知識の表現に関する一考察†

——情報理論的観点から——

大須賀 節 雄††

知識ベース・システムにおける知識表現の問題は知識情報処理の最も基本的なものであるにもかかわらず，その理論的な検討が十分進んでおらず，かなり経験的な方法で扱われている．本論文は，機械，電気・電子回路，材料，ソフトウェア等の設計や製作などの開発支援が可能なことを今後の知識ベース・システムの条件とし，情報量の概念を導入して知識表現のもつべき性質を考察する．知識表現としては対象モデルの記述が可能であることを要し，これには構造の記述とその構造のもつ性質等の記述を同時に行えることが必要であるが，さらに開発のプロセスが，所要の性質をもつ構造を試行錯誤によって求めることであることから，構造の操作によって，その構造のもつべき性質（要求）の表現が影響を受けず，しかもつねにその構造を正しく評価しうるものでなければならない．さらに，見方を変えると設計等は情報を逐次的に与えてモデルを精細化するプロセスであり，小刻みに情報の与えられる表現であることが望ましい．この論文では知識表現の情報量を定義し，定量的評価のもとで上述の要求を満たす知識表現として，論理表現の一変形である多層論理が適していることを示す．

1. まえがき

知識工学は近年多くの話題を提供しているが，これを定着した技術と見なすには，まだあまりに未知の部分が多い．最も基本的な知識表現の問題ですら，多くの提案があり，根拠のある比較がなされていないのが現状である[1),2)]．この問題は知識の利用，獲得，管理，保守など知識工学のすべての面にわたって検討を経て定めるべきものであるが，これらの諸検討がいまだ不十分なゆえであろう．そしてさらにいえば，知識表現自体の理解が十分でないことがその大きな原因であるように思える．

本論文は，知識表現の問題に対する一つの見方として情報量の概念を導入し，それを通して知識表現のもつべき性質を考察する．しかる後，望ましい性質を備えた知識表現の一形式を示す．なおこれに先立って，以後の議論の根底となる筆者らの考え方について述べる．

第一に，以下では知識情報処理に今後期待される応用として設計支援（機械，電気・電子回路，材料その他）やソフトウェア開発支援などの開発型の問題を考える．この問題はこのほかの応用，たとえば医療等への応用とは見かけ以上に異なる．後者では，利用する知識そのものに曖昧性が含まれる以上，知識ベース・システムから得られる結論にも曖昧性が含まれることを容認するという立場に立つのにたいし，開発型の問題は，最終結果として機械的に製造され，あるいは機械を駆動するための製品を創り出すものであり，原則として結果にたいして曖昧さの残ることは許されないし，結果の正当性が保証されたものであることが望ましい．このためには，結果を生成する知識ベース・システム自身が無矛盾性のチェックの可能な方式となっていることが必要である．知識ベース・システムの特徴の一つは知識のモジュール性が高いという点にあり[1)]，モジュラリティは個々の知識素（知識ベースを構成する個々の規則や事実をここでは知識素と呼ぶ）が相互に独立であること，あるいは全体的な規制から自由であることを要求するのにたいし，無矛盾性とは知識ベース全体の規制を意味するから，この二つの概念には相容れない性質があり，どこかで妥協が必要になる．現在，多くの知識表現が提案されているが[1),2)]，このような観点から，個々の知識素のみでなく，その集積された体系全体としての性質や振舞いを解析し，議論できるだけの背景をもっている方法としては筆者の知る限り述語論理のみである．この意味で，本稿では開発問題への知識表現としては述語論理もしくは述語論理と対応づけられるものを考える[3),4),9)]．

第二に，開発とは与えられた要求を満たす対象のモデルを創成するプロセスと考える．開発作業に知識ベース・システムを適用し知的な支援システムを実現するためには，対象のモデルをシステム内に構築する必要がある．ここに，広範囲の問題に適用できる汎用

† A Consideration to Knowledge Representation—An Information Theoretic View by SETSUO OHSUGA (Institute of Interdisciplinary Research, Faculty of Engineering, University of Tokyo).

†† 東京大学工学部境界領域研究施設

のモデル表現の問題が生ずる．本稿はモデル表現を論ずることが直接の目的ではないから詳細は省くが，モデル記述の情報としては，①モデル構造，②モデルおよびその構成要素の性質や相互の関係の記述の２種類の情報が必要である[6]．構造情報は要素とそれによって構成される構造体の関係を示すものであり，通常，データ構造として表されている．②の記述情報は述語によって表現されるが，本稿では一階述語の一変形である多類論理を発展させた形で，構造を含む論理形式を定義する．この方法は特定の構造を述語内に陽に含まずに構造の参照が可能であるため，構造の生成を，評価のための構造の参照と分離できる点で後述のように開発問題に向いている．

第三に，人間は知識を取り込むと同時に，つねにそれを洗練しているから，知識の表明にも多くの微妙に異なる段階がある．この知識を正しく受け取るには，知識ベース・システム自身が微妙な表現の可能なものであることが望ましい．このような知識表現の評価を行うには評価尺度を定義する必要があり，本稿では情報理論に基づく定量的尺度を導入する．定量的尺度は知識表現の評価のみでなく，知識の利用や獲得のプロセスにおいても有用である．しかし本稿における主たる目的は，このような定量的評価を通して知識表現の性質をよりよく理解し，それによって知識表現にたいする解釈の曖昧性を減少することにある．

以下，２章においては述語論理式の情報量を定義する．３章において知識表現とするための述語論理の拡張を述べる．拡張された論理を多層論理と呼ぶ．４章では多層論理式の情報量が，対象の構造に応じて広い範囲に変化することを示し，また多層論理式が開発型問題の知識表現に適することを示す．５章では情報量の概念が利用上からも有用であることを二つの場合で示す．

2. 述語論理式とその情報量

知識ベース・システムでは扱う問題領域に関するすべての対象の集り U を含む世界のモデルを計算機内に構成し，これを用いて問題解決をはかる方式を確立することが必要である．世界のモデル化に際し，対象の性質や関係の記述手段として，有限の述語記号の組 P をもつ述語論理を用いることとする．通常，一階述語論理が用いられるが，これは世界のモデル化には必ずしも便利な手段ではなく，現実の計算機の技法，とくに対象モデルを扱う上でデータ構造の利用に適したものとすることが望ましい．以下では論理形式の変更によってこの要求を満たすことを考える．なお，以下で世界モデルとは扱う問題領域の世界に関するすべての情報の集り，すなわち知識ベースそのもの，対象モデルは設計等でこれから作り上げようとする対象を記述する情報の集りであり，これは当然，世界モデルの一部である．

