

知識の構造化を重視した 学習のモデル

赤間 清

(北海道大学 文学部 行動科学科)

1. まえがき

人間の知的情報処理の全体像を把える理論を構築することは非常に困難な目標であり、現段階で細かいレベルまで具体化された動くモデルを作ることはもちろん不可能である。しかし逆に、細かい現象レベルで人間に似せたモデルを作る研究方法を採用したとしても、全体の構造に対する的確な考察を積極的に推進しなければ、モデルの全体の構造は人間の知的情報処理の構造とは全く異なるものになる可能性がある。

人間の知的情報処理の全体構造への手がかりとして何があるだろうか。人間の行なう非常に高度な知的情報処理を計算機上で実現しようとするとき、そのシステムの知的情報処理の全体構造は、利用できる時間的、空間的資源の範囲内で高度の機能実現を可能とするために満たすべき条件によって制約を受ける。人間の脳もまた現に動いている実在のシステムであるので、その知的情報処理の全体構造がそれらの条件のいくつかに束縛されていることは十分予想される。従って、人間の知的情報処理の持つ大きな構造を把えるための有望な方法の1つは、人間に要請される基本的な知的情報処理とはどのような種類のものがあるかを考察し、それらを実現するために利用できるアルゴリズムや知識表現の時間的、空間的コストを推定し、可能な実現の方法を比較検討することである。このようにしていろいろな実現方法の相対的な位置付けを与えれば、その知見は、実際の人間の情報処理を考えるうえで大きな役割を果たすと予想される。

この方法に基づいた考察とモデル作成 1) - 4) を継続してきた結果、我々は、人間の知的情報処理の主要が「帰納的学習」のための情報処理によって決定されているという仮説をより強く確信するに至っている。本論文では、その仮説をより詳細化、具体化するために、未知の言語による学習を扱う形式的な設定を改めて提案し、そのもとで動く帰納的学習システムLS/1を構成する。

我々が帰納的学習という言葉によって何を意味するかを簡単に説明する。人間は生まれたばかりの時には、他の人間と1単語の意味すら共有していない。それにもかかわらず人間は、ある言語世界の中で生活するうちにその言語を修得し、世界のいろいろな知識を獲得する。このことは、人間同士を結ぶ最も基礎的な基盤は言語ではないことを意味する。言語よりも下のレベルにあり、言語をも獲得させるような情報処理こそ、人間が持つべき基礎的で重要な知的情報処理であり、それは実は言語の獲得期だけでなく、人間の生涯の日常的な情報処理すべてを実現する基盤である、と我々は考える。本論文での帰納的学習とはそのような視点からみた人間の情報処理の総称である。これから見れば、狭義の帰納的学習、すなわち、具体例から規則を発見するアルゴリズムは、そのごく一部分にすぎない。

乳児の置かれた状況を考えてみると、任意の単語はほとんどどんな意味にも使われる可能性を持つ。また、乳児がこれから生きてゆく世界で、任意の情景が普通どのように分節化され、どのように意味付けられるのかもいろいろな可能性がありうる。このように乳児が学習すべき言語や世界の知識は莫大な自由度を持つ。乳児はそれらの多様性にうまく対処する生得的な能力を持たねばならない。その生得的な能力は、例えば自然言語処理の研究における知識表現のような、具体的な単語を用いて知識を記述する方法では付与できない。乳児の持つ「知識」は(いろいろな可能性のいずれにも適用可能という意味で)もっと一般性のあるレベルのものでなければならない。学習が人間にとって不可欠となる理由がここにある。学習を実現する方法でまず考えられるのは、あらかじめ生得的な枠組みを組み込んでおき、経験によってその中に各単語を位置付ける方法である。しかしそのような方法だけではまだ十分な学習能力は達成できまい。乳児の対処すべき状況を観察すれば、知識の枠組み自体の発見と構築が必要であることが推測されるからである。

我々はこう考える。乳児の行なう高い自由度の学習を実現するためには、少なくとも、受け取ったいろいろな情報から類似の対象、重要な構造を発見して抜き出し、一般化して仮説を作り、それらを積み上げて知識の体系を組み立てていく情報処理が必要である。我々は、言語学習や知識獲得など乳幼児期のごく日常的な情報処理は(意識下のレベルに)知識の構造化、仮説生成、一般化、発見などを必須の部分として含んでいると推測する。さらに我々は(人間の情報処理メカニズムの連続性を重視し)単に乳幼児期の学習だけでなく人間の一生のすべての活動の基礎にもそれらの情報処理が存在すると仮定する。それらの情報処理、特に、仮説生成や発見などの単語は(例えば専門的な研究をする科学者の思考などと結びついて)高級で非日常的なものともみなされることが少なくないが、ここでは基礎的で日常的なものとして位置付けていることを強調しておく。

