

解 説

知 識 の 獲 得 と 学 習†



大 須 賀 節 雄††

1. は じ め に

知識処理技術の発展と実用化に伴って、機械による学習や知識獲得などの機能に関心が高まってきた。知識を対象とする以上、知識を獲得し、記憶し、利用するという一連の関連機能をすべて備えたシステムを目指すことは必然のように思える。しかし、これまで人工知能の分野で開発された多くの学習システムは、学習の結果に相当するプログラムを人が直接コンピュータに与えることができ、その方が性能上は良いシステムになるという性質のものであり、その意味では実験システムの域を出ていない。すなわち、これらのシステムは人の行っている情報獲得のメカニズムを分析して機械に再現させることを目的としたものであり、人間でも困難な発見的学習はほとんど実現していない。もしこれが機械による学習の越すことのできない限界であるとしたら、人間の学習能力研究や、科学としての学習研究への貢献は大きくても、実用面では、人が可能な限り最良のプログラムを与える方が、多くの場合、学習によって機械自身にそこに到達させるより良い結果となることを意味する。

もちろん、このような限界が理論的に示されたわけではない。現在の学習研究は実用性の点においてもより高度な機能に到達することを目指しているし、機械による学習は上述のように実用面のみでなく、人間自身の学習能力の解明や、基礎科学としての面を併せもっている。この意味での学習研究の意義は大きいが、同時に、学習のように、これまで実現されてきたものより一層高度の推論を必要とする能力の開発には、その意義を考えるとともに研究目的を明確にすることが必要である。

本稿の目的は知識工学という、実用性の高いシステムに関する知識獲得と学習について論ずることにある

† Knowledge Acquisition and Machine Learning 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

から、当然、実用面を重視する。したがって以下ではまず学習の一般的考察を行った後、主として実用面から知識獲得と学習について述べることにする。

2. 学習に関する一般的考察

2.1 学習の定義

学習は広義の概念であるため、深い意味の検討なしにこの用語が安易に用いられる場合が多い、そこでまずこの定義から始めよう。学習の一般的意味は「同一の試行もしくは同一の母集団からの試行を繰り返すとき、次の回の試行は前の回のものより良い性能を示すようなシステムの変化」のように与えることができる²⁹⁾。これは学習システムの所定の環境下での動作を外部から観察したときのものであって、総括的な定義を与えてはいるが、機械による学習を実現するための手掛りにはなり難く、これをより機構的なものに転化する必要がある。

2.2 学習研究の目的

一方、学習機能の研究にはそれぞれの目的を明確にする必要がある。これまでの学習研究はおおよそ3種に分類することができる。第1は人間の認知および学習機能の研究とそのコンピュータ・シミュレーション、第2は科学としての学習機能の研究、第3は実用化を目指した学習研究である。これらの研究は将来的には統合されるべきものであるが、現状はいまだそこまでは到っていない。

〔1〕 人間の認知および学習機能の研究

学習とは本来生物に固有の機能であるから、一般生物および人間について学習研究が、古くから主として心理学において、最近では記憶のメカニズムに関連して、生物学や脳生理学でも、行われてきた。一方、コンピュータの発達と共に、人の学習プロセスをシミュレートし、それによって人の学習機能を解明しようという研究が進められるようになった。これは学習研究の重要な目的の一つである。この研究目的に関しては、それが人の能力を超えるか否か、実用的な意義が

あるか否かといった点を顧慮する必要はない。一般に学習は人間にとっても容易なものではなく、したがって人間の学習速度は遅く、非効率である。この原因を解明することは新しい学習方法の開発にも役立つ。これは CAI 開発の研究とも深い関わりがある。この分野の最近の興味ある発展は、観察された結果から学生のモデルを推論する学習能力の研究に向っている^{30), 31)}。

これらの研究を通して、新しい機械学習の方法の研究にも発展することが期待される。

〔2〕 基礎科学としての学習研究

一方、学習機能を情報論的に考察し、その一般的なメカニズムを理論的に解明しようという研究がコンピュータ技術と歩調を合わせて進められてきた。この中心的課題は心理学的な考察よりも、もっと機械的な機能の分析や性能に重点を置いて、一般的な学習方法の分類や分析、あるいは情報生成の機能などを解明する点にある。学習法の分類に関しても推論機能との関係、特に帰納推論や類似推論などの論理学的考察のほか、環境との相互関係による分類、たとえば教師ありの学習と教師なしの学習およびその比較なども個々の学習研究の位置づけに役立つ。これまでの学習研究の多くがすでに人間に既知の事柄について、人が学習環境を整備し³⁷⁾、あるいは人が指導しながら習得させるものであったことは教師ありの学習を行っていたことを意味する。

〔3〕 学習研究の実用的意義

実用的な立場からは機械学習の意義は学習を通して(1)人間がプログラムすることのできない(人の知らない)結果に到達する、(2)人が直接プログラムするより容易にそれと同じ結果に到達し、かつそれが実用的な価値をもつ、のいずれかの場合に大きい。

