

自然言語に内在する論理構造

安藤 司文

長崎大学工学部機械システム工学科

概要

意味言語の基本的な言語構造は自然言語から導出されたものであるが、意味言語は自然言語を記号化して正確に表現できるので、意味言語で自然言語を表現すると、自然言語の言語構造を明確に浮かび上がらせることができる。本論文では、このような方法によって自然言語の論理構造を検討し、次のような結果を得た。

自然言語の論理構造は含意と含意以外に分けられ、含意以外の論理構造として、AND構造、OR構造、NOT構造、THAN構造を、含意の論理構造として、THEN構造、BUT構造、IF構造、EVER構造、AS構造、FOR構造、BY構造を挙げた。そして、複数(PLU構造)、～も(TOO構造)、～さえ(SAE構造)、～だけ(ONLY構造)、～しか(SIKA構造)などもAND構造の一種であることを明らかにした。

The Logical Structures in the Natural Language

Shimon Ando

Department of Mechanical Systems Engineering,
Faculty of Engineering, Nagasaki University
1-14, Bunkyo-machi, Nagasaki-city, Nagasaki, 852 Japan

abstract

Fundamental linguistic structures of a meaning-language is derived from natural languages. Therefore it is possible for the linguistic structures to take shape distinctly, when the natural languages are expressed with the meaning-language. In this paper, investigations are carried on with the logical structures of the natural languages in accordance with the above-mentioned methodology, since the meaning-language can be expressed very precisely by symbolizing the natural languages. This has brought about the results shown below.

The logical structures of the natural languages are classified into two types, i.e the IMPLICATION-structure and non-IMPLICATION-structure. As the non-IMPLICATION-structure, AND-, OR-, NOT- and THAN-structures are taken up. On the other hand, THEN-, BUT-, EVER-, AS-, FOR-, and BY-structure are taken up as the IMPLICATION-structure.

Furthermore exemplification is made with the fact that PLURAL-, TOO-, EVER-, ONLY-, and SIKA-structures belong to the same kind of the AND-structure.

1 はじめに

単なるデータ情報でなく、知識情報を計算機に取り込み、その知識情報を活用しようという動きはエキスパートシステム始めとして、広くみられるようになってきている⁽¹⁾。

また高性能なコンピュータが一般に広く利用されるようになって、知識情報を活用して、法律や医療などに分野にもコンピュータの適用範囲を拡大しようという要求がますます高まっている。

コンピュータ上で知識獲得、知識表現、推論などの知識処理を行うために、各種の手法が提案されている。一階述語論理の演算処理が可能なPROLOG⁽¹⁾、Newellのプロダクションシステム⁽¹⁾、Schankの概念依存構造(CD理論)⁽²⁾、Quillianの意味ネットワーク⁽³⁾、Minskyのフレーム⁽⁴⁾などがよく知られているが、本研究では、コンピュータ上で知識情報を処理するためのプログラミング言語として自然言語の母体である意味言語が適していることを提案している。

我々人間は自然言語を用いて思索を練り、自然言語を用いて互いに情報を交換しているので、自然言語の中にこのような人間の知的活動を支援しているメカニズムが内在していると考えられるが、自然言語の原理⁽¹⁾ですでに指摘しているように、自然言語では必要最小限の内容しか表現されておらず、しかも個別言語の文法、話し手の運用上の判断、状況認識などによってその表現が著しく影響をうけるので、自然言語をプログラミング言語として用いることはできない。

しかし、話し手が自然言語を表出するためには、話し手の頭の中に明確な概念構造が存在しているはずである。そして、その概念構造は簡潔、明瞭、堅牢で、意味を正確に表現していると予想される。そこで、その概念構造を記号化して取り出し、メタ言語として利用することを検討している。これは意味を正確に表現していると考えて意味言語となす⁽⁶⁾。

意味言語は頭の中にあるので、自然言語のように文字列として実際に目で見ることはできないが、自然言語は意味言語から生成されるのであるから、自然言語の中に意味言語の骨格は残っているはずであると考えて、自然言語を手がかりに意味言語の言語構造の解明をすすめている。