帰納的学習は演繹的推論に比べて格段に空間的、時間的コストのかかる情報処理なので、帰納的学習が人間の知的情報処理の大意に大きな影響を与えていることが予想される。従って、人間の情報処理の全体像をバランス良くモデル化するためには本研究の視点が必要と思われる。

2. 学習状況の設定

我々は、次のような形式化された学習状況を考える。登場するのは教師と生徒である。教師は、どんな質問にどう答えるのがよいかすべて知っている。簡単のためそれは時間もに依存せず、対応 Γ で表現されると仮定する（質問に対する正解は任意個ありうるので写像では表現できない）。教師は生徒と質問 Q 、応答 R 、正解 A の3つのフェイズよりなる会話を繰り返す。フェイズ Q においては、教師は、対応 Γ から1つの質問文 Q_t を選んで生徒に与える。フェイズ R においては、生徒が質問文 Q_t に対して正しいと思う文のひとつを応答文 R_t として教師に答える。フェイズ A においては、教師が、質問文 Q_t に対する正解文の1つ A_t を生徒に教える。 A_t は次のように決定される。 R_t が、もし対応 Γ に照らして正しい応答ならば、 A_t は R_t に一致させる。また R_t が正しい応答でないならば、 A_t は Q_t に対する正解文のなかの任意の1つの文が選択される。教師と生徒はこのような会話 QRA を何回も繰り返す。生徒は毎回の会話において応答文を正解文に一致させればされるほど高い評価が与えられる。

より形式的な表現を用いて上の記述を明確化しておこう。すべての単語の集合を Σ とする時、 Σ 上の文を Σ の単語を1個以上並べた列で定義する。 Σ 上の文全体の集合を (Σ^+) と書く。教師の持つ、質問文と正解文の対応 Γ は $(\Sigma^+) \times (\Sigma^+)$ の部分集合と等価である。対応 Γ を用いれば、質問文 Q_t に対する正解文 A_t の決定の規則は次のように書ける。

(WHAT IS DOG) ----- (IT IS ANIMAL)	(1) $\Gamma(Q_i) \ni R_i$
(SAY EXAMPLE OF ANIMAL) ----- (CAT)	$\rightarrow A_i = R_i$
(DOG CAN RUN) ----- (YES)	(2) $\Gamma(Q_i) \ni R_i$
(HAWK IS BIRD) ----- (YES)	$\rightarrow A_i \in \Gamma(Q_i)$
(WHAT CAN BIRD) ----- (CAN FLY)	
(DOVE CAN FLY) ----- (YES)	
(WHAT IS DOVE) ----- (IT IS BIRD)	
(SAY EXAMPLE OF BIRD) ----- (DOVE)	
(WHAT IS HAWK) ----- (IT IS BIRD)	
(WHAT IS PIG) ----- (IT IS ANIMAL)	
(BIRD HAS WING) ----- (YES)	
(SAY EXAMPLE OF ANIMAL) ----- (PIG)	
(CAT CAN RUN) ----- (YES)	
(ANIMAL HAS LEG) ----- (YES)	
(WHICH IS ANIMAL : DOG OR DOVE) -- (DOG)	
(WHICH IS BIRD : HAWK OR CAT) ---- (HAWK)	

図1 訓練列 τ_1

Fig. 1 An example of training sequence τ_1

対応 Γ は実際にどのような順序で学習を進めていくかは示していない。問題の提示順序の情報まで含めた質問正解の無限列を訓練列と呼ぶ。簡単な訓練列の例を図1に挙げる。この訓練列 τ_1 は、左が質問文で右が正解文である。 τ_1 において問題文に対する正解はいくとりも想定しているが、ここでは記述の簡単のために、各問題に対して正解の例を1つだけ挙げてある。またこれは無限列の最初の部分である。

生徒は初めは対応 Γ を知らない。しかし一聞いて十を知る能力があれば、教師との質問応答を繰り返すうちに対応 Γ の様子をしいうまく推測し、以前に正解を教えてもらっていない質問にも正しく答えることができるようになるだろう。我々は、そのような生徒にあたるシステムをつくることを最も粗いレベルの問題設定とする。

