

## 自然言語理解システム E L L E における 動詞の意味表現について

安原 宏, 小松英二, 北 研二, 山本由紀雄, 加藤安彦  
( 沖電気工業 (株) 総合システム研究所 )

### 1. はじめに

自然言語を計算機に理解させることは、自然言語処理の大きな目標である。計算機での実験システムとしては Winograd<sup>(1)</sup> の積木の世界が有名である。国内でも電総研の先駆的研究がある<sup>(2)</sup>。これらは言語理解の範囲を非常にミクロな世界に限定して人間らしい理解を実現しているが、その方式を日常の言語世界での理解に適用すると種々の問題点が生じてきた。本稿では、広範囲な分野への適用をめざして現在開発中の自然言語理解システム E L L E<sup>(3)(4)(5)</sup> のうち動詞の意味表現を中心に述べる。

### 2. E L L E の概要

E L L E の目標は

- (1) 通常の日本語文が解析できること。
- (2) 文脈解析もできること。
- (3) マンマシンインタフェイス, M T, Q A 等の各種応用システムに使用出来ること。

等をめざしており、汎用で大規模なシステムになることが予想される。そのためにもシステムは統一した設計思想の下で開発する必要がある。

我々は従来の意味を軽視した言語理論にはあきたりないものを感じていた。そもそも自然言語とは、人間が発するものであるが、そういう背景を無視しては真の言語理解は不可能ではなからうかというのが、素朴な疑問であった。そのような基本的考え方にたって日本語解析用の認知モデルを図1のように設定した。

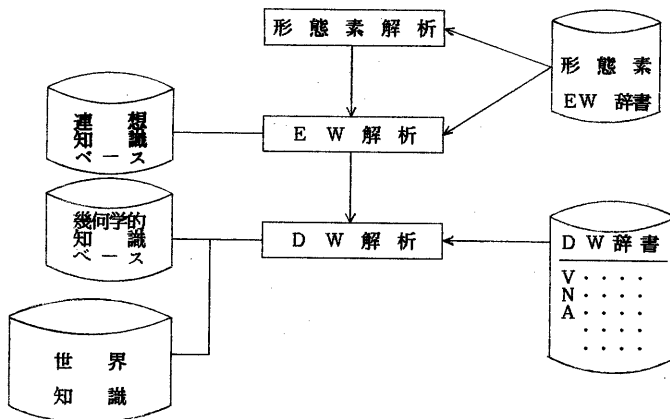


図1 E L L E のシステム構成

まず自然言語が表現する世界を基本世界 (Elementary World, 以下EWと略す) と叙述世界 (Descriptive World, 以下DWと略す) の2層に分割し, 第1層では言語の第1近似の意味を与え, 第2層では具体的な文意を認識するようにしている。EW, DWは共に意味理解に関係しており, 前者は係り受けを後者は具体的な意味理解を行なう。認知モデルを標榜している1つのことは, 自立語を{時間T, 空間L, 物体O, 人間H, 属性A, 行為B, 現象P}の7つの概念に分割していることである。これは自然言語にとって, 時空概念および人間に関する表現が重要であることに基づいているが, この詳細については本稿では述べない。

### 3. 動詞の意味

日本語の動詞の意味を如何に分類するかは色々な見解があるが, その1つとして寺村<sup>(6)</sup>による詳細な分類がある。本システムでは機械処理を目的としていることと, 一応正しい日本語が入力されているものと仮定しているため, できるだけ単純なものを心がけた。そのため例えば\*“彼は今京都でいる”といった文でも“彼は今京都にいる”というのと同等の解釈になってしまうという弱点はある。

格表現が動詞の意味表現として使われることが多いが, それでは, 次のような文の適否を決定するのは困難である。

		木	窓	字	葉書	手紙	
私は	本を	買った。	○	○	X	○	△
私は	本を	読んだ。	X	X	○	○	○
私は	本を	開いた。	X	○	X	X	X

我々は全ての動詞の意味を状況の時間的变化とみなす。

$$\langle V \rangle = d S = S ( t + \Delta t ) - S ( t ) \quad ; \text{但し } S \text{ は状況}$$

このように動詞の意味を状況変化と捉える方法は石綿<sup>(7)</sup>, 上原<sup>(8)</sup>等にもみられるが, それらは文と文の間の状況変化を記述している。一方ここでは1文内での状況変化とみなす点異なる。

動詞の意味記法としてV式を導入する。V式のシンタクスを図2に示す。

$$\{ \text{---} + | + \text{---} \} \left[ ( \langle \text{副詞コード} \rangle ) \right] \langle \text{用言コード} \rangle ( \langle \text{体言コード1} \rangle ( \langle \text{体言コード2} \rangle ) ) .$$

$$\langle \text{体言コード1} \rangle : = \langle \text{体言コード} \rangle [ . S | . 0 ]$$

$$\langle \text{体言コード2} \rangle : = \langle \text{体言コード} \rangle [ . S | . 0 ]$$

図 2 V式のシンタクス



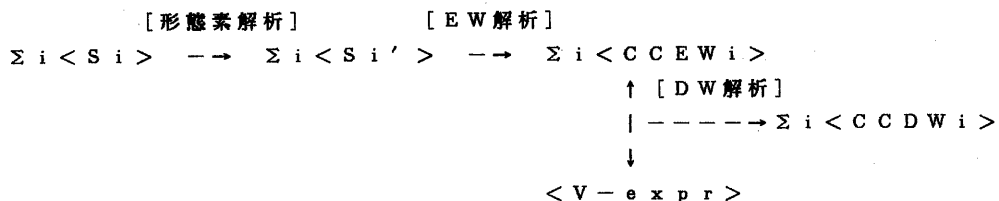


図4 解析の流れ

ここで、 $< S_i >$ は*i*番目の文、 $< CCEW_i >$ は*i*番目のCognitive Chart of EW、 $< CCDW_i >$ は*i*番目のCognitive Chart of DWである。