その結果、意味言語の言語構造は、多層格-論理構造であることが分かった。つまり、あたかも織物が縦糸と横糸によって模様織りなされるように、縦の格結合と横の論理構造によって、単語が織り込まれて、精緻で稠密な概念の世界が構築されている。

人間はこの意味言語を用いて、質問応答、知識獲得、推論、学習、物語理解、機械翻訳などを行っていると考えられるので、コンピュータ上で意味言語を用いて、このような知的処理を実現し、意味言語の妥当性を検証しようとしている。

すでに、意味言語による質問応答⁽⁶⁾、意味言語による対話と学習⁽¹¹⁾、自然言語に内在する仮説生成機構⁽⁹⁾、意味言語における知識獲得にメカニズム⁽¹⁰⁾、意味言語による多言語間機械翻訳⁽⁶⁾などについて報告した。意味言語によって自然文の意味が正確にコンピュータ上で表現でき、人間と同じようなプロセスで上述の知的処理が可能であることから、知識情報処理には意味言語が有効であることが次第に明らかになってきた。

上述の研究成果を報告する前に、意味言語の基本的な言語構造、意味の構築のプロセス、普遍文法としての意味文法について報告するべきであるが、本意味言語の目指す目標が分かりに難しく思われたので、まず意味言語をどのように利用するのか、それによってどのような効果があるのかを明らかにした。これからは一連の論文で意味言語の基本的な枠組について報告する。

意味言語の言語構造は先に述べたように多層格-論理構造(多層縦-横構造)であるが、本論文ではそのうちの論理構造(横構造)について報告する。

意味言語の基本的な言語構造は自然言語から導出されたものであるが、意味言語で自然言語を表現すると、自然言語の言語構造を明確に浮かび上がらせることができる。本論文では、このような方法によって自然言語の論理構造を検討したので、その結果について報告する。

しかしその前にまず意味言語の概要について簡単に説明する。

2. 意味言語の概要

人間の頭の中にある概念構造をコンピュータ上で表現するために、次の二つの要素MWとPSを設定した。

要素MWは<META WORD>の略で、記号()で示す。このMWに単語(WD)を埋め込む。単語は一つ概念を表しているので、要素MWは概念を表現することになる。この要素MWには、上下左右の4本の結合用の手がある。縦の結合を格結合あるいは縦結合、横の結合を論理結合あるいは横結合と呼ぶ。要素MWには単語WDの他に多くのメンバを持っているが本論文では説明を省略する。“あの”などの冠詞は助詞として取

り扱っているが、冠詞は半角で () の左側、格助詞は右側に示す。従って、" ぬ(太郎) が " のように表記する。" 太郎と次郎が " は " 太郎 " という概念と " 次郎 " という概念が AND という論理関係で結合され、それがまた新しい概念を生成すると考えて、次のように表現する。

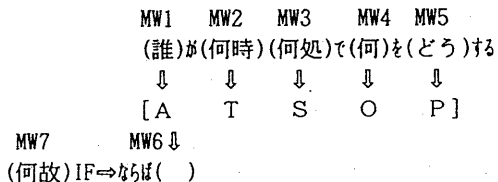
MW17(太郎) AND⇒と MW18(次郎)

↓

MW19() が

() の左側に要素番号 (必要のないときには MW を省略する) , ⇒ の左側には論理関係の種類, 右側にはその論理助詞を示す。MW19には何も単語は埋め込まれていないが、上に示した概念全体 " 太郎と次郎 " が埋め込まれていると考える。

