

# 人間の推論

## Human Reasoning

鈴木 宏昭  
Suzuki Hiroaki

青山学院大学文学部  
Department of Education, Aoyama Gakuin University  
susan@ri.aoyama.ac.jp

keywords: deduction, induction, analogy.

### 1. はじめに

Fodor は、人間の思考システムを、モジュール構造を持つ知覚や言語などの入力系と対比し、中央系と名付けた。そして、この中央系はありとあらゆる領域の知識、信念が関与するという意味で等方性を持ち、信念の全体的性質が関与するという意味においてクワイン性を持つと主張した [Fodor 83]。Fodor はこれに続けて、こうした性質を持つ中央系（特に類推）を理解することは不可能であるとし、これを Fodor の第一法則と呼んだ。

私は、この主張の前半つまり、思考が等方性とクワイン性を持つと言う主張は、人間の思考の研究にとってきわめて重要な意味を持つと考える。しかし経験科学の立場からすると、後半の第一法則は、穏やかに言えば当時の学問状況に縛られたものであり、忌憚なく言えば間違いである。たしかに、人間の思考は単一のモジュールとして記述できるようなものではなく、多様な認知資源の複雑な相互作用として捉えられるべきである。当時の認知科学の状況は、こうした相互作用の性質を解明・記述には十分な蓄積がなかった。しかし、近年の制約論的アプローチの展開は、これらが確実に認知科学の射程圏内に入ったことを示している（たとえば、本論文の 4 章や [鈴木 01a] をみていただきたい）。また、論理的には等方的であるとしても、人間の通常の思考においては、特定の制限があり、思考が参照する範囲をある程度まで限定することが可能だからであるし、参照する範囲を越えることが極めて稀であれば、それは誤差として取り扱うことが可能だからである。

本論文では、人間の推論の研究を、大きく (1) 演繹、(2) 帰納、(3) 類推、(4) 確率推論、に分け、各々の分野における基本的な考え方や、研究のトピックと経緯を紹介する。人間の思考研究を支える、もう一つの大きな柱である問題解決については紙幅の関係から扱うことができなかったが、これについては [鈴木 01b] をご覧いただきたい。

### 2. 演繹

演繹推論は、心理学の中で、もっとも古くから研究が行われてきた領域である。人間の推論の規範を作ることが、論理学のパイオニアたちの目的であったことを受け、当初の研究は人間の演繹推論がどれほど規範理論に従っているか、もし従っていないとしたらどのような特徴があるのかを、三段論法や条件文推論の課題を用いて検討してきた。以下に、よく知られた人間の演繹推論の特徴を挙げる（詳しい説明は [高橋 96] を参照されたい）。

#### 2.1 推論における人間のバイアス

##### 1 雰囲気効果

三段論法の前提中に「～ではない (not)」や「いくつかの (some)」が用いられてると、結論中にもそれらが現れる。そうでない場合は全称肯定文となる。

##### 2 転換バイアス

三段論法中で「A は B である」が前提として与えられると、「B は A である」という転換した命題も真であると考えられるバイアスである。

##### 3 格バイアス

定言三段論法では、3つの名辞が2つの前提と1つの結論に現れる。この現れ方のパターンを格 (figure) と言う。格によって三段論法の難しさは異なってくる。

##### 4 信念バイアス

人は論理的帰結の中で、自らの信念に一致する結論を選択しやすい。

##### 5 双条件解釈

条件文推論において「もし A ならば B」という命題があり、かつ「B」が成り立っていることが分かると「A」も成り立っていると考え、人間の傾向性を双条件解釈という。

##### 6 確証バイアス

認知科学でよく知られた課題に 4 枚カード問題と呼ばれる問題がある。これは、表にアルファベット、裏に数字が書かれてあるカードを用いて「もし表が母音ならば裏は偶数」という命題の真偽を確かめるというものである。

被験者は「母音」、「子音」、「偶数」、「奇数」の4枚のカードが見せられ、どれをめくって見る必要があるかが問われる。もっとも多い間違いは、「母音」と「偶数」のカードをめくるといものである。これは、論理的には間違いであり、正解は「母音」と「奇数」である。人間のこうした傾向は、確認バイアスと呼ばれる。これは、人間は仮説を確認しようとする方向で専ら推論を行い、仮説を反証するための推論を行わないことを示している。「偶数」を選択する誤りは、もし偶数の裏が母音であれば、なおさらその仮説の妥当性が高まると、人が考えるからである。