EW解析は各自立語毎に定義されている概念マーカ（7つの概念のこと）と付属語の意味とで文節の意味と文節間の係り受けを仮決定し、リスト構造にする。この時点では正確な係り受けを決定できないものもあるが、それは次のフェーズのDW解析にもちこす。図5にEW解析の出力例を示す。

\*\*\*\*\* 認知チャート \*\*\*\*\*

Midashi  
Concept, Label

I	L 空は 3	O	H	P	B	A	<p>-----</p> <p>Bun No. = 9</p> <p>Slot 6 is absent. Default number (255) is used!</p> <p>***** 係り受け *****</p> <p>[[空は,低く ],[雨や,雪は ],ふりません ]].</p>
	雨や 7 雪は 8						低く 10
							ふりません 5

< CCEW >

図5 EW解析の出力例

DW解析のキーポイントは<CCEW>と<V式>とのユニフィケーション $\cup$ である。  
<V式>は状況Sの差分dSであるから

$$S = \cup (< CCEW >, < V式 >) = \cup (< CCEW >, dS) = < CCDW >$$

と表現できる。状況Sが<CCDW>に相当しているのである。

$\cup$ の作用のさせ方は、次のような規則に従う。

- (1) 法情報（受身、否定等）によりV式を変形する。
- (2) V式中の . S, . Oのものと、CCEWの S, O属性（基本的には助詞ガ, ヲ）の語とをbindする。
- (3) V式中の概念語は必ずしもCCDWの実体語の有する概念と一致しなくても否定しない。

(4) 2つ以上のAND列は左から順に作用させる。

(5) 現CC EWに存在しないものは、それ以前のCC EWを探す。

(6) 得られたCC EWの認知モデルとしての妥当性をV式中の用言コードのセマンティクスでチェックする。

(7) もし妥当性がなければ、第2の解釈に対して(1)から再度試みる。

V式中の概念語と文中の実体語が、シンタクス上からかなりの確度でマッチングできる時(例えば、ガとヲが助詞として出現しており、述語が主語と目的語を要求するような場合)は問題はないが、そうでないときは、実体語の概念と概念語の親和性の高いものを選ぶ。

例

“ 実力で道路の建設に反対する。 ”

反対する: - + 反 (物 . S (物 . O) ) .

( 反対する: Hum が N ( a b s ) に V する。 )

V式中の‘物 . O’は‘実力’か‘道路建設’のどちらかが候補となるが、‘実力’は概念Aで‘建設’は概念Bのため‘建設’がより‘物 . O’に近いものとして選ばれる

(6)の規則は意味解釈の本質的部分である。それは、世界の構造(幾何学的知識ベースと呼ぶもの。)からみて、得られたCC DWが正しいかどうかをチェックするものである。その時には、L軸は、2次元、3次元に拡大し、その中で、OやHがどのようにふるまっているかを認識することになる。

用言コードのセマンティクスは幾何学的操作で記述されており、bindされたOやHの概念マーカをもつ語は、その意味表現中に形状が記述されているため、その幾何学的操作が論理的に正しいかどうかは、幾何学的知識ベースの命題群で判定することができる。

## 5. おわりに

本稿での動詞意味表現は動詞の意味を動詞をとりまく句の格関係として表現するだけでなく動詞そのものの意味も意味コードで記入している。意味コードの体系だけでは完全に言語の意味を表現することは困難であるが、そのために幾何学的知識ベースといわれるものを導入し、推論部で意味を補完しようとした。文の意味としてラムダ式を導入するものもあるが、これは機械的に意味を表現しようとするものであり、自然言語のもつ複雑な意味をとりだすには不十分であると考えられる。本稿での意味表現は簡潔な表現系であり、それを補う豊富な知識ベースと推論系が必要であるが、それは具体的な応用システムの開発に伴って徐々に進化させる計画である。

## — 参考文献 —

- (1) WINOGRAD, T: 「言語理解の構造」(淵, 田村, 白井 訳) 産業図書 1976
- (2) 田中穂積: 「日本語の意味構造を抽出するシステムEXPLUSについて」, 信学論D. Vol. 61-D, No. 8, 1978
- (3) 安原, 小松, 北: 「自然言語理解システムELLE (1)」  
情報処理学会第29回 (昭和59年後期) 全国大会
- (4) 小松, 安原, 北: 「自然言語理解システムELLE (2)」

情報処理学会第29回（昭和59年後期）全国大会

(5) 北，安原，小松：「自然言語理解システムELLE（3）」

情報処理学会第29回（昭和59年後期）全国大会

(6) 寺村秀夫：「日本語のシンタクスと意味I，II」 ころしお出版 1984

(7) 石綿敏雄：「結合価から見た日本文法」

（水谷他；朝倉日本語新講座 3「文法と意味I」第2章第1，2節）

(8) UEHARA, K; OCHIAI, R; MIKAMI, O: 'Steps toward an actor-oriented integrated parser.' Proceedings of Integrated Conference of FGCS 1984

(9) WILKS, Y: 'An intelligent Analyzer and Understander of English'

CACM 1975. May. vol.18. No.5 pp264~274

(10) 中村捷：「解釈意味論」（安井稔他；英語学大系 第5巻 意味論 第4章）

大修館 1983