一階述語論理式を一般形で

$$(Q_1x_1)(Q_2x_2)\cdots(Q_mx_m)M(x_1, x_2, \cdots, x_m) \quad (1)$$

と表す．x_i, Q_i, $i=1, 2, \cdots, m$ はそれぞれ変数および限量記号（$\forall$ または $\exists$）を表す．$M(x_1, \cdots, x_m)$ は $F_j \in P$ なる述語記号を用いた有限個の述語から論理結合子 $\wedge$(AND), $\vee$(OR), $\sim$(NOT), $\rightarrow$(IMPLY), $\leftrightarrow$(EQUIVALENT) を用いて作られる論理式である．定数と変数を項とするが，本稿では関数は述語で表すことにする．論理式としては自由変数を含まない閉論理式（文）を対象とする．

実世界では（ある性質や関係を示す）対象の集合が定義されるが，通常の一階述語論理ではこれを性質・関係として（内包的に），述語によって表している．たとえばある対象 x が集合 d に属することは，$D(x)$ なる述語を満たすことと同義であり，集合の代りにこれを用いることとしている．これは計算機内に実世界のモデルを作る上で必ずしも便利な方法ではない．

これにたいし，集合の概念に直接的に対応できる述語論理として定義されたのが多類論理 (Many Sorted Logic) である．この論理系では類 (sort) の概念を述語の構文に含め，x が $D(x)$ を満たすことを接頭部の (Qx/d) によって表す．Q は限量記号（$\forall$ または $\exists$），d は $D(x)$ を満たす x の集合である．すなわち，一階の一類述語では

$$(\forall x)[D(x) \rightarrow F(x)] \text{ あるいは}$$
$$(\exists x)[D(x) \wedge F(x)] \quad (2)$$

で表される，“すべての x について $D(x)$ なら $F(x)$ である”あるいは“ある x について $D(x)$ かつ $F(x)$ である”は，多類論理のもとではそれぞれ $(\forall x/d)\ F(x)$, $(\exists x/d)F(x)$ と書かれる．たとえば一階一類述語で $(\forall x)[\mathrm{MAN}(x) \rightarrow (\mathrm{MORTAL}\ x)]$,（“すべての人は死す”）は多類論理では $(\forall x/\mathrm{man})\mathrm{MORTAL}(x)$ である．多類論理を実世界に適用するとき，類は世界のある部分集合に対応づけられる．以下，とくに混乱のない限り，多類論理における類と実世界においてこれと対応する集合を同一の記号 (d) で表す．

集合 d と述語記号 F を考えよう．以下では d は

有限とする．簡単のため F を一変数述語 $F(x)$ とする．$F(x)$ は対象に関する記述を与えるが，見方を変えると，これは集合 d の要素を，F を満たすものとそれ以外のものに分割する働きをする．以下 $F(x)$：真および $F(x)$：偽をたんに $F(x)$ および $\bar{F}(x)$ と表す．

さて，F を用いた述語論理式（以下たんに式という）が与えられるものとして，その前後における d の状態というものを定義する．事前では，$x \in d$ なる任意の x が $F(x)$ か $\bar{F}(x)$ かが不明であり，事後では d の要素のどれかまたはすべてについて F または $\bar{F}$ が明らかになるという変化がある．

$d=\{a_1, a_2, \cdots, a_N\}$ とする．a_i, $1 \le i \le N$, は d の要素である．F に関する d 自体の状態を，d のすべての要素 a_1 に関して $F(a_i)$ か $\bar{F}(a_i)$ の組合せで表すとする．F が与えられる以前には，d の状態は S_1：$F(a_1) \wedge F(a_2) \wedge \cdots \wedge F(a_N)$ から S_{2^N}：$\bar{F}(a_1) \wedge \bar{F}(a_2) \wedge \cdots \wedge \bar{F}(a_N)$ まであらゆる可能性を含む．このような可能な状態の集合を S_F^1 とし，これを

$$S_F^1=\{F_{\bar{a}_1\bar{a}_2\cdots\bar{a}_N},\ F_{\bar{a}_1\bar{a}_2\cdots\bar{a}_{N-1}a_N}, \cdots F_{a_1a_2\cdots a_N}\} \quad (3)$$

と表す．$F_{\times\times\bar{a}_i\times\times a_j\times\times}$ は $\bar{F}(a_i)$，$F(a_j)$ なることを添字によって表したものである．集合 x の濃度（要素数）を $|x|$ と表す．$|d|=N$ とすると，$|S_F^1|=2^N$ である．

S_F^1 の各要素は相互に排反であり，かつ網羅的である．S_F^1 を F に関する d の事象系と呼ぶ．F の記述が与えられる以前には d はこのなかのどの状態にあるかがまったく不明であるから，この事前確率は等しく，2^{-N} とする．

S_F^1 に情報理論の情報源のエントロピの概念を導入する[10]．これを $I_{S_{F_1}}$ とすると，

$$I_{S_{F_1}}=-\sum_{S_F^1} 2^{-N}\log 2^{-N}=N \quad (4)$$

である．

次に，ここに論理式で表された知識素が与えられたとする．多類論理のもとでは，この可能な表現形式は $F(a_i)$, $a_i \in d$, $(\forall x/d)F(x)$ および $(\exists x/d)F(x)$ の3種である．これらはいずれも S_F^1 内の可能な状態を限定する作用をもつ．まず $F(a_i)$ を考えよう．一般性を失わずに $i=1$ とすることができる．$F(a_1)$ を与えることは S_F^1 の状態のなかで $F_{\bar{a}_1\times\times\cdots\times}$ の形の状態はすべて生じえないことを宣言したことを意味するが，$a_2, \cdots, a_N$ については何も述べていないから，この事後の状態集合を S_F^2 とすると，

$$S_F^2=\{F_{a_1\bar{a}_2\cdots\bar{a}_N}, \cdots, F_{a_1a_2\cdots a_N}\} \quad (5)$$

である．$F_{\times\times\cdots\times}$ の形式の表現において，与えられた知識素により固定された部分を (,) でくくり，固定されず a_i と $\bar{a}_i$ のあらゆる組合せが取られる部分を [,] で示す記法を用いると，S_F^1, S_F^2 はそれぞれ

$$S_F^1=\{F_{[a_1\cdots a_N]}\},\quad S_F^2=\{F_{(a_1)[a_2\cdots a_N]}\}$$

と表される．S_F^2 の濃度 $|S_F^2|$ とエントロピは

$$|S_F^2|=2^{N-1} \quad (6)$$

$$I_{S_{F_2}}=N-1 \quad (7)$$

になる．$I_{S_{F_1}}$ と $I_{S_{F_2}}$ の差 $(=1)$ は知識素 $F(a_1)$ のもつ情報量に相当する．以後知識素のもつ情報量を K と書く．