このように、人間の演繹推論はさまざまなバイアスを持っており、誤った推論、あるいは危ない推論を行うのが常である<sup>\*1</sup>。1970年代までの研究者の多くは、こうした傾向性を列挙したり、それらの傾向性の確認などを行ってきた。

この現象に対する理論的なアプローチの一つは、心理論理 (mental logic) と呼ばれるものである。この立場の研究者たちは、人間はまったく形式的で、文の意味や内容には無関係な推論図式 (たとえば Modus ponens や、連言の削除など) を持っており、これを適用することで演繹推論を行っている。簡単に言えば、人間は限定付きの prolog のインタプリタということになる。心理論理アプローチからすると、人間の誤りは、ほとんどすべて自然言語を推論図式が適用される論理形式に翻訳する際の誤りに起因させられることになる [Braine 91, Rips 94]。

## 2.2 メンタルモデル

Johnson-Laird は心理論理説と真っ向から対立するメンタルモデル理論を提出した。この理論では、人の推論過程はルールの適用ではなく、メンタルモデル<sup>\*2</sup>の生成として捉えられる [Johnson-Laird 83]。

この理論は次のようなきわめて単純な仮定をおく [Johnson-Laird 93]。

- (1) (問題中の) 対象はトークンとしてメンタルモデル内に表象される。
- (2) 対象の性質はトークンの性質として表象される。
- (3) 対象間の関係はトークン間の関係として表象される。

ここで大事なことはメンタルモデル内のトークン、性質、関係は、我々が頭の中でイメージ可能な具体的な形式で表象されているということである (つまり変数として表象されているわけではない)。そして推論はこうして形成されたメンタルモデルを変更、解釈することにより行われるというのが Johnson-Laird の主張である。

\*1 ただし、これらのバイアスがいつでも人を間違いに導くわけではないことに注意されたい。たとえば「どの A も B ではない」に転換バイアスが作用しても誤りにはならない。

\*2 メンタルモデルという概念は様々な文脈で用いられる。特に、装置などのメンタルモデルのことをデバイスモデルと呼ぶこともある。

彼はこの考えを三段論法を例にとって説明している。メンタルモデル理論に従えば、三段論法は次のようなプロセスで推論が行われるとされる。

- (1) 第一前提についてのメンタルモデルを形成する。
- (2) 第二前提の情報を先のメンタルモデルに付け加える。
- (3) 出来上がったメンタルモデルの全てに共通する関係を抜きだし、それを結論とする。
- (4) 結論に対する反例を探す。

メンタルモデル理論においては、人は既存の知識を用いて、問題中の対象の意味内容を分析し、意味的に一貫したモデルを構成するものとされており、問題文の意味に応じた人間の推論パターンの変化を説明するための枠組みが用意されている。これにしたがえば、人間の推論における誤りは、メンタルモデルの作りやすさ、モデルの数、および複数のモデルに共通した事柄を抜き出す作業に関連している。

上のプロセスの 1, 2 において背景知識、一般的な知識の介入が想定されている。たとえば、数字とアルファベットを用いた典型的な 4 枚カード問題 (偶数の裏は母音でなければならないという問題) と Griggs & Cox によって行われた「ある人がビールを飲んでいれば、その人は 19 歳以上でなければならない」という問題 [Griggs 82] を比較してみよう。典型的な 4 枚カード問題では規則の違反例を想定することが相対的に困難であり、その結果メンタルモデルの中にそうした対象がトークンとして表現される可能性はきわめて低い。一方、後者では規則に違反している仮想的な人物 (19 歳未満なのに酒を飲んでいる人) を容易にトークンとしてモデルの中に含めることができ、規則違反の人間がどの例に該当するのかを容易に推論することができる。また、前提から作り出されるメンタルモデルは複数であることが多い。このため、モデルの数が多いと作業記憶に負荷をかけることになり、成績が低下することが予測できる。実際に、モデルが多くなればなるほど、人間は誤った推論を行いがちになる。

