

心像意味論に基づく マルチメディア・コミュニケーション理論

横田 将生

福岡工業大学 言語情報工学研究所

我々は、心像意味論と呼ぶ自然言語意味論の構築を行ってきた。心像意味論は、(1)自然言語、(2)心像、(3)知識表現、(4)対象世界、および、(5)それらの間の対応関係(人間の自然言語理解過程を含む)を取り扱う総合的な理論である。我々は、既に、心像意味論に基づいた自然言語理解システムIMAGESを作成している。コミュニケーションは、人間やIMAGESのような人工的システムを含む複数の知的システム(情報体系と呼ぶ)が種々の媒体(文字、音声、画像など)を介して情報(知識、信念など)を交換することにより、共通の部分情報体系を形成することであると定義することができる。本稿では、心像意味論との関連においてマルチメディア・コミュニケーションについて考察を加える。

A Theory of Multimedia Communication based on Mental Image Directed Semantic Theory

Masao YOKOTA

Language and Information Laboratory, Fukuoka Institute of Technology

The authors have been developing a theory of natural semantics named "Mental image directed semantic theory (MIDST)", which is a total theory that deals with (1) natural language, (2) mental image, (3) world, (4) knowledge representation, and (5) the correspondence relations among them, including the human process of natural language understanding. We have already constructed natural language understanding systems named IMAGES based on MIDST. Communication may be defined as information exchange among intelligent systems, so called "information systems", including humans, artificial systems like IMAGES, etc. in order to share a certain common partial information system. This paper presents some consideration on multimedia communication related to MIDST.

1. まえがき

コミュニケーション(communication)は、情報体系(information systems)間の相互作用と定義できる。ここでの相互作用は、複数の情報体系が種々の媒体(言語、画像など)を介して情報(知識、信念など)を交換することにより、共通の部分情報体系を形成することである。異種の情報媒体は、相補的に情報伝達に寄与するが、主導的役割を担うのは、それらの情報体系に共通のシンタクスとセマンティクスを持ったもの、通常、言語である。それら二つのものを明確には持たない媒体(例えば画像)は、それを受けとる側にして

みれば、どこに意味があるのか理解できないのが普通であるので副次的な情報伝達媒体と考えることができる。ここで議論するコミュニケーションは、そのような事実を考慮し、言語を核とする多重媒体(multimedia)によるものである。本稿における議論の特徴は、情報伝達媒体のセマンティクスの基盤を心像(イメージ)にしている点にある。我々は、既に、心像に基盤を置く自然言語意味論(心像意味論と呼ぶ)を提案している。以下では、心像意味論について述べ、それに基づき、コミュニケーションなるものの形式的取扱いについて考察する。

2. 心像意味論

人間における自然言語理解の本質的な特徴は、言語の表層構造に依存しないイメージに基づく処理形態にある。即ちイメージおよびそれに対する操作の形式的記述が本質的に重要である。今日までに、自然言語の意味に関する研究¹⁾⁻³⁾が数多くなされてきている。しかしながら、従来の多くの自然言語意味論は、人間の心的過程を余り考慮していないため、その応用として構築される言語理解システムが人間と高次コミュニケーションを行うことができるようなもの、あるいはその原型となりうるか大いに疑問である。また、意味論それ自体のシンタックスおよびセマンティクスを実質的に与えていないため、局所的あるいは観念的にしかその妥当性を評価することができず、発展性や計算可能性に関して問題がある。また、このような問題を解決する一つ方法は、Miller, G.A.³⁾が指摘するように心理学的に根拠のある有限個の基本概念を導入し、それらの論理的構成により再帰的に自然言語概念の記述および解釈を与えるシンタックスおよびセマンティクスを構築することである。筆者はこのような状況を踏まえて、述語論理の基盤に立ち、自然言語の意味を人間の心像現象と対応づけて説明でき、計算にも適した理論(心像意味論(MIDST: Mental image directed semantic theory)と呼ぶ)の構築を行っている。