3. 置換不変、構成的アプローチ

我々は、訓練列 τ_1 をうまく学習するだけでなく、例えば図2の訓練列 τ_2 の場合にもうまく学習することをシステムに要請する。訓練列 τ_2 は我々にとって理解しにくい。実は、 τ_1 の各単語に現われるアルファベットの順序を単語ごとに逆転して得られる。従って、既存の知識を仮定しなければ τ_1 と τ_2 は等価であると考えることができる。その意味で等価な訓練列に対する学習過程（学習の結果得られた、質問 Q_t 、応答 R_t 、正解 A_t の組 (Q_t, R_t, A_t) の無限列）が互に対応したものであるとき、システムは置換不変性を持つと言う。

置換不変性を明確に述べる。単語の集合を Σ 、 Σ 上の訓練列全体の集合を T 、 Σ 上の学習過程全体の集合を L とする。 Σ 上の任意の置換 p は、 T 上の置換と L 上の置換を引き起こす。それらも p で表わすことにする。システムの決める T から L への写像を S とする。システムが置換不変であるとは、任意の訓練列 τ と任意の置換 p に対して、

(TAHW SI GOD) ----- (TI SI LAMINA)
 (YAS ELPMAKE FO LAMINA) ----- (TAC)
 (GOD NAC NUR) ----- (SEY)
 (KWAH SI DRIB) ----- (SEY)
 (TAHW NAC DRIB) ----- (NAC YLF)
 (EVOD NAC YLF) ----- (SEY)
 (TAHW SI EVOD) ----- (TI SI DRIB)
 (YAS ELPMAKE FO DRIB) ----- (EVOD)
 (TAHW SI KWAH) ----- (TI SI DRIB)
 (TAHW SI GIP) ----- (TI SI LAMINA)
 (DRIB SAH GNIW) ----- (SEY)
 (YAS ELPMAKE FO LAMINA) ----- (GIP)
 (TAC NAC NUR) ----- (SEY)
 (LAMINA SAH GEL) ----- (SEY)
 (HCIHW SI LAMINA : GOD RO EVOD) -- (GOD)
 (HCIHW SI DRIB : KWAH RO TAC) ---- (KWAH)

図2 訓練列 τ_2

Fig. 2 An example of training sequence τ_2

がどの程度の一般性を有するものかを吟味し、極力 ad hoc 性を排除するように務めながら構成的なシステムを作成し、理論を改善して行く研究方法を置換不変、構成的アプローチと呼ぶことにする。学習とは何か、意味とは、文法とは、…などの根本的な問題をより深く突き詰めて考察するために、このアプローチは大きな役割を果たす可能性がある。知識を詰め込んで安易に人間に「似せる」わけにはいかないからである。

4. 帰納的学習システムの基本原理

最初に考えなくてはならないことは、一を聞いて十を知るための基本メカニズムである。我々は、次のようなアイデアを採用する。システムがある時点で

$\Gamma \ni ((\text{WHAT IS DOG}), (\text{IT IS ANIMAL}))$
 $\Gamma \ni ((\text{WHAT IS SALMON}), (\text{IT IS FISH}))$
 $\Gamma \ni ((\text{SAY EXAMPLE OF ANIMAL}), (\text{DOG}))$
 $\Gamma \ni ((\text{SAY EXAMPLE OF BIRD}), (\text{EAGLE}))$

という情報を得ていると仮定しよう。これは次のように書き直せる。

$\Gamma \ni \{((\text{WHAT IS } \$X), (\text{IT IS } \$Y)) \mid (\$X, \$Y) \in S_1\}$
 $\Gamma \ni \{((\text{SAY EXAMPLE OF } \$P), (\$Q)) \mid (\$P, \$Q) \in S_2\}$
 $S_1 \ni ((\text{DOG}), (\text{ANIMAL}))$
 $S_1 \ni ((\text{SALMON}), (\text{FISH}))$
 $S_2 \ni ((\text{ANIMAL}), (\text{DOG}))$
 $S_2 \ni ((\text{BIRD}), (\text{EAGLE}))$

ここで S_1 と S_2 の間の関係が、なんらかの根拠（例えば、(DOG) と (ANIMAL) のペアが S_1 と S_2 で逆順になって入っている）に基づいて、

$S_1 \ni (\$X, \$Y) \leftrightarrow S_2 \ni (\$Y, \$X)$

と推測されたとする。この推測は S_2 から S_1 へ、また、 S_1 から S_2 へ、新しい元を供給する。それらは、それぞれ、

$((\text{EAGLE}), (\text{BIRD})) \text{ into } S_1$
 $((\text{FISH}), (\text{SALMON})) \text{ into } S_2$