意味は通常各種の意味が複雑に合成された構造をもっているが、それ以上分解できない意味を基本意味単位と考え、これを要素 PS で表した。これは <PRIMITIVE SENTENCE 基本文> の略で、記号 [] で表す。要素 PS は A 格 (Agent case), T 格 (Time case), S 格 (Space case), O 格 (Object case), P 格 (Predicate case) の 5 個の主要な格と, X 格 (Auxiliary case), Y 格 (Yes-No case), Z 格 (Zentai case) の 3 個の補助的な格から構成されている。X, Y, Z の補助的な格はここでは用いないので、説明を省略する。要素 PS の格は要素 MW と結合し、さらに下の要素 MW とも結合するために合計 9 本の結合の手を持っている。要素 PS 1 の A, T, S, O, P の格に要素 MW 1~5 を結合させ、その PS 1 を下の MW 6 と結合させ、さらに、その MW 6 を要素 MW 7 と論理関係 IF の " ~ならば " で横に結合させると、次に示したような概念構造が得られる。



これらの要素 MW には (誰) が, (何時), (何処) で, (何) を (どう) する, (何故) などのデフォルト値が埋め込まれるので, (why), (who), (when), (where), (what), (how) の 1H5W が基本的な意味単位になる。意味単位には、<~がある>, <~は~である>, <~は~をする> という 3 種類があるが、これについては別の機会に詳しく報告する予定である。

要素 PS と要素 MW が縦一横に多層に組み合わせられて、複雑な意味が表現される。我々人間は日常の生活で、あるまとまった概念構造を使ってい

るが、これに動詞や形容詞などの単語を割り当てて用いている。意味フレームにはスロットがあり、ここに単語を埋め込んで意味を確定する。文章の意味は、このような意味フレームや要素 PS や MW がさらに多層に縦一横に結合されて、別の概念構造を表現する。

人間は未知の事象に遭遇したときには、これまでに獲得した知識を活用して、より安全で快適な生活を確保しようとする。そのためには正しい知識を獲得しなければならないが、人間はいろいろな方法で仮説を生成し、それを検証して、有効で信頼性のある仮説だけを知識として記憶しているように思われる。

後述べる { ~ならば, ~である } という IF 構造によって、知識が記述されるので、知識情報を記述するのには、この IF 構造がもっとも重要である。そこで、論理構造を含意とそれ以外に分けて説明する。

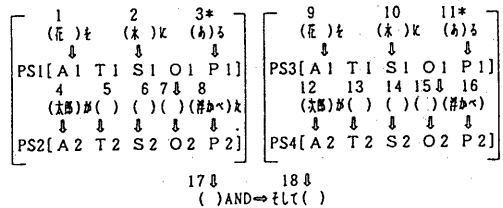
3. 含意以外の論理構造

含意以外の論理構造として、AND 構造、OR 構造、NOT 構造、THAN 構造を挙げる。

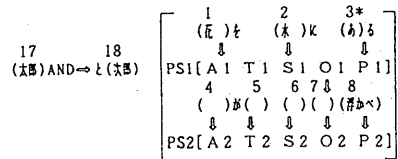
3.1 AND 構造

{ 太郎が花を水に浮かべた } という自然文は意味言語では、図 1 に示すように、MW 1~MW 8 と PS 1~PS 2 の意味フレーム " 浮かべる " ([imi-1]) のスロットの MW 4, MW 5, MW 6, MW 1, MW 2, MW 8 に " 単語や助詞の " 太郎, " が, " 花, " を, " 水, " に, " 浮かべ, " た " を埋め込むことによって、表現される。

{ 太郎が花を水に浮かべた。そして次郎は花を水



(a) { 太郎が花を水に浮かべた。
そして次郎が花を水に浮かべた }



(b) { 太郎と次郎が花を水に浮かべた }

図 1 構造文

に浮かべた}

という自然文は意味言語では、(a)で示すように、{太郎が花を水に浮かべた}という意味フレームがMW17に、{次郎が花を水に浮かべた}という意味フレーム([imi-2])がMW18に縦結合して、MW17とMW18が次に示すように、

MW17([imi-1]) AND ⇒ それ MW18([imi-2])
AND構造で横結合する。意味フレーム[imi-1]と意味フレーム[imi-2]を比較すると、MW4の“太郎”とMW12の“次郎”が異なるだけで、あとは全て同じデータが格納されるので、メモリーの節約のために、(b)に示すように略記する。即ち、次に示すように“太郎”と“次郎”をAND結合して、新しい概念として、それをMW4に縦結合する。