同様にして，$(\forall x/d)F(x)$ は d のすべての要素 a_i, $1 \le i \le N$, について $F(a_i)$ であることを示すから，

$$S_F^2=\{F_{(a_1a_2\cdots a_N)}\}, \quad (8)$$

$$|S_F^2|=1 \quad (9)$$

$$I_{S_{F_2}}=0 \quad (10)$$

$$K=I_{S_{F_1}}-I_{S_{F_2}}=N \quad (11)$$

$(\exists x/d)F(x)$ では $F_{\bar{a}_1\bar{a}_2\cdots\bar{a}_N}$ 以外はすべて可能性として残されるから

$$S_F^2=S_F^1-\{F_{\bar{a}_1\bar{a}_2\cdots\bar{a}_N}\} \quad (12)$$

$$|S_F^2|=2^N-1 \quad (13)$$

$$I_{S_{F_2}}=\log(2^N-1) \quad (14)$$

$$K=N-\log(2^N-1) \quad (15)$$

となる．これからわかるようにこの情報量は非常に小さく，知識素として与えることは効果は少ないが，5章に述べるように，質問の表現としては好都合である．なお表現としては上記3種の基本形式のほかに，F を $\sim F$ に置き換えたものも基本形式になりうるが，情報量に関しては F でも $\sim F$ でも同じなので省略する．

上述の考え方は2変数以上の述語に関しても同様に適用される．たとえば $(\exists x/d_1)(\forall y/d_2)\ F(xy)$, $|d_1|=M$, $|d_2|=N$ なる述語を考えよう．この情報量は $d_1 \times d_2$ の空間のなかで，この式を成立させるあらゆる要素の組を求めることにより求まる．このような組の濃度は $|S_F^2|=2^{M\times N}-(2^N-1)^M$ であり，このエントロピ $\log[2^{M\times N}-(2^N-1)^M]$ を事前のもの $M\times N$ より差し引いて次のように求まる．

$$K=M\times N-\log[2^{M\times N}-(2^N-1)^M] \quad (16)$$

3. 述語表現の拡張—多層論理

3.1 モデル化のための知識表現

前章の結果からすぐ気がつくのは，(a) 一階多類論

理がきわめて不完全な（エントロピの大きな）表現法を許していること，(b)許されている表現間の情報量が非常にかけ離れていること，である．ここで意図している開発型の問題は逐次的にモデルに情報を与えて，この構造を精細化してゆくプロセスであり，これはモデル全体の情報量を増大することといえる．(a)の性質は人間にとって記述を容易にするが，情報量を一度に飛躍的に増大することは困難で，小幅に増加してゆくほかない場合が多い．この点(b)の性質に制約される従来の多類論理のみでは不便であり，小幅の増分で情報を与えうる知識表現に変えることが望ましい．しかし，そのために知識表現の構文規則が著しく複雑化し，論理の理論体系が崩されるものであってはならない．

さらに，モデル表現は構造と，性質・関係等の記述から成ること，開発とは望ましい性質・関係等の記述が要求として与えられ，それを満足するモデル構造を見いだすプロセスであるという考え方についてはすでに述べたが，この実際的な方法は，(1)仮の構造を造り，(2)それに基づいて性質等を評価し，(3)これが要求と一致するか否かを調べ，もし一致しなかったら構造を修正する，という繰り返しのプロセスである．このプロセスにおいて，目標に達するまで構造表現はさまざまに変えられるが，性質等の記述は不変に保たれねばならない．このため，性質等の記述内には具体的な構造表現を含まず，構造と性質等の記述を分離することによって構造を独立に操作しうること，しかし，作られた構造にたいしてつねにその評価が可能なものであるような表現形式が必要である．

以上の二つの要求を満たすための知識表現として，以下では多類論理の概念を発展させ，構造表現と密接な関係を保ちながら，これと独立な述語論理を考える．

3.2 多類論理による構造

一階の一類論理と多類論理の相違は，前者が項に関して構造もしくは相互の関係を含まないのにたいし，後者は項間に類という概念を通して一定の構造を含み，この論理を実対象の集りである世界に適用することによって，この世界の中に対象の構造（集合-部分集合関係）を定義していることである．多類論理における類との対応によって，世界の中に $\cdots u\subset v\subset w\cdots$ のような，集合関係の多段構造を定義することができる．これは構造の一形態であり，一般的概念構造を構造化して表すのに便利である．たとえば，生物⊃哺乳動物⊃人間…のような構造を骨格とし，そこに知識素を関連させて（たとえば $(\forall x/\mathrm{man})\mathrm{MORTAL}(x)$ を“人間 (man)”と関連づけて）記憶する形で知識構造を構成することができる．これはいわば世界のモデルのなかで概念構造としての静的な知識ベース構造であり，上記構造は世界内の対象の分類によって自然に作られるものとしてこれに役立っている．

しかし，世界モデルのなかでも，実際の物としての，いわば人工的な構造を現し，しかも動的に操作される対象モデル構造部分の表現には上記構造は不十分で，これとまったく別の形式の構造を必要とする．

3.3 論理の拡張—多層論理

$d=\{a_1, a_2, \cdots, a_N\}$ なる集合と，その部分集合，$b_i=\{a_{i1}, a_{i2}, \cdots, a_{im_i}\}$, $i=1, 2, \cdots, k$, を考えよう．さらにこの各部分集合を，ある特定の実対象と関連づける機構を想定しよう．たとえば，各 a_i は三次元空間内で定義された線分を表すものとする．このとき，**図1**のように，この線分の組 b_j によって面 s_j を表すという手法はこれまで形状処理の分野で多用されてきた．同様に，集合 $c=\{b_1, b_2, \cdots, b_6\}$ によって多面体 h を表す．

この構造関係が前記のものと異なる点は，$a_i\in s_j\in h$ なる集合関係があることである．この関係が実体の構造を表すとするのは一つの約束であり，これが対象の

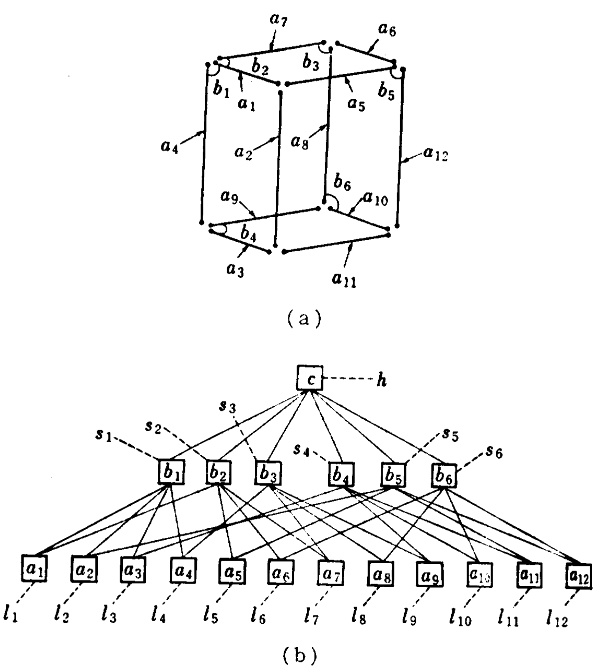


図1 3次元構造表現
Fig. 1 Representation of a three dimensional structure.