## 2.3 実用的推論スキーマ

人間の演繹推論がおもしろいのは、人間がいつでも誤ったバイアスを適用してしまうわけではない点にある。たとえば、過去に経験のあったことであれば、人は 4 枚カード問題と同様の問題を正しく解決できる [Griggs 82, Johnson-Laird 72]。しかし、実際に経験したことしか解決できないのでは、その推論能力は甚だ弱いと言わざるを得ない。

ところが、Cheng & Holyoak は、問題に「許可」という情報が含まれている限り、人間は自然な推論で正解にたどり着けることを実験的に明らかにした [Cheng 85]。ここで許可情報を含む問題とは、「ある行為を行うためにはその前提条件が満たされていなければならない」とい

う形で表現できる。我々はこうした状況に対処するためのルールを日常生活から獲得している。それらは次のようなものであると考えられる。なお括弧内はいわゆる論理学の規則である。

- (1) もし前提条件が満たされているならば、その行為を行ってもよい。(後件が真ならば前件は真あるいは偽である)
- (2) もし前提条件が満たされていないならば、その行為を行ってはならない。(後件が偽ならば前件は偽である)。
- (3) もしその行為を行うならば、前提条件を満たしていなければならない。(前件が真ならば後件は真である)。
- (4) もしその行為を行わないならば、前提条件を満たす必要はない。(前件が偽ならば後件は真あるいは偽である)。

これらのルールは条件文推論のための規則と一致している。したがって、こうした内容を持つ問題に対しては、仮にその内容が未知のものであったとしても、人間は正しく推論を行うことができる。

Cosmides はこの実用論的推論スキーマの起源について、進化論的な観点から興味深い考察を行っている。人間は集団生活を行う動物である。こうした生き物は、獲物の分配などの援助的利他行動を少なからず行う。この行動は遺伝子淘汰の観点から見て合理的であるが、もし援助を受けるだけで、他者に対しては援助を行わない裏切りものがいた場合には適応的ではない。したがって、集団生活を行う人間には、この裏切りものを検知するためのモジュールが進化的に獲得されているという。つまり許可のスキーマは、こうしたモジュールなのだというのが彼女の主張である。これだけだと、単なるお話の域を出ないが、彼女は社会的交換の文脈を加えた 4 枚カード問題とそれ以外の形式の 4 枚カード問題のパフォーマンスの比較を通して、この仮説を検討している。実用論的推論スキーマ説でも、進化論的な説明でも、通常の許可の文脈の問題についての予測は同じになる。しかし、進化論的な説明では、許可の問題の条件部と結論部を逆にした場合でも逆にする前と同じ選択肢が選ばれることが予測できる。実験を行うと、実際に人間は逆にする前と同じ選択肢が選ぶことが明らかになった。[Cosmides 89]。

## 2.4 確率的アプローチ

最近になって演繹推論への確率論的なアプローチが現れてきた [Oaksford 99]。これは、人が 4 枚カード問題のような条件文推論を行う時には、演繹ではなく、確率推論を行っているという主張である。「 $p$  ならば  $q$ 」という規則が与えられた場合、これが成り立つという仮説  $H_D$  と、成り立たないという仮説  $H_I$  が存在する。被験者となる人間は、どのカードをめくることが、この 2 つの矛盾する仮説からなる世界の曖昧さを減少させることがで

きるのか、という課題を解いている。もし規則が成り立たないとすれば、一般的に  $p$  や  $q$  である確率は低いはずなので(つまり特定の  $p$  や  $q$  であるよりは、その他一般である  $not - p$ ,  $not - q$  の確率が高い)、ベイズの定理を用いた各々の仮説下における可能性の確率を質的に表現すれば表 1 のようになるだろう。さて、どのカードをめくることが不確実性を減少させるかは、この表の  $H_D$  の列と  $H_I$  の列の差の大きさに関係する。この間に大きな違いがあれば、そのカードをめくるとは仮説の識別にとって informative である。この表からわかるように、 $p$  と  $q$  をめくれば、 $H_D$  と  $H_I$  を識別するのに、もっとも informative ということになる。こうしたことからすれば、人間が条件文推論において  $p$  や  $q$  を選択するのは、非合理的と見なすことはできなくなる。人間の非合理性を示すと言われてきた現象は、実験者の想定する課題の性質と、被験者が定式化する課題が異なっているためなのである。