(1)は新しい事実の発見が学習を通して行われる場合である。発見と学習は同義ではない。発見は、たとえば実験観測データの集まりから物理法則を見出すような機能、学習はそれを記憶し、検索用索引を自動生成することによって、次に同種の問題を扱う際に、ここで得た結果から出発できるようにする機能である。これら両機能は相互の組み合わせによって大きな効果を発揮するが、これが教師なしの発見的学習である。これを実用レベルで機械的に実現したシステムはいまだ存在しないが、近年、実例からの学習として、類似推論や帰納推論など、この方向の研究に目が向けられつつある。また発見的学習の結果は人間にとって

も未知の事実の発見として大きな意義をもつものであるから、今後これを人と協調的に行う発見支援システムの研究が重要になると予想される。なお、発見的学習は新しい一般的な物理法則の発見や、設計問題のように所与の要求を満たす構造をもった特定対象の発見など、必ずしも一様ではないが、いずれも情報の体系化という機能が含まれている。

(2)は人がプログラムすることも可能であるが、その準備を含めて煩わしい作業を伴うのに比して学習が効率である場合で、音声認識の際の特定話者の音声パターンを記憶して、以後の認識に利用するもの、大量のデータを含むデータベースに頻繁に新しいデータが付加されるようなとき、自動索引化を行うものなどがある。知識工学においてエキスパートの知識を機械に与えることも実用性という意味でここに含めておこう。このクラスの学習は実例も多いが、システムとして成功するためには、学習した結果を含むシステムに実用的意義が与えられることが不可欠である。

以上、学習研究の目的を3種に分類したが、第3の実用化研究は第2の一般的な学習機能の研究成果として得られる体系化の中に自己を位置づけることにより研究目標を定めることができるし、第2の学習の基礎的研究は第1の人間の認知プロセスの研究から多様な学習現象を見出すなど、これらの諸研究が相互に関連し合っていることはすでに述べた通りである。なお、以上の記述に含まれているように、学習には環境からの新しい情報の取得と共にそれを後に利用するための体系化が重要な機能として含まれている。

2.3 知識獲得と技能修得

学習に関する第3の視点として、学習という用語には知識の獲得・構造化と、技能の修得・熟練という2つの面があることも考えておく必要がある。前者はこれまでも触れたように、外部から記号化された情報を取り込み、それを体系化することである。この知識には物理システムの観測値や画像データのような単純な情報から、物理法則やエキスパート知識のように構造化され、テストされた情報までさまざまであり、それに応じて取り込む側の学習機能も異なる。この型の学習は環境や対象についてどの程度広くかつ正しく記述し得るかにより判定される。

技能修得はたとえばゴルフやピアノの練習のように、繰り返し訓練することによって技能が上達するが、記号で表現できる種類の知識と異なり、意識下の学習と呼ばれるものである。この種の学習は動作や認

知を司る生物器官の構造に直接関与するものとして、情報学的には低レベルの学習であるが、コンピュータ学習の対象としてはむしろ困難なものである。

このように学習と一口に言っても、含まれる概念は多様である。次章以下では知識工学という本特集の主旨に鑑み、また「知識の獲得と学習」という表題から、対象を記号的に表現される知識に限定し、かつ応用面に重点を置いて学習機能を考察する。

3. 知識の利用と獲得

前章では、(1)学習とは知識の獲得を通して、再試行時の性能が向上するようなシステムの変化であり、(2)知識の獲得は新しい情報の取得と、その体系化を含むことを述べた。(1)は、学習はシステムの性能と不可分であること、知識工学的に言えば、知識獲得は知識の利用と切り離して議論できないことを意味する。それには、(i)知識表現、(ii)知識利用方式、特に推論機能との関連を考察しておくことが必要である。

3.1 知識表現

このうち知識表現は知識処理の最も基本的な概念であり、それを基盤に設計される知識処理システムが、知識利用—獲得を含めた総合的なシステムとして、所要の性能を発揮することを保証するものでなくてはならない。もしその表現能力が不足であったり、必要な推論機能が十分に発揮できないものであったりするとシステムの性能に直接的に影響する。

逆に、知識表現は利用と獲得を含めたシステムの分析によって定められる要求を満たすように設計されるべきものである。知識表現については本特集号に含まれているので、本稿では主として上記の項目(ii)について考察する。なお参考文献22)に主として知識利用の面からシステムを分析し、知識表現への要求を抽出し、かつそれを満たすものについて議論がなされているので参照されたい。

3.2 知識の利用—単調推論と非単調推論

知識利用の中心に推論がある。知識処理で主要な役割を演じているのは演繹推論であるが、これは既存知識のもつ意味の範囲内でこれを利用する機能である。したがって対象とする問題領域をカバーするだけの完全な知識が得られていれば演繹推論で十分であるが、多くの場合、必要な知識をすべて集めることは困難である。そこで、存在する不十分な知識のみを用いて、それが含む内容以上の結論を導き出す推論が考えられ