である。これは Γ に対して、

$\Gamma \ni ((\text{WHAT IS EAGLE}), (\text{IT IS BIRD}))$
 $\Gamma \ni ((\text{SAY EXAMPLE OF FISH}), (\text{SALMON}))$

という2つの推測を生み出し、

$$p(S(\tau)) = S(p(\tau))$$

が成り立つことである。

言語では、各単語がどのような発音やスペルを持つかに（大幅な）任意性がある。人間は、他人と1単語も共有していない誕生直後の状態から出発して、しだいに自分の知識体系を組み立てていく。もし英語とよく似た言語を話す架空の世界があって、英語の世界で SOME と言うところを OF、また、OF と言うところを SOME と言うだけの違いしかなければ、その架空の世界における人間の生涯の行動は、やはり、SOME と OF の逆転以外は実際の英語の世界と同じだろう。これは置換不変性に対応している。

システムに課された置換不変性はシステム作成に際して厳しい制約となる。置換不変性により、少なくとも単語レベルでシステムに知識を入れておくことは無効になり、与えられた質問応答 (QRA) を素材にして、一般的な操作だけで知識を組み立てていく構成的なやり方をするのが、十分な適応能力を確保する前提としてどうしても必要になる。置換不変性を尊重して、注意深く、何

(WHAT IS EAGLE)

(SAY EXAMPLE OF FISH)

という2つの新たな質問にうまく答え得る可能性をもたらす。

この例の示すアイデアを核として、置換不変、構成的アプローチにより、一般的かつ具体的なシステムに発展させたのが帰納的学習システムLS/1である。LS/1はよりよい応答をするために、各回の質問応答(QRA)から得た情報をもとにして仮説的知識体系を逐次的に作り、未知の問題に対する答えをよりの確に推測し、それに答える。

ここでシステムの一般性について特に注意したい。我々はLS/1によって直接的にある特定の学習過程を模擬しようとする訳ではない。システムの説明にはいろいろな例が用いられるが、それらは分かり易くするためにたまたま選ばれたものである。システムはそれが扱う単語(例えばDOG, ANIMALなど)によって人間が連想する意味内容のレベルとはまったく独立なレベルの原理で作られる。

そのようにして組み立てられたシステムに特定の「意味ある」訓練列を与えた場合、十分うまく扱い得る可能性はありうるのだろうか。我々はその可能性は十分にあると推測する。7節の記述は、この可能性を実現するために我々が提案する有望な糸口である。少なくとも我々は、その可能性の範囲をはっきりと見極めることこそが、知能を研究して行く上で最も重要な課題であると確信する。

5. LS/1の核となる情報処理の概要

(Σ +)のn個の直積の部分集合を、単に、n項関係、または、RELATION(以下RELと略記する)と呼ぶことにする。また、RELの要素をTUPLE(以下TPLと略記する)と呼ぶ。例えば、学習中の生徒が、それまでの質問応答によって、質問と正解の関係になり得る文のペアの集合

((WHAT IS CAT), (IT IS ANIMAL))

((DOG CAN RUN), (YES))

((WHAT IS CAR), (IT IS VEHICLE))

を得たとする。これは1つのREL(2項関係)であり、質問正解ペアの1つ1つがTPLである。このRELから、例えば、

((WHAT IS \$X), (IT IS \$Y))

の形を持つTPLを取り出し、変数\$X, \$Yに入るべき単語列の組を列挙すると、

((CAT), (ANIMAL))

((CAR), (VEHICLE))

という新しいRELが得られる。このように任意のRELから別のRELを生み出すものを一般に文脈と呼ぶことにする。

文脈やRELは無数に存在するが、知識をうまく構築したり、その結果新しい質問に対する正しい推測を導くために有効に利用できるものはそのうちの一部分だけである。システムの学習能力は、有効な文脈やRELをうまく選択する方法に大きく依存する。ただし、その方法は、システムに課せられた置換不変性の条件を損なわないような一般性を持たねばならない。

LS/1は、すでに得られたRELから新しい有用なRELを生成する可能性のある文脈を見つけ出す。その1つの方法はマッチングを用いる方法である。例えば、RELの2つのTPL

TPL1 = ((WHAT IS DOG), (IT IS ANIMAL))

TPL2 = ((WHAT IS CAR), (IT IS VEHICLE))

