

知識工学とモデル構築

大須賀 節 雄

1. ま え が き

知識工学は、われわれが知識とよんでいる種類の情報を形式化して表現し、計算機に貯えて、問題解決に利用することを目的とした技術分野である。知識情報の蓄積を知識ベースとよび、これを利用する機構をそなえたシステムを知識ベース・システムとよぶ。

知識ベース・システムは潜在的に大きな将来性を秘めているが、これを真に有効な手段とするためには知識ベース・システムという情報処理機構がもつ固有の機能は何か、それをどの程度実現し得るか、そしてそれがどのような利用形態において有用性を発揮するかを十分検討しておく必要がある[5]。

本稿の目的は知識ベース・システムをモデル構築の手段として用いることがきわめて効果的であることを示すことにある。モデル化技術は問題解決の汎用的手法であり[1][10]、知識ベース・システムがそれを実現する汎用的手段であることが示せれば、この2つの概念は、実は問題解決という大きな目的の2つの面を表わしているといえる。モデル化技術に関しては上記議論に続いていかにモデル化するかを述べねばならないが、紙幅の都合でこの部分の議論は別の機会にゆずる。これに関しては文献[3]~[8]を参照され

たい。

2. 知識工学とは何か

知識工学は人間と計算機の間、従来の情報処理方式と異なる新しい関係を可能にする。人間が行なう情報処理の形態にはさまざまなものがあり、そのある部分は従来の、いわゆる von Neumann-Turing 型の計算機方式によって効果的に代替されてきた。しかし、この方式は人の情報処理機能のすべてをカバーするほどの適用性はないこと、すなわち、従来の計算機方式が適している情報処理の型と、そうでないものがあることが、経験を通じてある程度直観的に理解されるようになってきた。これらは、通常、定型的・非定型的という言葉で簡単に区別される場合が多い。しかし情報処理の方式を正しく分類することは、実はむずかしい仕事である。これは情報処理という総合的な活動に含まれている基本的な機能を、変換、記憶、学習、演繹、帰納などのように純粋な形でとり出し、個々の情報処理活動がそれらのどのような組合せで構成されているかを明らかにすることから始めねばならない。これによってはじめて、どのような機能を現在の技術で実現できるか、それがどのような利用分野をカバーし得るかを明らかにすることができる。

しかし、情報処理の基本機能の機械化はおろか、それを明確に定義すること自体に多くの困難が含まれていることは、これが単に情報工学のみ

おすが せつお 東京大学 工学部

でなく、古くから心理学、数学、論理学、言語学等の諸学問の主テーマであり続けていることから明らかである。とはいえ知識工学が、その名のとおり、工学としての地位を確保するためには、このような体系化のための努力が今後一層必要とされよう。

現在の知識工学はこれら基本機能のうちの推論の一部を機械的に実現したものである。この点で従来の計算機とは異なった機能をもっていることは確かであるが、特定の応用を中心に発達してきた知識工学という言葉が表わす技術の、これまでの定義は必ずしも一般的ではない。これまで特定問題ごとに、経験的に構成されてきた知識工学のシステムの定義は、個々の応用分野に依存したものにありがちだからである。知識ベース・システムが潜在的にもつ処理能力を考えるなら、その定義はできるだけ応用分野に依存しないものとしておくべきである。これは必然的に、機能中心に、あるいはその機能を実現する構造中心の定義になる。これを定義に忠実に実現すれば汎用性の高い知識ベース・システムが得られよう。

このためには、まず知識そのものの定義から始めねばならない。これまで多くの知識工学システムが問題依存になったのも、実は「知識」の定義が包括的でなかったことにも起因しているからである。

たとえば、“知識とはその知識工学システムを適用する分野の専門的知識 (expertise) の集積である”という定義が、エキスパート・システムとよばれるクラスのものについてなされている。しかし専門的知識を使う際に誰でも知っている常識が不可欠なこともある。また、最近では自然言語解析や図形解析などに知識ベース方式を用いる各種の試みがなされている。この場合、言語の使い方などを知識の形で記憶することになるが、前述の定義ではこれらはいずれも知識でなくなってしまう。したがって、「知識」を応用分野における情報の意味とは切り離して、すなわち、それが表わす