た。それは非単調推論と呼ばれ、それを含む論理は非単調論理と呼ばれる。非単調論理は「新しい公理を持ち込むと、それ以前の定理が成立しなくなる可能性を含む論理¹⁸⁾」として定義されている。これは演繹推論に基づく従来の論理が「既存公理のもとで証明された定理は新しい公理の追加に無関係に常に真である」という単調性を満たすのに対するもので、単調性の制約を受けない論理として非常に包括的に定義されている。したがってここには従来の論理を超えた多種の論理が含まれる。これらは推論方式の相違により、類似論理、帰納論理、default 論理²⁴⁾、truth-maintenance システム⁸⁾、circumscription¹⁷⁾、closed world assumption²³⁾などと呼ばれる。このうち前2者は他のものと多少異なるので取り扱いも異なり、通常、非単調論理と言うときはこれ以外のものを指すことが多い。

非単調論理はいまだ完全な体系を形成するまでには到っていないが、これは通常の論理(単調論理)での知識の利用と獲得の中間に位置する機能をもつものとして知識獲得の問題を論ずる上で興味深いものである。以下では、まず通常の(単調)論理の場合の証明可能な論理式(以下、単に式という)の組の形式的表現を示し、次いで非単調論理の場合を求め、これらの比較を通して知識獲得の意味を考察する。

3.3 単調論理における証明可能性

一般的に論理は言語、公理系および推論規則から成る^{3),9)}。言語は構文として定数、変数、関数を含む「項」、「述語」および「式」を定義する。公理は普遍的な式の関係を示す基本公理式(logical axiom)と、そこに外部から付加された式(Proper axiom または non-logical axiom)から成る。単調論理における推論規則は① $[A, A \rightarrow B]/B$ (A と $A \rightarrow B$ から B が導かれる)で表される「モーダス・ポネンス」と、② x を自由変数としたとき、 A から $(\forall x)A(x)$ を推論する「一般化規則」である。(知識処理に対応させると、公理系は知識ベースで、推論規則は推論マシンで実現されている。ただし、基本公理は推論マシンに組み込まれ、明示的に知識として表現されるのはそれ以外の公理である場合が多い)。

A を公理系、 S を任意の式の集合とし、 σ を S と A から推論規則により導かれる式とする。これを $S \vdash_A \sigma$ と書く。 A として基本公理のみを考慮したとき、単に $S \vdash \sigma$ と書く。また $Th(S) = \{p : S \vdash p\}$ と定義する。 $Th(S)$ は基本公理と推論規則を繰り返し適用することによって、ある与えられた式集合 S から導かれるす

すべての式（これを S の定理と言う）の集合を表している。 A, B によって、異なる公理系を表し、 S を A または B とすると、単調論理の特徴は

- (a) $A \subseteq Th(A)$
- (b) もし $A \subseteq B$ なら、 $Th(A) \subseteq Th(B)$

で表される単調性と、

- (c) $Th(Th(A)) = Th(A)$

が常に成立するという、良い性質をもっている。(c) は公理集合から単調的に導かれる定理の集合が、推論規則のもとで式の閉包を求める演算に関して固定点を形成することを示している。実は単調論理に関してはさらに、 $Th(A)$ は最小固定点であることが示されている。

$Th(A)$ は知識ベースを公理系とすると演繹推論のもとでこの知識ベースから導かれるすべての式を表している。すなわち、 $p \in Th(A)$ なる結論 p のみがそれから導かれる。したがって、 $Th(A)$ が解こうとする問題の全領域をカバーするだけ十分に大きければ良いが、多くの場合、知識 A が不十分である。

3.4 非単調論理における証明可能式

単調論理のもとでは、 $p \notin Th(A)$ なる結論は得られないが、これに対し、“ p を否定する明確な根拠がないければ p を仮の結論とする” 推論を考えることができる。これは“ p が真であると証明されること” をもって結論の条件とする単調論理に対し、“ p が偽であるという証明がないこと” を結論の条件としており、前者を含んでいる。これを体系化することは単調論理の場合ほど容易ではない。単調論理も非単調論理も、いずれも式で表された公理や定理についての関係が示されているが、これには言語が必要である。ここでは一階述語言語を想定している。以下これを L とする。

非単調論理演算子 NM を次のように定義する¹⁸⁾。任意の式の集合 $S \subseteq L$ に対し、公理系 A のもとで S から作られる仮説式の集合 $AS_A(S)$ を定義する。

$$AS_A(S) = \{Mp : p \in L \text{ かつ } \sim p \notin S\} - Th(A)$$

Mp は“ p と仮定しても矛盾は生じない”を表す。 $\{ \}$ は言語 L で表現される式 p で、その否定が、与えられた式集合 S に含まれていないものの集合を表す。 S は仮説式の除外条件を与えている。すなわち S と矛盾しない限り Mp である。このような Mp の集合には単調論理の場合の $Th(A)$ も含まれるが、これを除いた残りは“単調論理では扱えないが非単調論理の枠組に含められている部分”である。これを用いて、単調論理の場合と同様にして証明可能式集合を