から、マッチングにより、文脈

CTX0 = ((WHAT IS \$1), (IT IS \$2)) $\rightarrow$ ((\$1), (\$2))

が得られる。

生成されたRELのうちからより有効なものを選び出すために、LS/1では次のようなヒューリスティックスを用いている。

- ・RELのTPLの数が多いほど重要
- ・応答生成に使われる回数が多いほど重要
- ・他のRELとの関連が強く推定できるほど重要

ここで他のRELとの関連とは、1つのRELが、他のRELから置換や合成などによって作られることを指す。例えば、図3で、REL1はREL2から(TPLの成分の)置換によって得られる。また、REL1はREL3とREL4の合成である。学習中に得られるRELは「不完全なもの」であり、厳密にこれらの関係が成り立つことは稀である。実際には、RELどうしの関係は、含まれるいくつかのTPLどうしの関係から推測される。

REL1	REL2	REL3	REL4
((DOG), (RUN))	((RUN), (DOG))	((DOG), (ANIMAL))	((ANIMAL), (RUN))
((CAT), (RUN))	((RUN), (CAT))	((CAT), (ANIMAL))	((FISH), (SWIM))
((SALMON), (SWIM))	((SWIM), (SALMON))	((SALMON), (FISH))	
REL5		REL8	
((DOG), (RUN))		((DOG), (ANIMAL))	
((HAWK), (FLY))		((CAT), (ANIMAL))	
		((CAR), (VEHICLE))	
REL6	REL7	((DUCK), (BIRD))	
((DOG), (ANIMAL))	((RUN), (ANIMAL))	((HAWK), (BIRD))	
((EAGLE), (BIRD))	((FLY), (BIRD))		

図3 RELATIONの例 (REL1 - REL8)
Fig. 3 Examples of relations (REL1-REL8)

例えば、図3の3つのTPL、

((DOG), (RUN)) in REL5
((DOG), (ANIMAL)) in REL6
((RUN), (ANIMAL)) in REL7

は、合成関係

REL5 = (EQL REL6) (REV REL7)

(これは、REL6と、REL7の逆転との合成がREL5に等しいという意味である)を推測する根拠の1つである。
これによりREL5の中に

((EAGLE), (FLY))

も入っていることが推測される。推測の根拠となる TPLの数は多いほどよい。またこの推測は(4節の原理に基づいて)新問題への応答を可能にする。

LS/1はRELから射影や逆像によっていろいろな集合を作り出す。例えば図3のREL8から射影により

集合1 = { DOG, CAT, CAR, DUCK, HAWK}

集合2 = { ANIMAL, VEHICLE, BIRD}

が得られる。また、集合2の要素の逆像により

ANIMALの逆像 = { DOG, CAT}

VEHICLEの逆像 = { CAR}

BIRDの逆像 = { DUCK, HAWK}

が得られる。

これらは、上記の、文脈とREL (n項関係)による抜きの特別な場合として位置付けることができる。
例えば、図3で、集合1を得る射影には、文脈

CTX3 = ((\$1), (\$2)) → ((\$1))

が対応する。またANIMALの逆像を得るには、文脈

CTX4 = ((\$1), (ANIMAL)) → ((\$1))

を用いればよい。

生成された集合は、互いの排反性や包含関係を重視して構造化される。その知識表現には階層構造が用いられ、それを利用してある集合とある集合が等しいことが推測される。このようにLS/1では、集合に対して(n項関係という共通の枠組みで考え得るもののほかに)その特性や重要性を考慮した抜きの実現している。

6. LS/1の知識表現

LS/1は、RELを扱う知識表現構造を3種類用いている。それらは標号網(LN)、関係データベース(RDB)、概念階層(CH)と呼ばれ、さまざまなRELをそれぞれの特徴にあわせて処理するために導入された。LNは、質問文と正解文の対応を表すRELのように、質問文や応答文に比較的近いレベルのn項のRELを扱う。RDBは、REL1-REL8のように、LNの扱うRELから生成され、TPLの各成分の単

```

i18 ----> 80 k2 (p v1 (($1 CAN $2)(yes)))
          (p v2 (($1)($2)))
i19 ----> 80 k2 (p v1 ((WHAT IS $1)($2)))
          (p v3 (($1)($2)))
i20 ----> 100 k1 (p v2 ((ANIMAL)(RUN)))
i21 ----> 100 k1 (p v2 ((FISH)(SWIM)))
i22 ----> 100 k1 (p v2 ((BIRD)(FLY)))
i23 ----> 100 k1 (p v3 ((DOG)(ANIMAL)))
i24 ----> 100 k1 (p v3 ((HAWK)(BIRD)))