我々は人間の自然言語理解の基本的機構は図1に示すようなものであると考え、これを自然言語体系(のモデル)と呼んでいる。図1は自然言語(N)とそれが担う概念の体系(C)との対応づけ機構(自然言語解釈・表出機構)および、その概念が指示する対象の世界(W)との対応づけ機構(対象認知・指示機構)のみから成る人間の自然言語理解機構の原始的モデルである。勿論、詳細には概念体系(C)において概念構造を操作する機構(推論機構)など更に多くの複雑な機構が関与している。このモデルに従うと、我々の仕事はN、C、Wおよびそれらの間の対応関係を明確にし、計算可能な形式に整理することである。ところで、眼前に起こっている事件のような特定の事例に関する言及や理解結果として対象世界に対する働き掛けなどが存在する場合を除いて、人間は自然言語理解に際しては実際に対象世界Wに関係することはほとんどなくCのみにおいて処理を完了すると考えられる。特に、Wが空想世界の場合はCのみが処理の対象である。即ち、人間は自然言語文章で与えられる実在の事物同様に実

在しない事物(例えば「一角獣」)の心像(イメージ)まで想起し言及する事ができる。また、既知の事物同様未知(未来)の事物に関しても、文章を構成している全ての単語が既知ならば、単語の担うイメージを結合する事により想起する事ができる。結論として、人間の自然言語処理においては、自然言語文章が想起させるイメージおよびそれに対する操作が本質的である。このような観点から、筆者は、事物のイメージを実在と考える自然言語意味論、すなわち、心像意味論を提唱する。この立場では、イメージは人間の脳内に存在する事物であり、集合を形成し限量の対象となる。それぞれのイメージは属性空間における軌跡によって特徴付けられ、外界に実在する事物のイメージかどうかは、「物理的场所」に関する属性空間に軌跡が存在しうるか否かかで表示される。したがって、この場合の事物変項あるいは定項はイメージを互いに区別するための標識としてしか意味を持たないことになる。また、このイメージの集合の要素数(基数)は時間と共に変化しうるとする。但し、人間の感性経験と対応する属性値の個数だけは有限であり、構造を持った属性空間という抽象空間の座標として記述される。



図1 人間の自然言語理解の基本機構

図1の概念体系Cを客観的に取り扱うことを意図して2種類の道具を用意した。これらが図2に示すDおよびMである。Mは人間が概念として持っている対象世界のイメージの視覚的モデル、Dは形式的に概念を記述・操作するための体系(システム)である。一般的にはDは記号による意味・知識表現(格構造表現など)のシステム全てを指示している。本研究ではDの基盤として、概念記述に対して体系的な推論処理が可能となる述語論理を選択している。また、Mは人間に理解しやすい線図形(属性空間における軌跡)により、Dにおける形式的概念記述あるいはWにおける事物に対して、それぞれ直観的解釈または表現を与えるものである。すなわち、Nの要素である自然言語表現の内包をDは計算機向きに、Mは人間向きにそれぞれ表現する道具である。

心像意味論のセマンティクスはDを形成している応用述語論理のものである。一方、心像意味論のセ

マンテックスは、DとMの構成要素間の対応関係（Dの解釈）を与えるものである。Dには基本概念として原子軌跡式が導入されており、Mに置けるイメージの基本単位である属性空間の原子軌跡と対応が付いている。一般的な自然言語概念はDにおいて原子軌跡式の論理的結合列で表現され、Mにおける一般的な軌跡と対応付けられる。

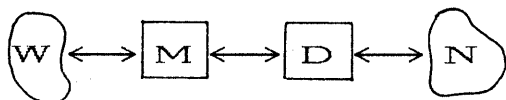


図2 心像意味論の枠組み

3. 心像意味論のシンタックス⁵⁾

自然言語の概念体系は人間の感覚知覚に依拠する部分が大きいと、本来曖昧で各人において異なる主観的なものとされている。しかしながら、これはそのような体系に含まれる単独の概念内容に関してであり、概念間に存在する構造および自然界の法則等の認識に関してはそうであってはならない。たとえば、色彩の概念および概念構造に関していえば、「赤」という色彩感覚（概念）は各人において異なるが、「青」より「橙」の方が「赤」に似ているなどという構造は万人に共通であると考えられる。また、「同じ人間が同一時刻に異なる場所にいることはない。」などという自然界における法則もそうであると思われる。もし、そうでないと人間同士の正常な情報の授受は不可能である。このことは、個人の持つ概念体系間の同型性(isomorphism)を仮定することになり、ここに本質的に曖昧とされている自然言語概念体系の論理的構成への可能性と意義が存在する。我々は、現在、人間が日常、感覚知覚できる範囲の世界に関する概念体系が論理的構造を持った事物概念の集合と法則の集合より成っていると考え、それらをコンパクトで、しかも計算可能な形式で記述することを当面の目標にしている。