$$NM_A(S) = Th(A \cup AS_A(S))$$

と定義する。もしこれが最小固定点 $S = NM_A(S)$ を持つものとする、

$$TH(A) = NM_A(S) \text{ の最小固定点}$$

と定義された $TH(A)$ は単調論理の $Th(A)$ に対応するものとなる。しかし非単調論理では NM_A が最小固定点を持つという保証はない、このため定義できるのは

$$TH(A) = \bigcap_s (\{L\} \cup \{S : NM_A(S) = S\})$$

までであるとされている。

非単調論理の議論は与えられた公理系からいかに結論を導くかという議論であり、知識利用の問題であるからこれ以上細部には立ち入らない。本節の主旨は非単調論理によって公理系から導出可能とされる式の範囲は、単調論理の場合には知識獲得を通して公理系 (=知識ベース) を拡大することによってのみ処理可能となる範囲であるという事実から、知識の利用と獲得の関係を見直すことにある。

3.5 知識の獲得—不完全知識の完全化

図-1 は単調論理および非単調論理について定義された諸概念の関係を示したものである。この図で A が公理系、すなわち知識ベースを示す。斜線部分は単調論理で処理可能な部分、太線で囲まれた部分は非単調論理により処理可能な範囲である。この差に相当する部分は従来の単調論理の場合、処理不可能な部分である。この部分を処理可能にするには知識獲得により A を拡大することによって $Th(A)$ を拡大するほかない。すなわち非単調推論の機能は知識獲得の結果として得られる機能と重複している。

一方、公理系 A が拡大していても図-1 の太枠が拡大することはないから、原理的にはいつか斜線部分と太枠部分がほとんど同じものとなる。斜線部分は推論により完全な結論を保証している部分、太枠内は結

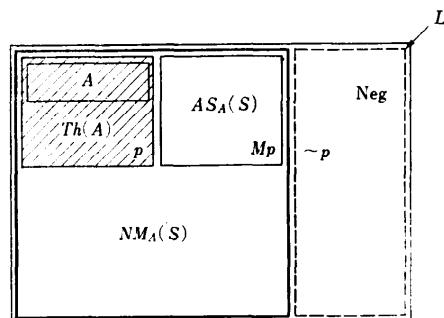


図-1 推論による証明可能領域

論が、新たな情報を得ることによって変わる可能性があるという意味で不完全ではあるが、既存の情報を最大限に利用している部分であるから、知識獲得は不完全情報を完全化する機能ということができる。単調論理では知識獲得は不可能な結論から一挙に完全な結論にまで変えるものだったのに対し、これは中間段階を置いたことになる。さらに今後、情報の効果的利用を図るために知識自体に曖昧性を許すことになると、この中間段階における情報の完全性の程度も一層細分化され、知識獲得は情報の完全化の機能として位置づけられることになろう。このことは従来経験的に行われてきた知識獲得に一つの統一的な視点を確立し、将来、獲得された情報が、知識ベースの完全化への寄与の程度によって評価されるようになるものと予想される。

3.6 知識の評価

簡単な例として、表現できる式の数が有限の言語 L を考えよう、この数を M とする。この一部が恒真および恒偽（図-1の斜線部分および Neg 部分）として与えられ、その部分の式の数をそれぞれ K および N とする。このとき、真偽の定まっていない式の数は $M - (K + N)$ である。個々の式は中間的な曖昧性なしに直偽の値を等確率でとるものとする、系全体としてはこの時点で $2^{M-(K+N)}$ 通りの可能性を等確率でもっている。すべての可能な表現の真偽値が定まったとき、系の状態もユニークに定まるが、これは、知識を用いて曖昧性なしに結論が求まる完全な状態である。これに対し上述の状態は系として不完全なものである。このような系の評価を定量化する方法としてよく用いられるエントロピーの概念を導入してみよう^{33), 34)}。すなわち $\Sigma 2^{-p} \log 2^{-p} = P, (P = M - (K + N))$ で系の不完全性を表すことにする。このとき、新しい知識 p が A に加えられたとしよう。前述の通り $Th(A \cup \{p\})$ が求まり、斜線部分が新しく定まる。集合 X に含まれる表現の数を $|X|$ と表すと、 p によって新しく恒真となる部分の数は

$$|A \cup \{p\}| + |Th(A \cup \{p\})| - |A| - |Th(A)| \\ = 1 + f(A, p)$$

である。 $f(A, p) = |Th(A \cup \{p\})| - |Th(A)|$ は p と既存知識との相互関係によって定まる $Th(A)$ の増加部分であり、 p, A の両方に関わる。これの大きな知識は系の不完全性をその分だけ減少させるものとして価値がある。

以上は極端に単純化した例であるが、知識そのもの

に曖昧性のある場合でもエントロピー評価が適応できよう。もちろんこれが唯一の評価ではないが、知識獲得をより合理的な根拠に基づいて論ずる上でこの種の評価法の確立が必要と考えられる¹⁴⁾。