内容とは無関係に形式によってのみ定義することが必要である。

形式的に定義する際の常用の方法は、まず表現の構造を与えることである。これを知識表現とよぶ、これは表現の基本になる文字ならびに記号とそれらを律する有限個の構造規則 (文法) から成る。

知識の形式的定義をさらに先へ進めるためには、まず推論機能を、これも形式のみによって定義しておかねばならない。推論とは「有限個の表現からこれらと異なった表現を形式的に導く機能」とするのが最も包括的である。たとえば“人は死ぬ”と“山田太郎は人である”から、“山田太郎は死ぬ”を導くことは最も単純な推論である。ここで前提から結論を導く過程は、対象が‘山田太郎’でなく‘鈴木二郎’であってもまったく同様である。これは内容でなく、山田太郎 \in 人という形式的関係にのみもとづいて推論が行なわれ得る例を示している。

この形式的推論過程は知識表現という形式に対して定義される。こうすると、(1)情報表現の構造規則と、(2)この形式にしたがって記述された表現に対して形式的に定義されている推論規則が与えられた時、「原則として推論の対象になることを前提として、(1)の形式によって表現された情報」が知識であり、「知識の集積」が知識ベースである、とするのが、知識および知識ベースの最も包括的な定義であろう。またこれを用いて知識ベース・システムを「知識ベースと、それを問題解決に利用するための推論機構を基本要素として含むシステム」とすると、これも最も包括的な定義を与える。

ここで推論という機能についてもう少し踏み込んで考えてみよう。なぜならこれが知識ベース・システムの能力限界を定めているからである。上述の一般的定義のもとで、前提となる情報と、結論として導かれる情報の性質によって、推論の性質が非常に異なってくる。

知識表現で記述される知識は内容的にはある対象に関する属性、性質、関係者を表わしている。ここで対象としては、‘山田太郎’とか‘富士山’のように個別的なものと、‘人間’、‘山’のように一般的な概念で表わされるものがある。これに応じて、知識にも個別の対象に関する記述と、もっと一般的な対象に関する記述とがある。前者を個別(的)知識、後者を一般(的)知識とよぼう。たとえば“山田は死ぬ”は個別知識，“人は死ぬ”は一般知識である。一般知識にも、より一般的である(ない)知識がある。たとえば，“りんごは落ちる”より“物体は落ちる”のほうがより一般的である。‘りんご’も‘物体’も一般的概念であるが、後者のほうがより包括的だからである。

推論のうち、ある一般的な知識から、それより一般的でない知識を導くものを演繹的推論とよぶ。仮にある特定の症状をもつ一群の患者についての医療診断・治療規則が与えられたとして、それを特定の人‘山田さん’に適用して山田さんの診断・治療法を導こうとするなら、その行為は演繹的推論である。演繹的推論は機械的手順が得られており、推論機能の中でも最も低レベルのものである。現在の知識ベース・システムで実現されているのは、ほとんどがこのレベルである。なお用語の通常の意味では推論ではないが、多数の個別的データの中から指定された特定のものを選び出す検索機能を推論機能の最も下位に位置づけておく。これにより、データベース・システムを機能的に知識ベース・システムのサブシステムと見なす。これはともに計算機内の半永久的記憶情報を用いる点で共通の性質があるからである。

一方、一般的に、複数の、個別的あるいはより一般的でない知識から、それらより一般的な知識を導くことを帰納的推論という。実験等の観測値にもとづいて一般法則を作り出してゆくのはこの機能である。一般的な帰納的推論の手順を見出すことは困難であり、現在の知識ベース・システムではほとんど実現されていない。通常、人の行な

う帰納の手順は個別的知識から類推される仮説(求めようとする一般的な知識の候補)を作り、これが正しいか否かを演繹的に検証するというものである。ここで帰納的推論の困難な部分は仮説の生成に集約されるが、このメカニズムは解明されていない。