形式的概念体系Dの定式は純粋述語論理（個体と述語については変項のみをオブジェクト記号に含み、定項をもたない）のものとして上述の自然言語概念体系に関するものから成っている。自然言語の担う概念のモデルとして属性空間およびそこにおける軌跡を導入しているため、属性空間との対応において7種類の個体を用いる点およびその空間における原子軌跡を指示する述語定項(L)が存在するのが特徴である。

3.1 構成記号

形式的体系Dを構成するオブジェクト記号の一覧を下の(i)~(ix)のように与える。この他にメタ記号をいくつか導入する。

- (i) 文章変項：適宜、導入する
- (ii) 論理的結合子： $\wedge, \vee, \supset, \equiv, \sim, \forall, \exists$
- (iii) 補助定項： $\cdot, (,)$
- (iv) 述語変項：適宜、導入する
- (v) 個体変項
 - a) 事物変項： x, y, z または、これらの添字付き
 - b) 表現変項： r または、これらの添字付き
 - c) 情報変項： i または、これらの添字付き
 - d) 属性変項： a, b, c, d または、これらの添字付き
 - e) 属性値変項： p, q または、これらの添字付き
 - f) パターン変項： g または、この添字付き
 - g) 基準変項： k または、この添字付き
- (vi) 文章定項：適宜、導入する
- (vii) 述語定項： $L(\#, \#, \#, \#, \#, \#, \#), \mathcal{J}_i(\#, \#)$ (但し、 $i \in \{0, \pm 1, \pm 2, \pm 3, \pm 4\}$), $\# = \#, \# < \#$, 他適宜、導入する
- (viii) 個体定項
 - a) 事物定項：適宜、導入する
 - b) 表現定項：適宜、導入する
 - c) 情報定項：適宜、導入する
 - d) 属性定項： A_1, A_2, \dots
 - e) 属性値定項：適宜、導入する
 - f) パターン定項： Gt, Gs
 - g) 基準定項：適宜、導入する
- (ix) 関数定項： $\mathcal{D}(\#)$, 他適宜、導入する

但し、 $\#$ は述語または関数が支配する項を表している。また、種々の定項は議論世界(日常世界や気象通報世界など)の違いによって、その内容が異なってくる。

3.2 軌跡式の形成

属性空間における最も基本的な軌跡（原子軌跡と呼ぶ）と対応づけられる式を原子軌跡式、原子軌跡式の論理的結合列を単に軌跡式と呼んでいる。原子軌跡式の形成規則は以下のものである。ただし、実体とは事物、表現、および情報を意味する。

$$L(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7) \quad (3-1)$$

式(3-1)は次の条件を満足するとき式(formula)である：

「 ω_1, ω_2 が実体変項または実体定項、かつ ω_3, ω_4 が属性値変項または属性値定項または実体変項または実体定項、かつ ω_5 が属性変項または属性定項、かつ ω_6 がパターン変項またはパターン定項、かつ

ω_7 が基準変項または基準定項または実体変項または実体定項である。」

原子軌跡式の基本形は式(3-2)で与えられる。

$$L(x_1, x_2, p, q, a, g, k) \quad (3-2)$$

このとき解釈は次のようになる：

「実体 x_1 の作用により実体 x_2 の属性 a が時間的 ($g=G_t$)あるいは空間的 ($g=G_s$)に変化 ($p \neq q$)あるいは持続 ($p=q$)する」

(但し、 p および q は基準値 k に対する相対値である。また、 ω_3, ω_4 または ω_7 の場所に実体が入った場合は、その時点で、その実体が持っている属性値に等しいことを表す。)