```

図4 L Nにおける獲得知識の表現例 (部分)
Fig. 4 An example of acquired knowledge
(part) in the label net (L N)

```

REL9
((DOG),(ANIMAL))
((SALMON),(FISH))
((EAGLE),(BIRD))
EQL CTX5=((WHAT IS $1),($2)) -> (($1),($2))
EQL CTX6=((($1 IS A KIND OF WHAT CLASS),
          (IT IS AKO $2)) -> (($1),($2))
REV CTX7=((SAY EXAMPLE OF $1),($2))
          -> (($1),($2))
COM (EQL REL5) (EQL REL7)

```

図5 RDBにおける獲得知識の表現例
(部分) : REL9

Fig. 5 An example of acquired knowledge (part)
in the relational data base (RDB)

CHは包含関係を基礎として集合を表現する5)。

3種の知識表現を相対的に次のように位置付けることができる。RDBとCHは、LNに比べてより小さい (RELの各TPLの成分の数nや各成分の単語数が小さい) RELを扱う。小さいREL同士は関係が付きやすいので、あるQAの経験を他のQに応用するための連結点として利用できる。LNはRDBやCHに比べて、外部世界 (質問応答文) に近いRELを扱う。これらは内部に蓄積されたRDBやCHの知識を外部形式である文に結びつけるのに役立つ。これらは、RDBやCHがより意味的な構造を扱い、LNがより構文的な構造を扱うと、大まかに対照することもできる。

帰納的学習システムLS/1は、3種類の知識表現、LN、RDB、CHを管理する3つのサブシステム: 中央処理システム (CPS)、関係管理システム (RMS)、概念管理システム (CMS) からなる。

7. 帰納的学習の高自由度とその制御

LS/1の知識形成過程の基本は、有用なRELを次々に作り出す過程である。これは構造化、仮説生成、発見などに対応する情報処理を実現するために必須の過程である。広い範囲のRELを対象として、重要なRELを作り出し、それらの関係を適切に扱うことは、帰納的学習システムにとって本質的な特徴である。しかしここには重大な問題がある。それは、重要なRELの発見のための探索は組み合わせ爆発を引き起こす危険性が非常に高いことである。あらかじめシステムの扱う領域を限定して、領域固有の知識を埋め込むことによって探索範囲を限定する方法は、そのような組み合わせ爆発を回避するために現在知られるほぼ唯一の方法である。しかし、領域固有知識を利用する方法は置換不変性の要請によって利用不可能であ

語数が少ないn項のRELを扱う。CHは1項のREL, すなわち、集合を扱う。

LNはRELを頂点、RELどうしの関係を枝とするネットワークである。LNには常に特別な頂点V1が含まれている。それは質問文と正解文の対応Tに対する推測を表現するRELにあたる。枝には2種類ある。1つはRELの間に成り立つ関係の推測を表現する枝である。もう1つは単にRELどうしの関係を記述するだけでなく、あるRELから新しいRELを生成する役割を受け持つ特別な枝であり、文脈を表現する。

図4にLNの知識表現例の一部分を示す。LNの各枝にはその推測の確からしさを示す数値が付加されている6)。これによって、知識には各TPLごとに重みがつき、LNは全体として、1つの仮説的な知識を表現することになる。特に頂点V1は、任意の文S1と文S2に対して、それらが質問文と正解文のペアとしてどの程度確からしいかを示す知識に対応する。確実度の付加は、より良い応答をするための帰納的情報処理に是非とも必要なものである。

RDBはRELを推測された相互関係とともに管理する。その知識表現は、厳密にはいわゆる関係データベースとは違うが、ここでは便宜上RDBと呼ぶ。図5はRDBによる知識表現例の一部分である。相互関係はEQL、REV、COMなどの記号を用いて表わされている。このREL9では、CTX5、CTX6など、EQLで示される2つの文脈の生成するRELは、REL9と等しいと推測されていることを示す。また、REVで示される文脈CTX7の生成するRELは、TPLの成分の順序を逆転すれば、REL9と等しくなると推測されていることを示す。COMの後ろに書かれた式は、REL9が、REL5とREL7の合成だと推測されていることを表わす。