帰納的推論は実現が困難とはいえ、その機能は定義が与えられる程度には明確になっている。したがって汎用性を捨て、知識の内容に依存した推論を行なうなら、多少帰納的推論に近いものを部分的に行なうことは可能である。しかし人間が通常行なっている知的活動の中には、これ以外に多様な思考機能を含んでいる。その中には、存在することはわかっている、厳密に定義できるほどに分析がなされていないものもある。人間が行なう判断はこのようなすべての方法を駆使した結果であることは十分考えておかねばならない。

3. 知識ベース・システムの有用性

演繹的推論と知識ベースを基本とする知識ベース・システムにはどの程度の有用性があるだろうか？あるいは、大きな有用性を示す利用法はどのようなものだろうか？この間に答えるには客観的なデータが乏しいため大半は考察に頼る他ないが、これまで開発されてきた応用向きの知識工学システムが予期されたほどには利用されていないのに対し、記憶情報を利用するという点で知識ベースの一特殊形態として位置づけられたデータベースのほうは早くから有用性を認められて比較的短期間に実用化された事実は検討に値する。データベースの特徴はデータ量が大きく、データの形式が単純なことだが、大量の画一的なデータの記憶は人間にとって不得手な機能であるから、そこに早くからニーズがあったのである。

しかしこの事実は推論機能が不要なことを意味するものではない。データベースの分野でも、まず上記のニーズが最小限満たされた後、次の段階としてユーザー・ビューの導入や自然言語による

アクセスなど、推論機能を含まざるを得ない要求が強まっている。

一方従来の知識工学システムには強いニーズがあったわけではなく、いわば発明が先行して必要を喚起してきたといえる。現状はようやくその効果が現われ、関心もたれ始めた段階といえよう。

ニーズという点で注目されるのは、最近、設計・研究・開発などの分野で、対象が複雑化し、技術革新の速度が早まっていることなどの理由で機能の高い計算機システムへのニーズが生じていることであろう。

一方、有用性を考察するうえで知識ベース・システムの特徴を分析する必要がある。これについては[5]を参照されたい。ここでは、利用上最も大きな特徴は知識という形で与えられる情報のモジュラリティが高いことにあることを指摘するとどめる。

これらの状況を考慮し、知識ベース・システムの代表的な利用の型として次の3種をあげておく。

1. マン・マシン・インターフェース
2. コンサルテーション型
3. モデル構築による問題解決支援型

(1) マン・マシン・インターフェース

人と計算機のインターフェースをよくすることは1つには人間にできるだけ、日常用いているものに近い表現法を許すことであろう。しかしこれは、通常考えられている以上にむずかしい。このむずかしさは、人間と計算機間で情報表現の形式すなわち言語が異なるのみでなく、人間の場合その言語の使い方がきわめて柔軟であるため、この変換が単純に形式の変換にとどまらない点にある。特に人間同士の日常の表現にしばしば表現の省略や隠喩の類が入り、これを正しく変換するには、会話のなされた状況を理解せねばならない。それには当事者が会話の内容に関連する知識をもつ必要がある。この意味でマン・マシン・システムは基本的に知識ベースを必要とする。

マン・マシン・インターフェースにおけるもう

1つの問題は、人間が自然言語、図形、画像など多様な情報表現を組み合わせて使っていることである。現在の知識ベース・システムの技術では文字情報以外に受け入れ得る表現形式は未だ限定的で、これを改善する努力が必要である。

インターフェース技術は今後知識工学の主要問題の1つになることは確実と見られているが、ユーザーにとって計算機利用の主目的な別において、マン・マシン・インターフェースはこのための支援技術である。したがってインターフェース改善のための投資は、それによって計算機利用の主目的がその投資を上まわるほどに改善されることが前提である。この点、この主目的にかかわる以下の2種の応用とは多少性格を異にする。