3.3 時間的結合子

属性空間における軌跡は原子軌跡単位で時間軸に関して分節される。その際、人間の認知過程においては、(時間の属性空間を除き)時計が逐一参照されるわけではないので分節の起こる時点を値として認知・記憶することはありえない。それよりも、各分節間の時間的位相関係(前、後、間など)が論理的関係と一緒に認知・記憶されると考える方が自然である。このような考察に基づき、軌跡式間の論理および時間的関係を同時に表示する時間的結合子 C_i を以下のように定義(Dと表示)し導入する。

$$D_i: X_1 C_i X_2 \triangleq (X_1 C X_2) \wedge \mathcal{J}_i(\mathcal{D}(X_1), \mathcal{D}(X_2)) \quad (3-3)$$

但し、 $X, C, \mathcal{J}_i, \mathcal{D}$ はそれぞれ軌跡式、論理的結合子(\wedge, \vee, \dots)、時間関係、時間帯抽出関数である。時間帯抽出関数 \mathcal{D} は軌跡式の時間帯を陽に取り出す働きをする。

4. 心像意味論のセマンティクス

心像意味論のシンタクスは応用述語論理のものである。通常、述語論理のセマンティクスは静的な集合論に基づいている。しかしながら、前述したように、自然言語概念は人間の動的認知理解過程の産物であるため、そのような方法では時間的連続変化に対応することが困難である。例えば、Shoham, Y.⁷⁾の時間論理では(例4-1)の意味表現は式(4-1)のようになり、その式の解釈は式(4-2)で与えられる。

(例4-1) HOUSE17 was red over the interval $\langle t_1, t_2 \rangle$.

$$\text{TRUE}(t_1, t_2, \text{COLOR}(\text{HOUSE17}, \text{RED})) \quad (4-1)$$

$$S \vdash \text{TRUE}(t_1, t_2, \text{COLOR}(\text{HOUSE17}, \text{RED})) [VA] \text{ iff } \langle MVA(\text{MVA}(t_1), \text{MVA}(t_2), \text{HOUSE17}), \text{MVA}(\text{MVA}(t_1), \text{MVA}(t_2), \text{RED})) \rangle \in M_5(\text{MVA}(t_1), \text{MVA}(t_2), \text{COLOR}) \quad (4-2)$$

ここで、問題にしたいのは式(4-2)(意味表現の解釈)である。時間帯ごとに、個体記号に個体、関係記号にそれを満たす個体の組合せの集合の割り付けを行っている。これは、Kripke, S.⁸⁾の可能世界モデルの宿命であり、これをあらゆる時間帯に関して行うのは原理的には可能でも実質的に不可能である。しかも、世界の設定が離散的にならざるを得ない。例えば、"HOUSE17"の色が赤(t_1)から黄色(t_2)に徐々に(連続的に)変化した時、途中の橙色(限り無く赤に近いものから、限り無く黄色に近いものまで)の取扱いが可能でない。心像意味論では、属性値がその類似度に応じて配置されている属性空間およびそこにおける線分の軌跡と一対一に対応する原子軌跡式を導入することにより、連続変化の記述を可能にしている。さらに、何度も述べたようにこのような、自然言語文の真偽判定だけを目的とする外延的意味解釈は自然言語理解処理には余り役に立たない。

5. 情報体系とコミュニケーション

「まえがき」でも述べたように、コミュニケーションは、情報体系(information systems)間の相互作用と定義できる。ここでの相互作用は、複数の情報体系が種々の媒体(言語、画像など)を介して情報(知識、信念など)を交換することにより、共通部分情報体系を形成することである。

まず、明確にすべきことは、1)情報、2)情報体系、3)情報媒体、および4)共通部分情報体系の定義とその性質である。

情報は、媒体によって表現・伝達される意味内容(すなわち概念)でありセマンティクスの取扱い対象のことであり、ここでは心像(イメージ)に帰着されるものである。

情報体系は、図3に示すような、1)対象世界(W)、2)概念体系(C)、3)表現体系(R)およびそれらに係わる処理体系である(心像意味論の枠組みでは表現体系は自然言語体系(N)となっている)。



図3 情報体系の概念図

我々は、情報体系における処理体系(対象認知・指示、表現理解・生成、概念推論など)として個々の情報体系に依存しない普遍的なものを想定してい

る。したがって、個々の情報体系を特徴づけるのは、対象世界、概念体系および表現体系である。そこで、以後は、情報体系(sで表示)を式(5-1)で示すような3字組で表現する。

$$s = (W, C, R) \quad (5-1)$$

情報媒体は、ある情報体系が情報を他の情報体系に伝達するための表現手段である。ここでは、情報体系に共通のシンタックスとセマンティクスを持った言語が主となり、そうでない画像や音声などが補助的に働く多重媒体(multimedia)を想定している。このような媒体は、言語を核として図3に示す体系(すなわち表現体系(R))をなしている。