る。それに代わるより一般的な、組み合わせ爆発回避の方法はあるだろうか。我々は、獲得から利用までのいろいろな情報処理の巧妙で柔軟な連係の可能性を追及することが、RELの生成を含む情報処理の組み合わせ爆発を回避するための最も有望なやり方だと思う。

これは「人間の情報処理は（基本的には）それほど深くはないが、非常に広く柔軟である」という観察に対応する。おそらく人間は、数個のS式のペアから、それらに対応づける難しいLISP関数を一挙に構成するような狭くて深い帰納アルゴリズムなどは持っていないだろう。そのかわり、いろいろな種類（抽象度の違いを含む）の膨大な知識を関連づけてうまく操ることにたけている。

以下では、現バージョンのLS/1がこの方向へそって達成した情報処理の例を示す。LS/1は、すでに獲得した知識を利用して、新知識の構造化における組み合わせ爆発を抑制する。例えば、重要な文脈の発見の場合、最初は単に2つのTPLを任意に選んでマッチング3)し、その結果得たマッチングの良さを基準にして知識を構築するために使う文脈を選択する。しかし、知識がたまってくるとそのような初歩的なやり方から段々うまいやり方に移行し、より重要度の高いと思われる文脈をより少ない手間で見つけることが可能になる。それには、どの単語がどこに現われるかを示す単語帳(CPSが作る)や、文脈になり易い単語と重要なRELのTPLに現われる単語とを区別する知識(RMSが作る)や、RELの重要度の知識(RMSが作る)などが使われる。例えば5節の文脈発見の例で、もしDOGやANIMALが重要な単語であるとRMSがすでに推測していたり、あるいは更に強く、

((DOG), (ANIMAL))

の属するRELは重要であると信じている場合には、TPL1だけから文脈CTX0を作り出す事もできる。

RMSが管理するRELから射影によっていろいろな集合をつくり、それらの集合のうち共通の要素を持つもの同士をまとめて互いに排反な集合の族を作る。この知識も、以下のようにして、新知識の構造化における組み合わせ爆発を抑制するために使われる。得られた排反な集合全体の集合を、

$U = \{D1, D2, D3, \dots\}$

とする。各々のREL(2項)に、

$REL \subset D_i \times D_j$

によってペア(D_i, D_j)を対応させる。REL(2項)の間の合成関係を示すTPLの具体例が存在する場合、それに対応してU上の元の間の到達関係が存在する。従って逆に、合成関係発見のためのTPLの探索を、U上での到達関係が存在するRELの組み合わせの中だけに限定できる。

LS/1は、単一種の知識ではなく、いろいろな種類の知識を作り、そして、それらを使い分ける。幹となる知識は、n項RELやそれらの相互関係、1項REL(集合)やそれらの相互関係、RELと集合の関係などであるが、その他に、単語帳や連想リスト、RELやTPLへの重みなど少しレベルの違う知識も作成し利用する。

LS/1の応答生成アルゴリズムは、質問Qもが与えられると、その答えに最もふさわしいと思われる応答文を探索によって求める6)。全ての可能な応答を求めてからそれらの優劣を比較するという全数型探索ではないので、探索量は比較的小さく抑えられている。またループチェックをやって無限ループに陥るのを防いでいる。LNはRDBやCHと結合しているので、探索がRDBやCHと連絡している頂点まで来ると、RMSやCMSに対して問い合わせが発生する。LNでの探索はその問い合わせの答えが帰って来ると再開される。応答生成アルゴリズムは、その他に、CHなどの対象の類似関係の情報を使って、簡単な類推を行なう。これはマッチングの条件を緩和することによって、LNのルール(枝)を拡大解釈するものである。

LS/1は、知識の獲得から知識の利用までの情報処理を、長期または一時記憶を通して連係させている。例えば質問に対する応答の生成のための探索の結果得られた情報(一時記憶)は(最良応答探索アルゴリズム6)を利用して)その応答が正解か否かの判明を待って知識の更新に利用される。また獲得された各種の知識(長期記憶)は応答の生成の際に、またさらなる知識の獲得に利用される(この例についてはすでに述べた)。

8. むすび

人間の日常的な活動を理解するためには、それらの基礎に、知識の構造化、仮説生成、一般化、発見などの情報処理が存在すると仮定しなければならない。そのような視点から見た知的情報処理を総称して帰納的学習と呼ぶ。帰納的学習の実現は演繹的な情報処理とは比較にならないほど困難なので、(それらをうまく実現するために)人間の知的情報処理の概略はそれらの情報処理にとって都合のいい形に決定されている可能性が高いと推測される。従って人間の情報処理全体の理論を目指すためには、それらの情報処理を中心に据えたモデル(帰納的学習システム)を作成することが有効と考えられる。