MW19(太郎) AND ⇒ と MW20(次郎)

↓

MW4()が

(b)から自然文を生成するには次のようにすればよい。図示されているように、(MW)や[PS]を埋め込めば、
[[(太郎)AND⇒と(次郎)]が()] [(花)を(水)に(あ)る] (浮かべ)た]

となるが、これから、表現禁止(*印で示す)の単語と助詞や()、[]の記号を取り除き、半角の助詞を全角文字に変更すれば、

{太郎と次郎が花を水に浮かべた}

となる。このことから、この自然文は本来図1の(a)に示した論理構造を持っているが、(b)に示すように略記されたものであると考えることにする。

“太郎達”、“太郎も”、“太郎だけ”、“太郎しか”、“太郎さえ”なども次に示すように、AND構造の一種と考える。

a: PLU構造(複数構造)

{太郎達が花を水に浮かべた}

というのは<太郎が花を水に浮かべた。そして、太郎の他にも太郎として取り扱うべき人、つまり、太郎に相当する人(これを“相当”とする)も花を水に浮かべた>という意味であると考え、この構造文は図2の(a)のようになる。[imi-1]と[imi-2]の内容がほとんど同じであるので、メモリーの節約のために、AND構造の場合の同じような略記法を用いる。即ち、意味言語では[imi-1](PS1~PS2とMW1~MW8)とMW19とMW20で、自然文の{太郎達は花を水に浮かべた}を表現する。“相当”は“太郎達”で十分推測ができるので、自然言語の原理Zによって、表現を禁止(*印で

示す)にする。

b: TOO構造(も構造)

{太郎も花を水に浮かべた}

は<太郎が花を水に浮かべた。そして、その他にも誰かがいて、その人も花を水に浮かべた>という意味になる。“太郎達”は“太郎の他に太郎に相当する人も浮かべた”という意味であるが、“太郎も”は“誰か分からないが、少なくとも、誰かが、同じことをしたという意味であると考えられる。PS1のA1格に結合される部分は

19(太郎)TOO⇒も 20(誰か)*

↓

4()が

となる。上述の{太郎も花を水に浮かべた}という自然文は意味言語では、PLU構造の場合と同じように、意味フレーム[imi-1]とMW19とMW20で表現される。

c: SAE構造(さえ、すら、でも構造)

{太郎さえ花を水に浮かべた}

{太郎すら花を水に浮かべた}

{太郎でも花を水に浮かべた}

これらの文は<太郎が花を水に浮かべた。そして、誰であるか特定していなくても、多くの人も花を水に浮かべた>という意味であると考え。“多くの人”ができるものは、たいした能力は必要がないので、“この太郎”もたいした能力はないという印象を与えることになる。PS1のA1格に結合される部分は次のようになる。

19(太郎)SAE⇒は 20(誰か多くの人)*

↓

4()が

この場合もメモリーの節約のために、意味フレーム[imi-1]とMW19とMW20で上述の自然文を表現する。

d: ONLY構造(だけ、のみ構造)

{太郎だけが花を水に浮かべた}

{太郎のみが花を水に浮かべた}

というのは<太郎が花を水に浮かべた。そして、太郎以外の人(これを“他(以外)”と記す)は花を水に浮かべなかった>という意味であると考え、この意味言語は図2の(b)のようになる。否定は[imi-2 NOT なかた]であるが次のようにMW20(以外)に“NOT”を記入する。PS1のA1格には次のものが結合される。