共通部分情報体系(S_p)は式(5-2)のように3字組で形式化される。ただし、pは式(5-3)に示すようにコミュニケーションに参加しているn個の情報体系(s)の集合である。

$$S_p = (W_p, C_p, R_p) \quad (5-2)$$

$$p = \{s_1, \dots, s_n\} \quad (5-3)$$

W_pは、s_mの間で共通的に認知された対象世界Wの一部であり、式(5-4)で示すように事象(e)の集合である。通常のコミュニケーションでは、全てのs_mがWの全てを直接体験できることは希である。それよりも、概念的にだけコミュニケーションを行っている場合、すなわち、次に述べる概念体系(C_m)だけを得るレベル(心像レベル)のコミュニケーションがほとんどである。

$$W_p = \{e_1, \dots, e_k\} \quad (5-4)$$

C_pは、W_pの事象e_jに関してpの成員s_mが得た情報(i_{mj})よりなる概念体系(C_m)の集合であり、式(5-5)で表現する。C_mは、s_mのW_pに関する理解結果と考えることができる。ここで注意すべきことは、情報交換の結果生じた個々の概念体系は、必ずしも同一のものではないということである。ただし、個々の体系では無矛盾が保たれ、また、3章で述べたように体系間には同型性が存在しているとする。

$$C_p = \{C_1(i_{11}, \dots, i_{1k}), \dots, C_n(i_{n1}, \dots, i_{nk})\} \quad (5-5)$$

R_pは、C_pが成立するまでにW_pに関してpの成員s_mが情報を発信するのに用いた表現体系R_m(言語や画像表現の構造体)の和集合であり、式(5-6)のように表現できる。R_pには、pの成員同士が自己紹介したり、互いの矛盾点などを問い質したり、あるいは、訂正したりした表現など全てが含まれる。

$$R_p = \bigcup_{m=1}^n R_m \quad (5-6)$$

図4は、2つの情報体系間の情報交換によって生じる共通部分情報体系の概念図である。

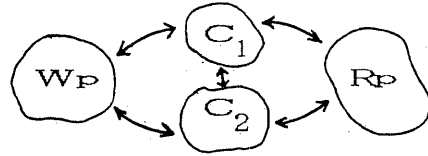


図4 共通部分情報体系の概念図

6. 言語主導型コミュニケーション

情報体系間における情報交換で主導的役割を担うのは、それらの情報体系に共通のシンタックスとセマンティクスを持ったもの、通常、言語である。それら二つのもを明確には持たない媒体(例えば画像)は、それを受けとる側にしてみれば、どこに意味があるのか理解できないのが普通であるので副次的な情報伝達媒体と考えることができる。ここで議論するコミュニケーションは、そのような事実を考慮し、言語を核とする多重媒体(multimedia)によるものである。

6.1 メッセージ

情報伝達の最小単位をメッセージ(message:通達)と呼ぶことにすると、コミュニケーション中におけるメッセージ(mで表示)は式(6-1)のような4字組で表現できる。

$$m = (D, A, H, B) \quad (6-1)$$

ただし、この式で、D、A、HおよびBは、それぞれ、通達時間帯、通達元、通達先および通達主部を意味する。これらの重要性は明らかである。R_pはこのようなメッセージの集合であるが、式(6-1)のように分節はなされていない。分節は後に処理体系が行う。

通達時間帯は、メッセージの発生順序とほぼ等価であるが、割り込み(interruption)や複数の情報体系が、時間的に重複して通達を行う場合などを想定しているためである。このような時間帯は、陽に表現されないのが通常であり、その抽出は他の記録に頼ることになる。