4. 知識獲得の方法

4.1 知識獲得の諸段階

知識獲得は外部からの情報の取り込みと自システム内での体系化という2つの機能から成るが、学習システムの備えている推論の能力によりさまざまな知識獲得の形態があり、取得し得る情報の形式も異なる。知識が利用されるための知識表現形式がシステム内で定められているから、学習システムがどの程度、入手した情報を知識化する能力を有するかによって、取得する情報の形式が定まるからである。もし学習システムにこの能力がまったくなければ、すべて人が外部で準備せねばならず、推論能力が増す程、人の負担は減少する。

なお知識獲得の際に学習システムに要求される機能はそのシステムの構成によって異なってくる。以下では知識工学システムとして、少なくとも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備えた図-2の構成のシステムを想定する。この形式では一定の知識表現形式の知識が知識ベースに加えられればそれは利用可能である。したがって学習機能に要求されるのは外部から与えられる情報を知識表現に変換して知識ベースに与えることである。

以下、この学習機能を能力的に段階別に分類してみよう。

A 推論なしの情報取得

- (i) プログラム入力
- (ii) 事実データ入力

B 外部情報源からの知識化された情報の取得

- (i) 内部形式の知識の直接取得
- (ii) 内部形式の知識のインタラクティブ取得
- (iii) 外部形式で表された知識の取得

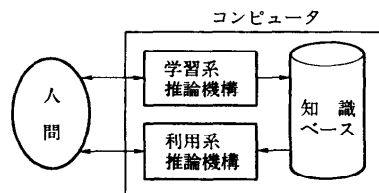


図-2 知識処理システム基本構成

C 事例からの学習

(i) パラメータ学習

(ii) 類似推論による学習

(iii) 帰納推論による学習—発見的学習

D メタ・レベル知識獲得

4.2 無推論学習および低レベル推論学習

Aは無推論学習(Rote learning)と呼ばれるもので、情報を取得するという意味でここに含まれているが学習のための推論機能はまったく必要なく、与えられる情報はプログラムやデータなど、そのままの形で問題解決に用いられる。すなわちこれは従来のコンピュータ利用のスタイルのもとでの情報取得形式である。

Bは外部から知識の形式、すなわち利用のための推論が適用できる形式で表された情報取得である。学習システムは、少なくとも入力情報を、利用可能な形式で知識ベースに組み込むための変換処理を必要とする。

この段階の知識獲得も、(i)は外部で人が内部形式として準備した知識を取り込む最も単純な形式で、エキスパート・システムで最初に与える専門家知識の多くはこの形式のものとなる。実用型の知識工学システムでは、これでは分野ごとの専門知識を機械内の表現形式で表現せねばならず、そのために対象分野と知識工学の両方に精通した仲介者を必要とする。これを改善するにはシステム自体にこの翻訳を代行させるほかない。一般にはこれにはかなり高度の推論を必要とするが、それを避け、比較的lowレベル推論で行い、他は人に任せる方法がインタラクティブな知識の獲得である⁴⁾。この例としてTEIRESIASが我が国でもよく知られている。これは医療分野でのコンサルテーション・システムとして開発されたMYCINを対象として、問題分野の専門家が学習システムの先生という立場で、生徒である知識工学システムが与えられた問題に対して出す解答を見ながらその誤りの原因となっている知識ベース内の規則を修正してゆくというプロセスを実現するものである。エキスパート・システムで知識を準備するためには一般に知識エディタのような支援手段を必要とするが、インタラクティブな知識獲得は単に個々の規則や事実を統語論的に編集するのみでなく、既存の知識の不備を補うなど、知識ベースの編集を行う点に特徴がある。

上記(i)および(ii)は入力に内部形式である知識表現にもとづいて行われるが、知識が自然言語などの外

部表現で与えられる場合、これを知識表現の形式に変換せねばならない。このためには自然言語や図形を含む外部表現の理解が必要となる。実際、知識獲得と理解とは密接に関連している。この場合も問題は単に構文の変換では済まず、与えられた表現を利用できる形に変換せねばならない。

さらに、問題の解法にたいする助言などを内部で利用できる形式に変換する問題があり、これは“可操作化”(Operationalization)と呼ばれることがある。これは問題領域の言語で表された助言などを、最終的には手続きに変換するまでの範囲を含み、人工知能の中心的課題でもあることから、多くの研究がなされている^{19), 20), 32)}。たとえばUNDERSTAND²⁸⁾はハノイの塔の問題の英語による表現を、適切な状態表現とそれへの操作を構築することにより可操作化を行っている。

4.3 事例からの知識獲得に関する基本概念

Bの範疇の諸問題は形式は異なっても、知識表現と同等の記述が外部から与えられている。ここで必要とされるのは情報表現の形式の変換であって、発見ではない。これに対しCの範疇の知識獲得の特徴は個々の事実を集め、それを一般的な知識として利用できるようにすることである。このためにはその事例を生じる背景となっている概念を見出し、それを表すための形式あるいは情報の構造を作り出さねばならない。このクラスに入るものも、推論機能の程度によって比較的容易なものから実現の困難なものまでさまざまである。