(2) コンサルテーション型

いわゆるエキスパート・システムとかコンサルテーション・システムとよばれるもので代表される知識ベース・システムの利用形態であり、この目的は専門家のもっている知識を貯え、推論機能を通して、非専門家に利用できるようにすることである。この形態でさまざまな応用は考えられるが、ユーザーに提示される情報はシステム内の知識の範囲を超えないから、重要なのは知識そのものであり、この点で情報検索やデータベースの延長上にある。この利用形態で注意すべきことは、しばしば計算機が人の判断機能を代行する結果になることである。判断機能は最も高度な知的活動の1つであるから、人がこれを計算機にゆずるとしたら計算機による判断が何らかの点で人間より優れていることを認めた場合であろう。事実、エキスパート・システムでは、専門家の知識を与えることにより、計算機が非専門家より優れた判断をすることが期待された。

この型の利用は現実には多くの制約がある。前述のように情報処理活動において人の思考は実に多様であり、上記条件が整う場合が限定されるためである。

人間は①直接に対象とする問題領域以外の広い

分野にわたって多くの知識をもち、これらが重要な働きをする。医療診断に例をとると、対象分野がたとえば呼吸器であっても、これ以外の内科諸分野から、さらに外科、耳鼻咽喉科など他科の知識のみならず、時には心理学、哲学などの広範な知識が診療にかかわってくるかもしれない。②X線や顕微鏡写真など多様な形式の情報を総合的に用い、推論に結びつけている。③演繹以外の高度の推論を行なっている。など1つの判断を下すにもあらゆる機能を使っている。②、③項に関しては現在の計算機技術では人に太刀打ちできないから、知識ベース・システムが相対的に人より優れた判断を行なうとしたら、それは②、③の機能があまり必要でない問題であり、かつ十分な知識もっている場合である。すなわち、データベースの場合と同様、人間の記憶能力の欠陥を補うものができてはじめて知識ベース・システムがこの面で有用性を発揮することになる。

このような知識は時とともに変化してゆくことが予想されるので、この管理を完全に行なうことが重要であるが、大量の知識情報が動的に変化してゆく時の管理は、データベースのように単純な形式のものでも多くの問題を含んでいる。まして知識ベースのように形式の複雑なものは一層困難を増すので、対策を十分に考慮しておくこと、特に知識の管理が計算機システム自身によって行なわれ得ることが望ましい。

前に知識ベース・システムの特徴として、知識のモジュラリティをあげたが、管理の問題は知識全体の性質にかかわる。個と全体の条件を両立させるためには、システムとしての統一的な理論体系の確立が必要であろう。なお計算機がどれほどの知識をもてばよいかは、それを扱う人と、扱われる問題により、一様ではない。関連する知識の広がりや問題によって異なり、最終的には人との相対関係で決まるからである。判断力の低い人との比較では容易に計算機が勝る。実用面では、このような関係の時、人は計算機の判断に全面的に

頼りがちであることは最初から予想してかかる必要がある。

(3) モデル構築による問題解決支援

研究・開発とか問題解決という仕事は総合的的判断と同様に最も高度な知的活動の1つである。しかし創造的な仕事の最も高度の部分は、それ以外の、非常に多くのかつ多様な仕事の上に成り立っており、人がこの部分の仕事に忙殺されてしまう場合が多い。このような仕事の処理に計算機を支援させるという考えは新しいものではなく、マン・マシン・システムの原点とも言えるものである。それにもかかわらず、今なお問題にされるのは上記部分は多様な仕事を含み、生じ得るすべての可能性を前もって見通すことはほとんど不可能であるのに、従来の計算機の処理方式は、行なうべき処理過程を前もって手続きとして書き下しておかねばならないから、原理的に上述の要求と相容れないためである。知識ベース・システムのもつ柔軟性をここに活用して、低レベル推論で処理される範囲の多様な仕事を知識ベース・システムが引き受けることにより、人には高レベル推論を十分に発揮する機会をふやすという方式の可能性は高い。このさい知識ベース・システムの能力(推論レベル)に応じて、両者の分担を決められるようにすることが必要である。前記(2)の利用法に属するこれまでの知識工学システムは、人間の行なう多様な仕事を縦割りにして、分野としては狭い範囲を、知的レベルでは広い範囲をカバーしているのに対し、(3)の利用法は仕事を知的レベルにより横割りにし、分野として広い範囲、知的レベルとしては低い範囲をカバーしようとする点に大きな相違がある。