本論文では、形式化された学習の設定のもとで、置換不変、構成的アプローチによって帰納的学習システム $LS/1$ を構成した。それは未知の言語を用いた質問応答により、知識を獲得し、よりよく応答できるようになる。人間の日常的な活動を可能にする基本的で必須の情報処理としての知識の構造化、仮説生成、一般化、発見などを考察する場合は、 $LS/1$ の枠組みの中で有用な REL を次々に作り出す過程を扱うことによって得られる。この過程は生成される知識に高度の多様性をもたらす可能性があるもので、その制御には特別の対策が必要である。 $LS/1$ はそのために、獲得した知識を利用して知識の構造化の組み合わせ爆発を抑制したり、各種の知識を作成し使い分けるなど、知識の獲得から知識の利用にわたって情報処理の種々の連係を実現している。

紙面の制約のために本論文は、学習の設定（学習状況や置換不変性）や学習の基本原理解や帰納的学習システムの最も特徴的な部分（RELの生成とその自由度への対策）の概要に主眼をおいて記述した。その背後には基礎となる多くの理論（例えば、本論文で与えた学習の設定や置換不変性の必要性の根拠をなす理論、LNの知識表現の方法とその背景となる数学的理論など）やアルゴリズム（各部分の知識獲得と利用のアルゴリズム）が存在する。それらは別の論文で詳細に議論する予定である。

本論文により、帰納的学習システムの理論の構築という長期的課題の重要性が明らかになり、 $LS/1$ によってその課題へ接近する具体的な糸口が与えられた。帰納的学習システムの理論を完成するためにはさらに非常に多くの研究を積み重ねる必要がある。それらの中で、特に $LS/1$ をさらに発展させるための（本論文が記述したような大きなレベルの）問題としては、(1)この枠組みの中への文法の位置付け、(2)関係データベース (RDB) や概念階層 (CH) とは別の構造を扱う知識表現とそれを管理する新しいサブシステムの追加、(3)知識の構造化における誤り回復への本格的な対処、(4)視覚など言語以外の入力への導入などがある。

参考文献

- 1) Akama, K. and Ichikawa, A.: A Basic Model for Learning Systems, Proceedings of 6th IJCAI, P4-5 (1979)
- 2) 赤間清, 市川惇信: 知能のモデル化のためのシステム設計問題 (3) — 知能システムを把える2つの視点, 電子通信学会技術研究報告, PRL79-114, p35-41 (1980)
- 3) 赤間清: 帰納的学習を行なうシステムの初歩的なモデル, HBSR M(S), NO.7, P29 (1984)
- 4) 赤間清: 帰納的学習システム $LS/0$ を実現するプログラムの概要, HBSR M(S), NO.8, P94 (1985)
- 5) 赤間清, 小笠原秀美: 帰納的学習の側面から見た概念の階層構造, 日本認知科学会, 第2回大会発表論文集 B-2 P23-24 (1985)
- 6) 赤間清: 帰納的学習システムの最良応答探索, 情報処理学会, 知識工学と人工知能研究会資料, 41-12, P89-96 (1985)
- 7) 石崎俊, 井佐原均: 文脈情報学習 (同化) のシステム化について, 情報処理学会, 自然言語処理研究会資料, 52-7, P49-56 (1985)
- 8) 北上始, 麻生盛敏, 他: 知識同化機構の一実現法, 情報処理学会, 知識工学と人工知能研究会資料, 30-2, P9-18 (1983)
- 9) 国藤進, 麻生盛敏, 他: PROLOGによる対象知識とメタ知識の融合とその応用, 情報処理学会, 知識工学と人工知能研究会資料, 30-1, P1-8 (1983)
- 10) 高木朗, 芦沢実, 他: 簡単な日本語の意味と文法を学習するシステム, 情報処理学会, 知識工学と人工知能研究会資料, 39-3, P17-24 (1985)
- 11) 弦巻宏治, 国藤進, 他: メタプログラミングによる仮説生成システムの試作について, 日本ソフトウェア科学会, 第2回大会講演論文集
- 12) Poole, D., Aleliunas, R., Randy, G.: Theorist: A Logical Reasoning System for Defaults and Diagnosis, submitted as a chapter in the volume Knowledge Representation, N.J. Cercone & G. McCalla (eds.), IEEE press, in preparation.