なおCの範疇は入力情報が事実を表す事例であるという点でBと区別しているので、これ以外の点では他の範疇のものと共通の性質をもつものもある。たとえば事実データが外部表現で与えられたときにはその理解が必要であるが、これはB-(iii)と同様である。

Cの範疇の上位のものは2.2節(3)項で述べた発見的学習に相当するもので、それだけに実現は困難である。事実、真の意味で発見的学習は実現していない。

A, Bの範疇に比し、Cは学習の方法に大きな自由度がある。したがってまずこの学習に関わる基本的な概念を挙げておこう。

(1) 表現言語; 事例による学習は個々の事例を集めて一般規則化するプロセスであるから、事例と一般規則の両方を表現する共通の言語が必要である。学習結果である一般規則は知識利用の対象とならねばならないから、これは知識表現言語である。逆に言えば知

識表現言語は上記の知識獲得の条件を考慮して定めねばならない。

(2) 対象記述の方式; 実例による学習は個々の対象についての記述から、あるクラスの対象についての記述を作るものと言えるが、この対象のクラスをどのように記述するかは(1)の表現言語とも関連する重要な問題である。少なくとも、ある対象の全体のクラスの中に、共通の性質をもつより小さなクラス(このクラス内のすべての対象が、そしてそののみが与えられた性質をもつもの)を定義することが必要であるが²¹⁾、現状はもっと単純に、一つの表を定義し、その中のすべての対象が共通の性質をもつことのみを定義するのが普通である。ある種の問題にはこの表に含まれないことを積極的に表現する偽例もしくは反例が効果的に用いられる。しかし非単調論理の議論と同様、「偽」の条件を扱うのは一般には難しい。制御された反例を用いた例としてニアミスがよく知られている³⁵⁾。

(3) 一般化規則; 個別例を集めて一般的規則を作る上で一般化規則が必要となる。これは言語の中の定数を変数で置き換えるとか、制約的に働いている条件記述のいくつかを除去するなど、いくつかの方法が考えられている。これも(1)の言語と密接に関係することと言うまでもない。

なお、一般化も上例のようなものはすでに記述の基本形が与えられていて、その範囲内で変数化すべき定数や削除すべき記述を選択するという性質のもので、これは比較的容易である。さらに困難なのは前もって与えられている記述以外の新しい記述を作らねばならないときで、このための規則はいまだ一般論としては確立されていない。前者を選択的汎化則と呼び、後者を構成的汎化則と呼ぶ。前者を用いた学習システムは多いが、後者の例はそれほど多くはない。

(4) 学習制御; 学習が事実からの規則化、さらには情報の構造化とすると、その構造化にもさまざまな戦略があり入力に応じてその制御が必要になる。典型的なものは個別情報を順次取り込んで構造化を進めるボトムアップ法と、仮説を作って、それを個別情報によって修正するトップダウン法である。実際にはこれらを組み合わせて、最も効率の良い学習制御を行うのは容易な問題ではない。

4.4 実例からの学習の諸段階

(1) パラメータ学習; 実例もしくは観測値からの学習の最も単純な形式は推論結果となるべき一般規則

の形を定め、そこに含まれるパラメータ値を逐次的にデータによって修正してゆくものである。これは心理学の学習モデルや学習制御系などに用いられている。

知識工学の分野でこの範疇の学習システム例として Meta-DENDRAL²⁾ が挙げられる。これは訓練あるいは実測のスペクトル・データから、プロダクション・ルールを修正しつつ、新しい規則を導き出そうというもので、パラメータ学習としてもやや高度のものであるが、知識の基本構造が与えられており、それが個別データにより逐次的に修正されるという点でここに含める。

(2) 類似推論による学習; 求めようとする概念に似た既存の知識を変換して新しい概念を獲得することができる。これを類似推論あるいは単に類推による学習という。人間の場合でも類推によって新しい概念や技術を獲得する例は非常に多く、重要な機能である。たとえばエレクトーンを弾いたことのない人も、ピアノを弾けるならその技術を転用して、より速やかに新しい技術の修得を行うことができる。

類似推論は人工知能の主要な研究対象の一つであり、P. H. Winston らの興味ある成果がある^{16), 36), 37)}。Winston は“2つの状況がいくつかの点で類似しているなら、他の点でも類似している”という仮定に基づく、類似推論を用いる。2つの状況の類似性は、その状況のもとでの重要性に従った最良の一致を見出すことで見分けられる。

2つの状況 S_1, S_2 を考えよう。 S_1 では $r_1(a, b) \Rightarrow r_2(c, d)$ で表される。“ r_1 は r_2 を生ぜしめる原因である”が成り立っている。 S_2 では $r_1(a', b')$ が判っているとする。通常の演繹的推論ではこれのみではなんの結論も得られないが、類似推論では $r_1(a, b)$ と $r_1(a', b')$ 間に類似性を見出すことにより、 S_2 にも結果 $r_2(c', d')$ を推論する。 $r_1(a, b)$ と $r_1(a', b')$ の類似性は、ある第3の表現 (W, u) を見出すことにより判定される。 W は両状況の比較される表現に対応させる第3の表現で、これは変数を含んでいる。 u はこの W の変数に代入される値の対(ついで)を表し、一方の値を W に代入すれば S_1 が、もう一方を代入すれば S_2 が得られる。上の例では $(W, u) = (\{r_1(X, Y), \{X - \langle a, a' \rangle, Y - \langle b, b' \rangle\}\})$ とすればよい。このような (W, u) が求まったとき、 $r_1(a, b)$ と $r_1(a', b')$ は類似であり、これが原因となっている関係の結果、 $r_2(c', d')$ が導出される。