この方式においても、どのレベルから先を計算機システムにまかせるかは、人によって異なるし、それが自由に行なえることが重要である。そのためには人と計算機が常に同じ情報源、すなわち共同で処理する問題の表現としてのモデルをもつことが必要である。これがモデリングの意味で

あり、この技術は単に応用分野のひとつというのではなく、これからの情報処理技術の基本ともいべきものなのである。

4. モデリング技術

モデルは(広義の)問題解決システムにおいて、解決すべき問題、換言すれば行なうべき仕事の対象を情報的に表現したものである。現実の対象は複雑で、さまざまな面をもっているが、モデルはこのうち、処理しようとする面(見方)についてのすべての情報を含み、かつ可能なかぎりそれのみを含むように抽象化される。

このような一般的な定義は理解しにくいので、典型的な例としてCAD(設計支援)システムをあげよう。設計とは与えられた要求性能を満たす対象を情報的に作り上げてゆく仕事である。このためにはまず作られるべき対象の表現形式が定められねばならない。機械設計の場合、これには従来、三面図が用いられ、電子設計や分子構造設計ではそれぞれ、回路図や分子構造図(式)が用いられてきた。設計対象が備えるべき要求性能や性質(たとえば対象が航空機の場合、重量、空気力学的特性、構造強度など、人が注目する面について対象がもつ特徴)はすべて上記図面等が表現する対象の構造に依存する。この意味で、これらの表現は解決すべき問題に関して不可欠な情報を含んでおり、それぞれの分野におけるモデルの一部を構成している。これらは対象の構造を表わしているが、性能等の評価のさい、この他に構造情報から直接導くことのできない属性等を表わす情報をモデルに含めることも、しばしば必要となる。たとえば強度計算をするためには、対象の構造の他に材料等の指定が必要である、モデルにもとづいて所要の性能等を評価する手段は従来の手法でプログラム化することができ、現に、すでに多種類のもの(たとえば有限要素法による構造解析プログラム等)が開発されている。

もし、上記のモデルが計算機内に表現されてい

れば、これら評価手段をモデルに適用してただちに評価結果を得、人に表示することができる。モデルの構造をも随時ディスプレイすれば、人は常にモデルの状態を把握することができる。このディスプレイ機能も、性能評価と同様、これが“人にどのように見えるか”を表わすモデルの一面と考えられる。さらに、もし、モデルが、その構造を変更する手段をあわせもち、またどのように変更したら望ましい性能をもつものに近づくかに関して、各技術分野ごとにこれまで蓄積されてきたノウハウをもち、さらに人とのインタラクション手段をもてば、非常に効果的な開発支援システムが得られるであろう。

以上述べたことはモデル構築システムが満たすべき条件であるが、これらは知識ベース・システムに対する条件と類似な点が多い。この意味で知識ベース・システム技術とモデル構築技術は近い関係にあり、前者はすでに述べたようにシステムのもつ機能中心の定義、後者は前者が最も有効に用いられる問題解決の方式に関する定義といってもよい。なお上述のモデル構築システムに対する要求は、3.(1)マン・マシン・インターフェース技術を含む。また推論機能を中心とする点では(2)の検索型利用と共通部分も多いが、モデルの表現は対象の複雑な構造の記述、性質等の記述、これらの間の変換則の記述など、モデル化技術に固有の機能が含まれる。

4. むすび

人と計算機との関係は共同で1つの仕事をするという立場からいえばすべてマン・マシン・システムとして考えることができる。モデリングはこの広義の汎用マン・マシン・システムを実現するうえから不可欠な技術として早くから認識されていた。知識ベース方式をこれに用いることは、人と計算機との役割分担のレベルを1段上げることであり、一見、きわめて地味なものである。