分類によって自然に生ずる前節の構造と異なる．またこの約束に相当するものが，実体たとえば s_j と，集合 b_j との対応づけ機構である．この対応づけのもとで，図1(b)の構造は $a_i \in b_j \in c$ と解釈される．この集合関係に基づく構造を体系化するためには数学的概念として新しくべき集合を導入する必要がある．

集合 d が与えられたとき，$*d$ によって d のあらゆる部分集合の集りから空集合を除いた集合を表す．$d=\{1,2,3\}$ のとき，$*d=\{(1),(2),(3),(1,2),(1,3),(2,3),(1,2,3)\}$ である．さらに $*d$ が一つの集合であるから，$*(*d)$ を同様に定義できる．これを $*^2d$ と書く．これを $*^n d=*(*^{n-1}d)$ によって一般化する．

これを用いて $c \in *^2d$ とすると，c の中に，$c \ni b_j \ni a_i$ なる関係にある a_i, b_j が存在する．すなわち $c \in *^2d$ なる c を定めることは一つの構造を与えることである．一方，$*^2d$ は上記の関係をなすあらゆる可能な構造の集りであり，c はその一要素である．

この与えられた構造の構成要素についてある記述を与えるとしよう．たとえば“面 b_j の中のある線分 x の長さは 5(cm) である”なる記述がなされたとする．これはどの線分が長さ 5 cm であるかを特定するだけの情報はないが，その線分が“面 b_j を構成する線分のどれかである”という限定をしている点で立派な情報である．これを $x \in b_j$ を考慮して多類論理を用い，

$(\exists x/b_j)$ (LENGTH x 5)

と表すのは不自然ではない．(LENGTH u v) は“u の長さは v である”と読まれる述語とする．同様に，

$(\forall y/c)$ (AREA y 20)：c (多面体) のすべての面の面積は 20(cm^2) である，

も可能である．(AREA u v) は“u の面積は v である”なる述語である．これらを一緒にして，

$$(\forall y/c)(\exists x/b_j)[(\text{AREA } y\ 20) \wedge (\text{LENGTH } x\ 5)] \quad (17)$$

も妥当な論理式である．

この表現で y と b_j はともに面を表しているが，相互に何の関連もない．しかし，図1(b)の構造に基づいて，“多面体 h のすべての面 y について，y の面積は 20(cm^2) であり，かつその y のなかに長さ 5(cm) の線分 x がある”という表現をしようとしたら，(17)式で y と b_j を連結することができねばならず，

$$(\forall y/h)(\exists x/y)[(\text{AREA } y\ 20) \wedge (\text{LENGTH } x\ 5)] \quad (18)$$

なる表現が許されねばならない．この式は(17)で b_j を，接頭部の先頭の y で置き換えたものであるが，これを許すためには多類論理において，類自身を変数とするという構文の拡大が必要である．このように拡大した論理系を以下，多層論理と呼ぶ．ここで，固定された類と，変量としての類が生じ，この区別は文脈から求まるが，理解を早めるために補助的な記号 $\sharp$ を導入し，固定類に付す．$\sharp$ は多層論理の定義のための基本記号ではない．これを用いると，たとえば(18)は

$$(\forall y/\sharp h)(\exists x/y)\ [(\text{AREA } y\ 20) \wedge (\text{LENGTH } x\ 5)] \quad (18')$$

のように表される．固定類が世界内に集合-部分集合の関係のデータ構造化を可能にしたと同様，変量類は多段の集合-要素関係のデータ構造化を可能にする．結局，多層論理のもとでは世界内に2種類の多段構造が用いられる．第一は $u \subset v \subset w \cdots$ の構造で，これを，以下，I型構造と呼ぶ．I型構造は概念構造の表現に適し，長期記憶としての知識ベースの構造表現として用いられる．第二は $\cdots x \in y \in z \in \cdots$ の構造で，II型構造と呼ぶ．これは動的に構成される対象モデルの構造表現に適している．このような構造は人工的なものであり，上記集合関係も一つの解釈である．この解釈は集合を特定の対象に結びつける機構により実現されることはすでに述べた．この結合を記号 $\rhd$ によって表す．たとえば図1の例では実体である面 s_j が集合 b_j に結合されており，$s_j \rhd b_j$ である．したがって，$s_j \ni a_i$ の関係は，実は，$s_j \rhd b_j$, $b_j \ni a_i$ という2段の関係を通して表現される．多層論理を用いることは，世界内に上記の2種の基本構造を含むデータ構造を含むことを意味し，$\rhd$ や $\ni$ はこのほかのいくつかの記号とともに，構造化のための基本関係の表現記号として用いられている．このような，I型，II型の両基本構造を含むデータ構造については付録を参照されたい．この構造内で，すべての対象は一定長セルで表されるノードを有し，各ノードは上記基本関係を表す数種のポインタで他のノードに連絡している．この構造に多層論理式で表された各種の知識素が組み込まれ，全体として，データ構造と式から構成される大きな抽象データ構造として知識ベースが構成される．式は接頭部に $(Qx/\sharp d)$ なる表現を有するとき，ノード d にリンクされる．この接頭部は $x \in d$ なる範囲の対象を指定しているが，この対象の範囲はデータ構造から直接に求まる．この対象 d が特定のインスタンスであるときは，その要素はII型構造としての要素を意味するから，まず $\rhd$ を辿り，次いで $\ni$ の関係を辿ること

が必要であるが，d がインスタンスでなければⅠ型構造関係を辿ることになる．この区別は文脈から明らかであるが，記号 # と同様，理解を助けるために補助記号として二重スラントを用い，Ⅱ型構造に基づく，(18′) のような表現は

$$(\forall y//\#h)(\exists x//y)[(\mathrm{AREA}\ y\ 20) \wedge(\mathrm{LENGTH}\ x\ 5)] \qquad (18'')$$

のように表すことにする．

なおⅡ型構造の定義に際し，基本集合 d に関しては有限性以外は何も仮定していない．d の要素は均質である必要もない．Ⅱ型構造は与えられた要素から作られる上位構造関係のみが問題であって，各要素の詳細には立ち入らないからである．したがって，たとえば異なる部品から作られる機械，異なる原子から作られる分子などをこれによって表現することができる．また人工的構造というⅡ型構造の性質から，d の有限性の仮定は実用上の障害にはならない*.