最近、原口・有川はこの類似推論の形式化を行い、

これをさらに発展させ、演繹推論を組み込んだ推論アルゴリズムを求め、類似推論の理論と共に実用化へ一歩近づけた^{10), 11)}。上記の形式的表現も参考文献11)による。

(3) 帰納推論による学習—発見的学習

学習の中でも非常に高度の推論を用いるものが発見的学習である。この範囲でも人が学習システムの状況を見ながらデータを与える教師あり学習から、入力データがまったくランダムに与えられる教師なし学習、あるいは正のデータのみが与えられる場合と、否定のデータが含まれる場合などによって推論も異なり、難しさの程度も異なる。いずれにしてもこの範疇の学習は新規則の発見、理論の形成、構造の形成などを含み、しかも、形成されるべき理論や構造の原形が与えられていないため、与えられる正のデータおよび反例をすべて説明できるものを作り出さねばならない。このような構成的推論を目指した研究も散見されるが、一般論としてこの構成的推論を実現したものではなく、今後の課題として残されている。

推論の結果の表現の一部を入力情報の表現から求める形式の帰納的推論は可能であり、これを用いて、具体例から、それを満たすプログラムを生成するプログラムが最近注目を浴びている^{25), 26)}。この方向の研究も今後発展するものと予想される。

4.5 メタ・レベル知識獲得

以上はオブジェクト・レベルの知識獲得であるが、オブジェクト・レベルの知識を用いて問題解決を行う際の制御情報を始めとするメタ・レベルの知識の獲得の問題はさらに困難な問題である。メタ・レベル知識に関してはそれを表現し、利用する方式、あるいはオブジェクト・レベル知識との関係などシステム化の技術がまだ確立されていない。利用面からの表現形式が定まらない限り、実用的な意味でのメタ・レベル知識獲得を確立することは困難であり、これは知識工学自体の進歩を待つほかない。

5. む す び

知識獲得を含めた機械学習の問題は一部の特定分野での応用を除き、汎用技術としてはこれまでは主に学問的興味から研究対象とされてきた。知識工学の発達に伴い、この状況は変化しようとしている。知識工学の最大の特徴の一つは情報のモジュラリティにある。モジュラリティを最も必要とするのは知識獲得であり、これが最大の特徴とされることは知識工学が本

来、知識の利用と獲得を一体化することにより真に効果を発揮するものであることを意味している。

学習および知識獲得は多くの未解決の問題を残しており、今後の積極的な研究が望まれる。

参 考 文 献

- 1) Buchanan, B.G., Feigenbaum, E.A. and Lederberg, J.: A heuristic Programming Study of Theory Foundation in Science, Proc. Second IJCAI, London, pp. 40-48 (1971).
- 2) Buchanan, B.G. and Feigenbaum, E.A.: DENDRAL and Meta-DENDRAL: Their Applications Dimension, Artificial Intelligence, 11, pp. 5-24 (1978).
- 3) Chang, C.L. and Lee, R.C.T.: Symbolic Logic and Mechanical Theorem Proving, Academic Press (1972).
- 4) Davis, M.: Interactive Transfer of Expertise: Acquisition of New Inference Rules, Proc. fifth IJCAI, pp. 321-328 (1977).
- 5) Davis, M.: The Mechanisms of Non-Monotonic Reasoning, Artificial Intelligence, 13, pp. 73-80 (1980).
- 6) Davis, R.: Knowledge Acquisition in Rule-Based Systems; Knowledge About Representations As a Basis for System Construction and Maintenance, in Pattern-Directed Inference Systems (eds. Waterman, D.A. and Hayes-Roth, F.), Academic Press, pp. 99-134 (1978).
- 7) Dietterich, T.G. and Michalski, R.S.: A Computer Review of Selected Methods for Learning from Examples, in Machine Learning (eds. Michalski, R.S., Carbonell, J.G. and Mitchell, T.M.), Springer-Verlag, pp. 41-82 (1984).
- 8) Dogle, J.: A Glimpse of Truth Maintenance, Proc. Sixth IJCAI Aug. 1979, Tokyo, pp. 232-237 (1979).
- 9) Enderton, H.B.: Mathematical Introduction to Logic, Academic Press (1972).
- 10) Haraguchi, M.: Towards a Mathematical Theory of Analogy, Bull. Inform. Cybernetics, Vol. 21 (1985).
- 11) Haraguchi, M. and Arikawa, S.: Analogical Reasoning Based on the Theory of Analogy, Res. Rept. Res. Inst. Fund. Inform. Sci. Kyushu Univ., No. 105 (1985).
- 12) Hays-Roth, F.: A Structural Approach to Pattern Learning and the Acquisition of Classification Power, Proc. Fifth Int. Joint. Conf. on Pattern Recognition, pp. 343-355 (1973).
- 13) Hays-Roth, F. and McDermott, J.: An Interference Matching Technique for Inducing