しかし、情報処理技術の上から見れば、この1

段のレベル向上の意義は大きく、革新的とすらいってもよい。革新的というのは、従来の計算機と異なる機能を情報処理体系の基礎に置いており[5]、その結果、人間との接点での情報表現レベルが上がるからである[7]。しかし今回は紙数の都合でこの部分は省略する。

最後に1つだけ付加する。本稿で述べた汎用の知識ベース・システムの考え方は、1970年前後に米国において述語論理がもてはやされた頃の議論と多少類似性がある。その後述語論理は処理性や記述性などの点で欠陥があるとの批判を受け、後退した。しかしこれらの批判は表現手段が本質的にそなえるべき性質に無関係である。処理性に関しては当時と現在のハードウェア環境の比較にならない相違を考慮すべきであるし、その他の問題は表現のもつ論理体系そのもののもつ欠陥ではなく、計算機化の方式の問題だからである。これらについても機会を改めて述べたいと思う。

なお、モデリングに関してCADを例にしたが、DSS(意思決定支援システム)なども、技術の性質としては同種のものと考えられる。

参 考 文 献

- [1] Newell, M. & Evans, D. C. : Modelling by Computer. in *CAD System*(ed. Allan, J. J.). North-Holland, 1977, 291-308
- [2] Ohsuga, S. : Perspectives on New Computer Systems of the Next Generation—A Proposal for Knowledge-Based Systems. *Jour. of Information Processing*, Vol. 3, No. 3(1980), 171-185
- [3] Ohsuga, S. : A New Method of Model Description—Use of Knowledge Base and Inference. *Proc. of IFIP W. G. 5. 2 Working Conf. on 'CAD System Framework'* 1982. (To be published from North-Holland Pub. Co.)
- [4] Ohsuga, S. : Knowledge Based Man-Machine System. *Proc. of IFAC/IFIP/IFORS/IEA Conf. on Analysis, Design and Evaluation of Man-Machine System*, 1982. 345-350

特集に当って

丹 羽 清

人工知能の研究は、1956年に本格的に開始されたと言われるが、最近、応用人工知能ともよばれる知識工学への関心が高まっている。これは、現実の問題に適用できるという見通しが確かになりつつあるからである。

知識工学は、従来の人工知能が汎用の推論方式の研究を中心としていたのに対し、現実の問題解決のためには、知識そのものを重視せよと言う。この主張は、OR/MSの研究者・実務家の議論をよぶであろう。

ORと知識工学は、①現実の問題解決の方法に関する学問であり、②適用対象も重なり合うことが予想されるという点で、今後、ますます関係が深まることが考えられる。

本特集では、新世代コンピュータ技術開発機構の古川氏、淵氏に、知識工学のわかりやすい解説と、この技術の中核とする第5世代コンピュータについて、東京大学の犬須賀氏に、知識工学の定義と限界をふまえて、モデル構築による問題解決支援における知識ベースの有効性を、東京理科大学の溝口氏に人間の行なっている知識利用の方略を考察し、その計算機上への実現をねらいとするプランニングの研究について、日立製作所の佐々木氏、井原氏に、知識工学技術の実問題への適用の種々の試みを、おのおの執筆していただいた。

にわ きよし 日立製作所システム開発研究所

- [5] 犬須賀節雄：知識ベース技術の展望。情報処理 23-10(1982), 967-974
- [6] 犬須賀節雄：知識ベースおよびデータベースを用いたモデル構築技法。アドバンスト・データベース・シンポジウム予稿集, 1982, 89-104
- [7] 犬須賀節雄：知識工学とその周辺。電子通信学会研究会 AL82-63, PRL82-51(1982)
- [8] 犬須賀節雄：多層論理—モデル記述のための述語論理。ロジックプログラミング・コンファレンス, 1983
- [9] Zilles, S. Z. : Types, Algebras and Modelling. *Proc. of Workshop on Data Abstraction, Databases and Conceptual Modelling*, 1980. 207-211