3.4 合成問題

一般にモデルに関する問題には二つの型がある，第一は解析型で，対象の構造を与えて，その性質や行動を求める問題である．第二は合成型で，基本要素と，最終的に作られる対象のもつべき性質や行動を要求として与え，それを満たす構造を求める問題である．たとえば，図1に関して，立体の構造を与え，その容積を与えるのは解析型，逆に，与えられた線分を用い，与えられた容積をもつ立体の構造を求めるのが合成型である．設計等は合成型の典型であるが，これは構造化と，その評価のプロセスから成り，後者は解析型問題であるから，一般に合成型の問題のほうが困難である．さて前記の例の立体構造は基本要素（線分）集合を d とすると $*^2d$ の要素である．すると要求は

$$(\exists z//*^2d)(\mathrm{VOLUME}\ z\ 100)?$$

のような形式になる．さらに一般的な形式では，中間に生成される部分構造に関しても制約を付けることにより，合成のプロセスを効率的にしている．この形は

$$(\exists z^n//*^n\#d)(\forall z^{n-1}//z^n)\cdots (\forall z^1//z^2)[M(z^n, z^{n-1}, \cdots, z^1)]? \qquad (19)$$

のようなものである．要求記述のみでなく，知識素として構造を含む表現を作る場合も同様である．

この一例として前述の立体表現を考えよう．ただし今度は基本構成要素集合として空間内の点の集合を与え，これらの点を頂点にもつ多面体を定義する知識素を作る．これを $d=\{p_1, p_2, \cdots, p_N\}$，$p_i$：空間内の点，として，次のように表しておく．

$$(\forall z//*^3\#d)(\forall y//z)(\forall x//y)[(\mathrm{COUNT}\ x\ 2) \wedge(\mathrm{ELEM\text{-}CYCLE}\ y)\wedge(\mathrm{EULAR\text{-}EQ}\ z) \rightarrow(\mathrm{POLYHEDRON}\ z)] \qquad (20)$$

基本の要素集合 d が空間内の線分の集合であるとき，立体が $*^2d$ の一要素であることはすでに示した．d が点の集合のとき，線分は $*d$ の一要素となり，立体は $*^3d$ の一要素である．(20)は $*^3d$ の要素のなかで(20)を満たすもののみが多面体の条件を満たすことを示す．ここで各述語を次のように定義している．

(COUNT x y)：(集合) x の濃度は y である．

(ELEM-CYCLE y)：(集合) y は単純閉路をなす．

(EULAR-EQ z)：(構造) z はオイラー式を満たす．

(POLYHEDRON z)：(構造) z は多面体である．

(COUNT x 2) は空間点のペア（濃度2の集合）が線分を表すこと，(ELEM-CYCLE y) は線分の連鎖が閉じている（面の輪郭を作る）ことをそれぞれ示している．(EULAR-EQ z) は多面体に関する条件を与えている Eular 方程式である．

このように，多層論理は実世界（あるいはその像としての計算機内表現）に2種類のデータ構造を許す．この構造は論理式の接頭部により参照されるが，ここでは構造の階層間の関係のみを指定し，特定の構造（インスタンス）を含まない．構造自体はこれを参照する記述とはまったく別の指示によって作られ，あるいは修正される．構造が修正されてもこれを参照する論理式は影響を受けず，しかも常にその構造に関する正しい記述を与えているという望ましい性質をもつ．

なお多層論理表現(19)の定義はこのままではいまだ不完全である．多層論理表現の典型的な形は

$$(\exists x^n//*^n\#d)(\forall x^{n-1}//x^n)\cdots (Q_0x^1//x^2)M(x^n\cdots x^1) \qquad (21)$$

というものである．簡単のため

$$(\exists x^2//*\#d)(\forall x^1//x^2)F(x^1)? \qquad (22)$$

を考えよう．$x^2\in *d$ は $x^2\subset d$ と同じであるからこれは d の部分集合 x^2 で，そのすべての要素 $x^1\in x^2$ が $F(x^1)$ であるものにより満足される．このような x^2 が見つかったとしよう．このとき，$x^2\supset x^{2\prime}$，$x^2\neq x^{2\prime}$ なるすべての $x^{2\prime}$ も同時に(22)を満たしている．(21)の x^n についても同様である．これを避け，解の

* Ⅱ型構造はモデル構造の表現に不可欠ではあるが十分なものではない．これと，同一レベルの対象間の関係をグラフで表すことにより，広い範囲の構造を表現することができる．このような構造例については文献 6)，グラフを表形式で表したときの論理系による処理については 5) を参照されたい．

一意性を保つために，真の解 $\hat{x}^n$ は，(21)を満たすすべての x^n にたいし，$\hat{x}^n \supset x^n$ であるものと定める．この定義のもとで(22)式をより単純な表現の論理式 $(\exists x^1/\#d)F(x^1)$ と比較すると興味深い．後者は $F(x^1)$ を満たす $x^1 \in d$ が存在すれば真であり，その個数は指定されない．これはデータベースへの応用などには不便である．一方，前者は定義により，条件に合うものすべてを見いだすことを要し，曖昧さはない．多層論理ではこの両方を使い分けることができる．データベースへの応用に関してはさらに(22)と類似な

$$(\exists x^2//*\#d)(\forall x^1//x^2)[G(x^2)\wedge F(x^1)]? \qquad (23)$$

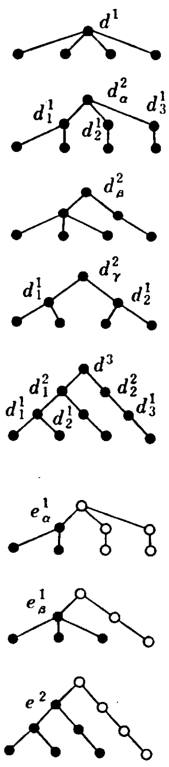
を考えよう．これは $F(x^1)$ なるすべての $x^1\in d$ の集合として $x^2\subset d$ を定め，その x^2 について $G(x^2)$ を求めることを示す．$G(x^2)$ が統計処理ルーチンと関連する述語とすると，(23)はデータ群中，$F(x^1)$ を満たすものの統計値を求めるといった目的に使うことができる．