19(太郎)ONLY⇒は 20(以外)*NOT

↓

4()が

この場合も意味フレーム[imi-1]とMW19とMW20で上述の自然文を表現する。

一括して含意とした。

4.1 THEN構造

{～をして、それから～をして、それから～をした}というように、物事が時間的な順序で徐々に発生する場合の表現に用いる論理関係である。例えば、{太郎が花を水に浮かべそれから次郎はそれを見た}は<太郎が浮かべた。それから、次郎がそれを見た>という意味であると考えられる。これを意味言語で示すと、図3の(a)のようになる。新たにMW17とMW18を設けて、両者を論理結合関係を示す⇒で結び、矢印の左側に論理記号THENを、右側に論理助詞“それから”(“それから”を論理助詞とする)を書き込んで、これらに間にTHENという論理関係が存在することを示し、そして、<太郎が花を水に浮かべた>という意味フレーム[imi-1]と<次郎がそれを見た>という意味フレーム[imi-3]を格結合したものである。この関係を表す記号として、英語の“THEN”を用いた。

上の関係は<太郎が花を水に浮かべた>そして<次郎はそれを見た>というように、AND関係(論理積)であるとも、受け取ることができる。ある出来事が時間的に前後して発生したのかあるいは同時に発生したのか判断することは難しいことがよくある。同時に発生した場合はAND構造であり、前後して発生した場合はTHEN構造である。人間がいろいろな現象を見たときには、時間的な順序関係があるのか、ただ単に並列的に発生したのか俄かには判らないので、とりあえず並列的であると考えてAND関係で表現することが多いと思われる。

明確に区別して表現する場合は、AND関係のときは例えば{太郎が浮かべたというなかで、次郎は見た}とか、THEN関係のときは{太郎が浮かべた後で、次郎は見た}というようにいろいろ工夫して表現されている。

4.2 BUT構造

{太郎が花を水に浮かべたが次郎はそれを見た}は<太郎が浮かべたら次郎は見ない>ことが期待されているのに、この期待通りには行かずに、<次郎は見た>という意味を表現している。これを意味言語で表現したのが(b)である。

4.3 IF構造

これが含意の本命である。
{太郎が花を水に浮かべれば次郎はそれを見る}を意味言語で示したのが(c)である。
現実の世界で検証して、ありとあらゆる場合に正確に実証できれば、ルールとして認められるが、場面、条件を限定して成立すれば、その場面、条件に限定した知識となる。IF関係では過去形や

動作の完了形は表現できない。

4.4 EVER構造

どんな条件でも帰結が決まっている場合の表現で、<太郎が浮かべても、浮かべなくても>、いずれにしても、次郎は見る>という意味を表現している論理関係である。これを意味言語で示したのが(d)である。

BUT関係は当然予想されていたのに、そのような結果にならなかった場合を表現する場合で、EVER関係はどんな条件でも決まった結果にしかならない場合の表現である。

4.5 AS構造

{太郎が浮かべたので次郎は見た}は<太郎が浮かべたことが原因で、次郎は見たのであり、太郎が浮かべなかったら次郎は見なかった>という意味になる。この部分を意味言語で示すと、(e)のようになる。

⇒の左側の(MW)が原因で、右側の(MW)が結果を表す。

4.6 FOR構造

{太郎が浮かべるために次郎は見る}は<次郎が見ると、太郎は浮かべる>という論理があって、その<太郎が浮かべる>という目的を達成するためには、<次郎が見なければならぬ>という動機を説明するための論理関係である。上の自然文の意味言語を(f)に示す。⇒の右側の(MW)が動機で、左側の(MW)が目的である。

4.7 BY構造

{手で、太郎は花を水に浮かべた}という文は{太郎が手を使って、太郎は花を水に浮かばせた}という意味にも取れるが、記述されていないので、分からない。“手”が関係していることだけをしめし、それ以上言及しないときの省略表現であると考えられる。

{車で、長崎に行く}、{バラで、花子を喜ばせる}の“車で”は<車に乗って>、なのか、<車を押して>なのか<車を運転して>なのか、また、“バラ”は<バラをあげて>なのか、<バラをちぎって>なのか分からない。話し手が表現していないから、判断しようがない。話し手がこのように表現したのは、とにかく“車”や“バラ”が関係しているということ表現したい場合の表現であると考えられる。