通達元は、話者や筆者などである。人間同士の通常の会話では、その参加者全員が顔や声などで話者認識を行っており、これも通達元自身により陽に表現されることは希である。

通達先は、聴者や読者である。人間同士の通常の会話では、呼びかけなどにより表現される。

通達主部は、対象世界に関する言及、すなわち、事象の陳述などである。この通達主部の細かい分析は、次の6.2で行う。

図5は、このような形式化が想定している情報交換の例を視覚表現したものである。この図において、"/"で区切られた部分がメッセージである。[t₁, t₂]が通達時間帯、": "の直前が通達元、下線部が通達主部、残りの部分が通達先である。

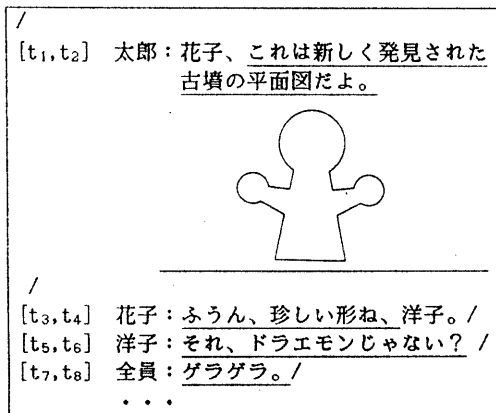


図5 情報交換の例

6.2 通達主部

コミュニケーションで主導的役割を担う言語は、メッセージの通達主部において、文の種類という観点からその役割を分類することができる。結論から先に述べる。言語の働きを分析すると、通達主部は、式(6-2)に示すように、2つの内容よりなることがわかる。

$$B = (E, T) \quad (6-2)$$

この式で、Eは、対象世界における事象の情報(概念)、Tは、Eに関する処理あるいは作業の通達元から通達先への依頼の情報である。また、通達主部の表現そのものの表示は、記号"で囲むことで式(6-3)のように一種の実体(表現)として論理式に取り込む。

$$L(\text{洋子}, r, \text{洋子}, y, A_r, G, t, k) \wedge \text{inf-sys}(y) \wedge r = \text{"花子、これは新しく発見された古墳の平面図だよ。"} \quad (6-3)$$

ただし、inf-sys および A_r は、それぞれ、情報体系および表現の存在位置を意味し、式全体としては表現の伝達を意味している。

文の種類を、平叙文、疑問文、命令文、感嘆文の4つに大分類すると、これらはTの情報の種類に関係があることがわかる。

1) 平叙文の場合

Eの真偽を確かめよ。偽または真偽不明の場合は、問い合わせるか信ぜよ。

(例7-1) 明日、雨が降る。

2) 疑問文の場合

Eの真値またはEを真とする情報を答えよ。

(例7-2) 今、5時ですか。今、何時ですか。

3) 命令文

Eが真となるように作業を行え。

(例7-3) 新聞を取って来い。

4) 感嘆文

Eの真偽を確かめよ。偽または真偽不明の場合は、問い合わせるか信ぜよ。

(例7-4) なんて速い車でしょう。

以上のTに係わる処理は、既に我々の自然言語理解システム²⁰⁾で実現されているものである。感嘆文の取扱いは実質的に平叙文と等価である。また、問い合わせは、情報交換のあらゆる場面で可能である。

7. むすび

心像意味論に基づき多重情報媒体によるコミュニケーションの概念化を試みた。このような研究を成功させるには、結局、人間の持っている常識をシステムに持たせることが不可欠である。人間の常識を論理的に構成しようとする試みは、国の内外に散見される(例えば、Kripke, S., Shoham, Y. など)が、実際の人間の心的過程を考慮したものになっていないため、局所的には妥当性を保てるが、発展性に問題がある。このような研究の問題点は、物事の時空間における連続的変化に関する概念の取扱いの難しさにある^{11) - 13)}。すなわち、我々が指摘するように、人間の感覚野に投影された時空連続の世界像(イメージ)の記述とその記述に基づく処理方法の開発が本質的であると考えられる。今後は、自然言語データの大量分析^{16, 17)}、心理実験等¹⁵⁾および多重情報媒体による知識の獲得運用過程の分析^{19), 21), 22)}などの結果を理論^{5, 9, 10, 18)}および理解システム^{6, 14)}に帰還することにより目標達成を図りたい。

なお、本研究の一部は、文部省科学研究費補助金(重点領域研究 05241105)によっている。

参考文献

- 1) Katz, J.J. et al : The structure of a semantic theory, Language, 39, 2 (1963)