以上は多層論理の非公式の説明である．新しい論理の形式を定める際には，これに基づく推論方式を求めねばならないが，本稿では紙数の都合でこの部分の記述は別の機会に譲る．推論方式の一部は文献6)および7)にあるので参照されたい．以下ではこの論理系が前に挙げた望ましい知識表現であるためのもう一つの条件を満たすことを示しておく．

4. 多層論理式の情報量

多層論理式は参照する構造に応じて異なった情報量を示す．これを簡単な例で示してみよう．

$d=\{a_1, a_2, a_3, a_4\}$ を4要素から成る基本要素集合とする．これから作られる各種のII型構造を図2(a)に示す．II型構造では構造内の各ノードが実世界と対応した意味をもち，それが基本要素から上方のどのレベルにあるかにより異なるし，また，同一の概念（たとえば'面'）の対象同士は同一レベルに置かれねばならない．これは意味をもたない記号の構造としてのリスト構造と異なる．ただしII型構造を用いてリスト処理を行うアルゴリズムを作ることは可能である．図2において，d^i のように末端ノード以外のノードには右肩にレベルを付して表している．同図でA群は基本要素集合 d をすべて用いた構造群であるが，B群はA群の構造の部分構造について記述された場合の例をいくつか参考のため示したものである．図中黒丸部分

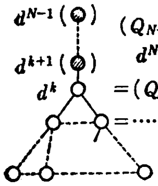


		(b)	(c)
〔A〕	(i)	$(\exists x// d^1)F(x)$	$4-\log 15=0.098$
		$(\forall x// d^1)F(x)$	4
	(ii)	$(\exists y// d^2_\alpha)(\forall x// y)F(x)$	$4-\log 13=0.300$
		$(\forall y// d^2_\alpha)(\exists x// y)F(x)$	$4-\log 3=2.415$
	(iii)	$(\exists y// d^2_\beta)(\forall x// y)F(x)$	$4-\log 9=0.830$
		$(\forall y// d^2_\beta)(\exists x// y)F(x)$	$4-\log 7=1.193$
	(iv)	$(\exists y// d^2_\gamma)(\forall x// y)F(x)$	$4-\log 7=1.193$
		$(\forall y// d^2_\gamma)(\exists x// y)F(x)$	$4-\log 9=0.830$
	(v)	$(\exists z// d^3)(\forall y// z)(\exists x// y)F(x)$	$4-\log 11=0.541$
		$(\forall z// d^3)(\exists y// z)(\forall x// y)F(x)$	$4-\log 5=1.678$
〔B〕	(i′)	$(\exists x// e^1_\alpha)F(x)$	$4-\log 12=0.415$
		$(\forall x// e^1_\alpha)F(x)$	2
	(ii′)	$(\exists x// e^1_\beta)F(x)$	$4-\log 14=0.193$
		$(\forall x// e^1_\beta)F(x)$	3
	(v′)	$(\exists y// e^2)(\forall x// y)F(x)$	$4-\log 10=0.678$
		$(\forall y// e^2)(\exists x// y)F(x)$	$4-\log 6=1.415$

(a) (b) (c)

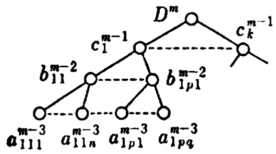
図 2 データ構造による情報量

Fig. 2 Information depending on data structures



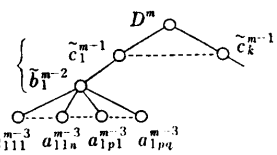
$$(Q_{N-1}x^{N-1}/\!/d^{N-1})(Q_{N-2}x^{N-2}/\!/x^{N-1})\cdots(Q_kx^k/\!/x^{k+1})\cdots(Q_1x^1/\!/x^2)F(x^1)$$
$$d^{N-1}=\{d^{N-2}\},\quad Q_N=\forall\ \mathrm{or}\ \exists$$
$$=(Q_{N-2}x^{N-2}/\!/d^{N-2})\cdots(Q_kx^k/\!/x^{k+1})\cdots(Q_1x^1/\!/x^2)F(x^1)$$
$$=\cdots\cdots=(Q_kx^k/\!/d^k)\cdots(Q_1x^1/\!/x^2)F(x^1)$$

図 3 等価表現（1）
Fig. 3 Equivalent expressions (1).



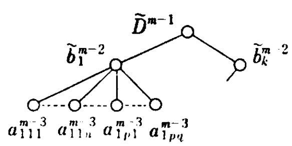
$$(Qz/\!/D^m)(\exists y/\!/z)(\exists x/\!/y)F(x)$$
$$D^m=\{c_1^{m-1},\cdots\cdots,c_k^{m-1}\}$$
$$c_i^{m-1}=\{b_{i1}^{m-2},\cdots\cdots,b_{ip_i}^{m-2}\}$$

(a)



$$(Qz/\!/D^m)(\exists y'/\!/z)(\exists x/\!/y')F(x)$$
$$D^m=\{\tilde{c}_1^{m-1},\cdots\cdots,c_k^{m-1}\},$$
$$c_i^{m-1}=\{\tilde{b}_i^{m-2}\mid i=1,2,\cdots,k\}$$

(b)



$$(Qy'/\!/\tilde{D}^{m-1})(\exists x/\!/y')F(x)$$
$$\tilde{D}^{m-1}=\{\tilde{b}_1^{m-2},\cdots\cdots,\tilde{b}_k^{m-2}\}$$

(c)

図 4 等価表現（2）
Fig. 4 Equivalent expressions (2).

