夫: 図形の意味解釈とその自然言語記述, *信学会論文誌 D*, Vol. J 59-D, No. 5 (1976).
- 17) Walz, D. L.: An English Language Question Answering System for a Large Relational Database, *C. ACM*, Vol. 21, No. 7 (1978).
- 18) Hendrix, G. G. et al.: Developing a Natural Language Interface to Complex Data, *ACM, Trans. on Data Base System*, Vol. 3, No. 2 (1978).
- 19) Walz, D. L.: Natural Language Interfaces, *SIGART Newsletter*, No. 61 (1977).
- 20) Winograd, T. et al.: GUS, A Frame-Driven Dialog System, *Artificial Intelligence*, Vol. 8 (1977).
- 21) Schank, R. C.: SAM—A Story Understander, *Research Report 43*, Yale Univ. (1975).
- 22) Winograd, T.: Frame Representations and the Declarative—Procedural Controversy, in *Representation and Understanding* (eds. Bobrow, Collins), Academic Press, (1975).
- 23) Buce, B.: Case systems for Natural Language, in *Artificial Intelligence*, Vol. 6 (1975).
- 24) 長尾, 辻井, 田中一敏: 意味および文脈を考慮した日本語文の処理, *情報処理*, Vol. 17, No. 1 (1976).
- 25) 防衛庁: 自動インデックス・システムに関する調査研究報告書, 防衛庁 (1979).
- 26) Woods, W. A.: Taxonomic Lattice Structures for Situation Recognition, in *TINLP 2* (1978).
- 27) 池田尚志: 一般化された格構造による意味表現を用いた日本語文の構文解析法について, *信学会論文誌*, Vol. J 60-D, No. 10 (1977).
- 28) Quillian, R.: The Teachable Language Comprehender, *C. ACM*, Vol. 12 (1969).
- 29) Wilks, Y.: Making Preference More Active, in *Artificial Intelligence*, Vol. 11 (1978).
- 30) Brown, J. S. and Burton, R. R.: Multiple Representation of Knowledge for Tutorial Reasoning, in *Representation and Understanding* (eds. Bobrow, Collins), Academic Press (1975).
- 31) 藤崎哲之助: データベース照会システム「ヤチマタ」と名詞句データ模型, *情報処理*, Vol. 20, No. 1 (1979).
- 32) Bresnan, J.: A Realistic Transformational Grammar, in *Linguistic Theory and Psychological Reality*, (eds. Halle, Bresnan, Miller), MIT Press (1978).
- 33) Wasaw, T.: Remarks on Processing, Constraints, and the Lexicon, in *TINLP 2* (1978).

(昭和54年7月9日受付)