- Abstractions, *Comm. ACM*, Vol. 21, No. 6, pp. 401-410 (1978).
- 14) Hilpinen, R.: Rules of Acceptance and Inductive Logic, *ACTA PHILOSOPHICA FENNICA FASC XXII* North-Holland (1968).
 - 15) Kitakami, H., Kunifuji, S. Miyachi, T. and Furukawa, K.: A Methodology of Knowledge Acquisition System, *Proc. 1984 International Symposium on Logic Programming*, Atlantic City, pp. 181-192 (1984).
 - 16) Kling, R.E.: A Paradigm for Reasoning by Analogy, *Artificial Intelligence*, 12, pp. 147-178 (1971).
 - 17) McCarthy, J.: Circumscription—A Formal of Non-Monotonic Reasoning, *Artificial Intelligence*, 13, pp. 27-40 (1980).
 - 18) McDermott, D. and Doyle, J.: Non-Monotonic Logic I, *Artificial Intelligence*, 13, pp. 41-72 (1980).
 - 19) Mostow, D.J.: Machine Transformation of Advice into a Heuristic Search Procedure, in *Machine Learning—An Artificial Intelligence Approach* (eds. Michalski, R. S., Carbonell J. G. and Mitchell, T. M.), pp. 367-403 (1984).
 - 20) Newell, A., Shaw, J. and Simmon, M. A.: Empirical Explorations of the Logic Theory Machine: A Case Study in Heuristics, *Proc. 1957 WJCC*, pp. 218-230 (1957).
 - 21) Ohsuger, S.: A New Method of Model Description—Use of Knowledge Base and Inference, in *CAD System Framework* (eds. K. Bø and Lillehagen, F. M.), North-Holland, pp. 285-312 (1983).
 - 22) Ohsuga, S.: An Integrated, Intelligent, Interactive and Incremental Model-Based Problem Solving System, *Proc. Artificial Intelligence and Advanced Computer technology*, Sept. 1985, Wiesbaden (1985 (To appear)).
 - 23) Reiter, R.: On Closed World Databases, in *Logic and Databases* (eds. Gallaire, H. and Minker, J.) Plenum Press, pp. 56-76 (1978).
 - 24) Reiter, R.: A Logic for Default Reasoning, *Artificial Intelligence*, 13, pp. 81-132 (1980).
 - 25) Shapiro, E. Y.: Inductive Inference of Theory From Facts, T.R. 192, Yale Univ. Dept. Comp. Sci. (1981).
 - 26) Shapirs, E. Y.: *Algorithmic Program Debugging*, MIT Press (1983).
 - 27) Shinohara, T.: Inductive Inference From Negative Data, *Bull. Inform. Cykernetics*, Vol. 21 (1985).
 - 28) Simon, H. A.: Artificial Telligence Systems That Understand, *Proc. Fifth, IJCAI*, pp. 1059-1073 (1977).
 - 29) Simon, H. A.: Why Should Machines Learn?, in *Machine Learning—An Artificial Intelligence Approach* (eds. Michalski, R. S., Carbonell, J. G. and Mitchell T. M.) Springer-Verlog (1984).
 - 30) Sleeman, D. H. and Smith, M. J.: Modelling Student's Problem Solving, *Artificial Intelligence*, 16, pp. 171-187 (1981).
 - 31) Sleeman, D. H.: Inferring Student Models for Intelligent Computer-Aided Instruction, in *Machine Learning An Artificial Intelligence Approach* (eds. Michalski, R. S., Carbonell, J. G. and Mitchell, T. M.), Springer-Verlag, pp. 483-509 (1984).
 - 32) Tappel, S.: Some Algorithm Design Methods, *Proc. AAAI 80*, Stanford Univ. pp. 64-67 (1980).
 - 33) Emden, M. Van.: Hierarchical Decomposition of Complexity, *Machine Intelligence* 5, pp. 361-380, Edinburgh Univ. Press (1969).
 - 34) Watanabe, S.: *Knowing and Guessing: A Formal Quantitative Study*, John Wiley and Sons (1969).
 - 35) Winston, P. H.: Learning Structural Descriptions from Examples, in *the Psychology of Computer Vision* (ed. Winston, P. H.) McGraw Hill, ch. 5 (1975).
 - 36) Winston, P. H.: Learning and Reasoning by Analogy, *CACM*, Vol. 23, pp. 689-703 (1980).
 - 37) Winston, P. H.: *Leaning New Principles from Precedents and Exercises*, *Artificial Intelligence*, 19, pp. 321-350 (1983).

(昭和60年8月15日受付)