が記述対象とされた部分であり，白丸は対応するA群の構造のうち，記述対象から省かれた部分を示す．

図2（b）はこれから構造をもった対象に関する記述例であり，同図（c）はこの記述のもつ情報量を示す．この求め方は2章で示した方法と同様であるが，一例として A-(ⅲ) の第一の場合を示してみよう．この記述に関しては

$$S_F^2=\{F_{(a_1a_2a_3)[a_4]},\ F_{[a_1a_2a_3](a_4)}\}$$

が可能な状態集合として求まり，$|S_F^2|=9$ である（$F_{(a_1a_2a_3)[a_4]}$ の数は 2，$F_{[a_1a_2a_3](a_4)}$ の数は8であるが $F_{a_1a_2a_3a_4}$ が両方に含まれるため，1を差し引く）．したがって $K=\log|S_F^1|-\log|S_F^2|=\log 2^4-\log 9=0.830$ だけの情報量をもつ．

この簡単な例が示すとおり，多層論理式は構造に応じてさまざまな情報量をもつ．$|d|=N$ のとき最大情報量は N であるが，N の増大に応じて可能な表現間の情報量の（最大情報量に対する）相対間隔は減少する．なお図2（b）において，構造に応じて論理式が異なるように見えるが，実際はA，Bの各群内では異なる論理式はそれぞれ2種類しかないことが簡単に示せる．

図2（b）には高さの異なる構造が含まれているが，基本要素が N 個の場合，最も高い構造は $N-1$ レベル (e. g. A-(v)) である．そこですべての構造に，**図3**のように上位にノードを加えて $N-1$ レベルの構造を作るとしよう．これにたいし，

$$(Q_{N-1}z^{N-1}/\!/\#d^{N-1})(Q_{N-2}z^{N-2}/\!/z^{N-1})\cdots(Q_1z^1/\!/z^2)F(z^1) \qquad (24)$$

なる式を作る．ここで $Q_i, i=1,\cdots,N-1$, は∀または∃である．この式は，もし $d^{N-1}=\{d^{N-2}\}$，すなわち単一要素集合であるなら，Q_{N-1} にかかわらず $z^{N-1}=d^{N-2}$ に限定されるから，(20)式は

$$(Q_{N-2}z^{N-2}/\!/\#d^{N-2})\cdots(Q_0z^1/\!/z^2)F(z^1) \qquad (25)$$

と，一段低い構造に関する記述に等しくなる．これを繰り返すと結局もとの高さの構造に戻る．

次に，**図 4**（a）のように，構造に沿った述語の接頭部表現で，同一の限量記号が二つ以上続いたとしよう．このとき，構造を同図（b）のように変形しても論理的には不変であり，さらに，これが同図（c）のように，段数の一段低い構造における表現と同一になる．したがって接頭部内の限量記号が ∀ と ∃ で交互に変化するもののみを考えればよく，結局，これら2種の等価表現間の変換を前提とすると，要素集合 d が与えられたときの論理式は2種類のみであることは通常の論理の場合と変わらず，論理式の種類が増えることは好ましくないという，本稿の最初に述べた要求にこの点でも合致しているのである．

次に合成型の表現の場合を考える．基本集合 d の濃度が N のとき，$|*d|$ は 2^N-1 である．一般に $|*^nd|=2^{|*^{n-1}d|}-1$ である．N がある程度大であるとき，これを $|*^nd|\fallingdotseq 2^{|*^{n-1}d|}$ として扱っても大きな影響はない．すると，$*^nd$ を変域とする表現が与えられるとき，事前のエントロピは $2^{|*^{n-1}d|}$ である．もし，この表現を与えることにより，これを満たす構造が M 個定まるとすると，この表現の情報量は

$$K\fallingdotseq 2^{|*^{n-1}d|}-M \qquad (26)$$

である．一般に合成問題がむずかしいのは，M は小

さく $|*^{n-1}d|$ がきわめて大きいため，その間に与えるべき情報量が大になるためである．

合成問題を実現可能なものにするには $*^n d$ の可能性をできる限り減らすことが必要である．多くの場合，(20)式のように，表現のなかに中間構造に関する条件 (COUNT や ELEM-CYCLE など) が含まれ，これを利用して可能性を減らすことができる．これに関しても文献 6) を参照されたい．

ちなみに(20)の (COUNT x 2) のような条件がどの程度の情報量をもつかを調べてみよう．$|d|=N$ とし，

$$(\exists x//*d)(\text{COUNT } x\ 2) \qquad (27)$$

を用いたとする．このときの事前のエントロピは $\fallingdotseq 2^N$ であり，事後には，2^N 個の要素のうち ${}_N C_2$ 個のみが選ばれるから，$K=2^N-{}_N C_2$ であり，N が大になるとこの情報量も大きい．しかし，中間構造に関する制約は最終構造の段階でいっそう大きく現れる．

(20)のように，$*^3 d$ のなかから最終構造を選ぶ際，$|*^3 d|=2^{2^{2^N}}$ であり，もし中間構造にたいする制約がない場合，このなかから条件に合うものを探さねばならない．しかし，(27)をこのような条件として用いた場合，最終構造の探索範囲は $2^{2^{{}_N C_2}}$ に減少する．したがって，(27)は実質 $2^{2^{2^N}}-2^{2^{{}_N C_2}}$ の情報量を有することになり，$2^N-{}_N C_2$ が著しく増幅される．なお，最終構造の選定に先立って，中間構造に関する制約を適用して選択を減らしても，論理的には差支えない（文献 6) 参照)．

5. 情報量評価の利用

これまで知識素の情報量を，知識表現のもつべき性質を明らかにするという意味で用いてきた．しかし，情報量の概念は知識ベース・システムにおいてもっと積極的な利用法がある．この問題を十分に論ずるにはすでに紙数が不足であるので，ここでは重要なものを二つだけ挙げておこう．

5.1 含意判定

二つの論理式 A, B があり，A のすべてのモデルが B のモデルでもあるとき (A が真なら B も真であるとき)，$A \Rightarrow B$ と書き，"B は A の論理的結論である" といおう．$A \Rightarrow B$ の判定は論理処理の最も基本的な処理であり，どのようなシステムにおいても（表現は異なるが）不可欠な機能である．すなわち，B が質問であった場合，知識ベースのなかに $A \Rightarrow B$ を満たす知識素を探索する必要があるが，この判定には一定の手続きを要する．知識ベースが実用問題に役立つ程度に大きくなったとき，このなかから条件を満たす知識をいかに速やかに見いだすかがシステム実現のきめ手になりかねない．これに関し，次の定理が成り立つ．

定理　論理式 A, B の情報量を I_A, I_B とする．$A \Rightarrow B$ なら $I_A \geq I_B$ である

これを用い，各知識素の情報量を前もって求められれば $A \Rightarrow B$ の判定は $I_A \geq I_B$ のものについてのみ行えばよい．この証明はほぼ自明といってよいが A, B が異なる構造を参照する一般の場合を含め，この証明は別の機会に譲る．

5.2 知識の獲得

知識ベース・システムにおいて知識の獲得は重要な問題であるが，とくに開発型の問題は情報を外部から与えつつ，逐次的にモデルを構築するという意味で，問題の本質が知識獲得のプロセスにある．試行錯誤によってモデルを構築するとき，与える情報量が状況に応じユーザの意思で制御しうることが最も望ましい．多層論理はこの理由のみでも有用であるが，さらに対象の集合の分類や構造を変えることにより情報量を増大してゆくというプロセスは人間によってごく自然になされる知識の洗練方法である．この点についてはさらに説明を要するが，詳細は別の機会に譲りこの事実を指摘するにとどめる．