{東京では、物価が高い}、{その点に関しては、知的である}の“東京では”、“その点に関しては”なども含意構造の省略形であると考えられる。フィルモアは“手で”は道具格としているが⁽¹²⁾、この意味言語では単に“()BY⇒()”の左側にある(MW)のことになる。

5. まとめ

意味言語は明確な言語構造をもち、しかもコンピュータ上で自然文の意味を正確に表現できることから、知識情報処理用プログラミング言語として用いることを検討し、実際に、質問応答、対話と学習、仮説の生成、知識の獲得及び機械翻訳などに適用して、その有効性が明らかになった。

まず意味言語の言語構造、意味の構築プロセス、普遍文法としての意味文法などについて明らかにすべきであったが、分かり難いと思われたので、先に適用例を示して、意味言語をどのように利用するのか、またそれはどのような効果があるのかを示した。

本論文では、意味言語の基本的な枠組としての多層格-論理構造の内の論理構造について説明した。意味言語の基本的な言語構造は自然言語から導出されたものであるが、意味言語で自然言語を表現すると、自然言語の言語構造を明確に浮かび上がらせることができるので、このような方法によって自然言語の論理構造を検討して、以下の結果を得た。

- (1) 含意以外の論理構造として、AND構造、OR構造、NOT構造、THAN構造を挙げ、その表記法を示した。
- (2) 複数(PLU構造)、～も(TOO構造)、～さえ(SAE構造)、～だけ(ONLY構造)、～しか(SIKA構造)などもAND構造の一種である。
- (3) 含意構造として、～であり、それから～である(THEN構造)、～であるが、～である(BUT構造)、～ならば、～である(IF構造)、～であるとしても、～である(EVER構造)、～であるので、～である(AS構造)、～のために、～である(FOR構造)、～によって、～である(BY構造)を挙げ、その表記法を示した。

今後、格構造、意味の構築のプロセス、普遍文法としての意味文法などを報告する予定である。

謝辞

長年にわたって多大のご支援を賜った九州産業大学工学部機械工学科楠本韶教授(元長崎大学工学部機械システム工学科教授)に心からお礼を申し上げます。また長期間にわたって本研究プロジェクトに参加していただいた多くの方々に感謝いたします。

参考文献

- (1) 人工知能学会編(1990):人工知能ハンドブック, オーム社
- (2) Schank,R,C(1975): Conceptual Information Processing, North-Holland
- (3) Quillian,M,R(1968): Semantic memory, in Semantic Information Processing (Minsky,M,L.ed)pp.227~270, MIT Press
- (4) Minsky,M(1975): A Framework for Representing Knowledge, in The Psychology of Computer Vision, edited by Winston, P.H., McGraw-Hill Company (白井, 杉原訳: コンピュータビジョンの心理, 産業図書)
- (5) 安藤司文(1989): 認知プロセスと自然言語-自然言語の構造(I), 認知科学会第6回大会発表論文集pp.128-129
- (6) 安藤司文(1991): 意味言語の提案, 情報処理学会自然言語研究会 情処研報Vol.91, No.80, pp.73-80
- (7) 安藤司文(1991): 自然言語から意味言語への変換, 情報処理学会第43回全国大会講演論文集(3), pp.3-137-138
- (8) 安藤司文(1991): 意味言語による質問応答, 人工知能学会ヒューマンインターフェイスと認知モデル研究会 研究会資料SIG--HICG-9102, pp.43-52
- (9) 安藤司文(1991): 自然言語に内在する仮説生成機構, 情報処理学会人工知能研究会, 情処研報 Vol.91, No.75, pp.1~10
- (10) 安藤司文(1991): 自然言語による知識獲得のメカニズム, 情報処理学会データベース・システム研究会, 情処研報 Vol.91, No.84, pp.1~10
- (11) 安藤司文(1991): 意味言語による対話と学習, 認知科学会学習と対話研究分科会, SIGLAL91-2, 学習と対話Vol.91, No.2, pp.1~10
- (12) Fillmore,C.J(1975): Toward a Modern Theory of Case & Articles, (田中春美, 船城道雄訳「格文法の原理-言語の意味と構造」三省堂)