6. むすび

知識表現の問題を情報理論的な観点から考察し，とくに開発型の問題に知識ベース・システムを利用する際の知識表現の具備すべき性質を挙げ，それを満たすものとして多層論理を導いた．これは曖昧さのない結果を導けるシステムを目指したものであるが，この基本方式の上に曖昧性を定義して，一般の問題に適用することは可能である．多層論理の有用性は第一に記述の対象が構造をもつとき，その構造表現と記述のための論理式を分離でき，論理式を変えずに構造を操作できることにある．これは構造の参照が論理式の接頭部においてのみなされうる構造になっているからである．しかし，これのみで十分なわけではなく，構造を操作するためには，構造表現を論理式の本体部分に含む，従来どおりの形式も必要となる．

多層論理の第二の利点は多様な情報量の表現が可能なことである．このことは現実の問題では重要である．

本論文では最後に情報量のさらに積極的な使い方を

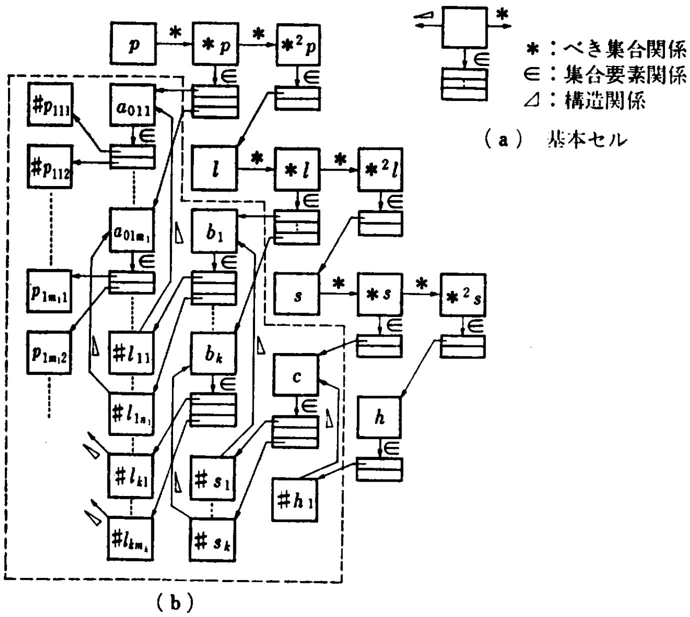
示唆している．しかしこの部分は紙数の都合で要旨のみ示した．この部分にかぎらず，本論文は知識表現にたいする考察の一部を示したにすぎない．多層論理の推論方式，推論機能の情報論的な意味づけ，同じく知識の曖昧性や冗長度の定量的評価等，知識表現問題をたんに経験的な技術にとどめることなく，科学的な立場から論ずべきことは多い．これらについても機会を改めて述べたい．

参考文献

1) Barr, A. and Feigenbaum, E. A.: *The Handbook of Artificial Intelligence*, Vol. 1, Chap. 3, pp. 141-222, William Kaufmann Inc., Los. Altos (1981).
2) Brachman, R. J. and Smith, B. C.: Special Issue on Knowledge Representation, *SIGART Newsl.*, No. 70, pp. 1-138 (1980).
3) Chang, C. L. and Lee, R. C. T.: *Symbolic Logic and Mechanical Theorem Proving*, Academic Press, New York (1973).
4) Enderton, H. B.: *A Mathematical Introduction to Logic*, Academic Press, New York, London (1979).
5) Ohsuga, S.: Perspectives on New Computer Systems of the Next Generation—A Proposal for Knowledge-Based Systems, *J. Inform. Process.*, Vol. 3, No. 3, pp. 171-185 (1980).
6) Ohsuga, S.: A New Method of Model Description—Use of Knowledge Base and Inference—, in Bo, K. and Lillehagen, F. M. (eds.), *CAD System Framework*, North-Holland, Amsterdam, pp. 285-312 (1983).
7) Ohsuga, S.: Predicate Logic Involving Data Structure As A Knowledge Representation Language, Proc. '83 Eighth International Joint Conf. on Artificial Intelligence, pp. 391-394 (1983).
8) Ohsuga, S.: Knowledge Based Man-Machine Systems, *Automatica*, Vol. 19, No. 6, pp. 685-691 (1983).
9) Robinson, J. A.: A Machine Oriented Logic Based on the Resolution Principle, *J. ACM*, Vol. 12, No. 1, pp. 23-41 (1965).
10) Watanabe, S.: *Knowing and Guessing—A Formal and Quantitative Study*, John Wiley and Sons Inc., New York (1969).

付録 対象の世界の構造表現

多層論理は世界内の対象間の集合関係に基づく構造を前提とした論理系である．このためまずデータ構造を定めねばならない．この構造を形成するのは3種の集合関係すなわち，要素-集合(∈)関係，部分集合(⊂)関係およびべき集合(＊)関係である．しかし，$x \subset y$ は $x \in (*y)$ と表されるから，表現のための基本関係としては⊂関係は除き，代りにII型構造を表すための構造(▷)関係を含めることにする．**付図1**はデータ構造の一例として，空間内の‘点’の集合 p から三次元形状の諸概念を構成する場合を示す．図中 p, l, s, h はそれぞれ一般的概念としての‘点’，‘線’，‘面’，‘多面体’を表す．記号♯の付いたものはそのインスタンスを示す．なお∈関係は1対 n 関係として同図(a)に示したように途中にポインタ表を経由する．図中，破線で囲った部分は図1に対応する多面体 $\#h_1$ を表すII型構造を，その外側はI型構造を表す．$\#h_1$ は p から作られるすべての多面体の集合 h の中の一要素である ($h \ni \#h_1$ で表されている) と同時に，$\#s_1, \cdots, \#s_k$ から成る集合 c としての性質を与えられている．この関係を示すのが $\#h_1 \rhd c$ の構造関係である．この図より，上位ノードからみて，I型構造は (＊, ∋) のペアを，II型構造は (▷, ∋) のペアを基本関係として作られる構造で表される．なお付図1は構造表現の説明のために単純化しており，現実の世界の構造化には＊, ∋, ▷のほかにもいくつかの基本関係が必要であるが，ここでは省略する．



付図 1 対象の世界の構造表現の一例
Fig. A. 1 An example of object structure in the world.

(昭和58年11月9日受付)
(昭和59年1月17日採録)