表 1 各場合における確率。ただし、 $H_D$  は規則が成り立っているという仮説、 $H_I$  は成り立っていないという仮説を指す。これらの列の下には、上の仮説のもとで考えられる確率を表す [Oaksford 99]。

表	裏	$H_D$	$H_I$
$p$	$q$	確実	低
	$not - p$	不可能	高
$not - p$	$q$	低	低
	$not - q$	高	高
$q$	$p$	高	低
	$not - p$	低	高
$not - q$	$p$	不可能	低
	$not - p$	確実	高

Oaksford & Chater は、こうした理論の下で、マッチングバイアスや実用論的推論スキーマが関与する条件文推論についての多くのデータを一貫して説明できることを示している [Oaksford 94]。また、彼ら以外にも、こうした確率的なアプローチによって条件文推論の研究を行うものが増えてきている。

## 3. 帰納

認知心理学の中で推論の形式としての帰納を純粋に取り上げた研究はさほど多くはない。ここでは、近年比較的まとまった研究がなされている「カテゴリーに基づく帰納」についての研究に絞って紹介を行う。ただし、概念、カテゴリー研究の多くは帰納的学習を想定したものとなっているので、これらについては本レクチャーシリーズの「概念の構造と処理」[河原 01] を参照されたい。

### 3.1 カテゴリーベースの帰納

帰納推論の中でもカテゴリーベースの帰納推論 (category-based induction) と呼ばれるものがある。これは、前提と結論におけるカテゴリーについての知識をもとにして、前提を一般化、拡張する推論を指す。カテゴリーに基づく一般化には、特殊論証と一般論証がある。たとえば、各々次のようなものである。

#### 特殊論証

まぐろの血液中の鉄分の濃度は人間よりも高い  
 かつおの血液中の鉄分の濃度は人間よりも高い

#### 一般論証

まぐろの血液中の鉄分の濃度は人間よりも高い  
 魚の血液中の鉄分の濃度は人間よりも高い

前提の種類によって、結論の確信度は様々に変化する。たとえば、上に挙げた「まぐろ」を「ひらめ」に変えれば、多くの人ははじめの結論(「まぐろ」)の方がより妥当であると考えらるだろう。Osherson らの研究では 90% 程度の人がこのような判断をしている [Osherson 90]。これは「前提の典型性」効果と呼ばれ、前提中の対象が結論部のカテゴリーにおいて典型的であればあるほど、その推論の確信度は強くなるということを表している。

では、「まぐろ 魚」と「ひらめ 魚」の各々の前提に「かつおの血液中の鉄分の濃度は人間よりも高い」という新たな前提を付け加えたらどうなるだろうか。この場合、おそらく、「まぐろ・かつお 魚」よりも、「ひらめ・かつお 魚」の方がより妥当であると感じられるだろう。Osherson らの実験では「ひらめ・かつお 魚」の方を強いと感じる被験者は逆の被験者の 3 倍にもなった。これは「前提の多様性」効果と呼ばれており、前提中の対象が多様であるほど結論の確信度は上がるというものである。

こうした現象は現在分かっているだけで、13 ほどある。これらの中には、「前提中にカテゴリーが用いられ、結論に含まれる対象がそのカテゴリーのメンバである時は、その論証は完全である」とか「前提が多いほど結論は妥当と感じられる」などという当たり前のことも含まれているが、「前提と結論を逆転すると確信度に差が生じる(論証の非対称性)」とか「結論部の対象を特殊化すると確信度が下がる場合がある」などのそれほど自明でもないことも含まれている。

### 3.2 類似・網羅度モデル

Osherson らは類似・網羅度モデル (similarity-coverage model) が提案した [Osherson 90]。これによれば、帰納

的論証のもっともらしさ(強さ)は次の式で与えられる。  

$$\alpha \text{Sim}(\text{cat}(P_1) \dots \text{cat}(P_n); \text{cat}(C)) +$$

$$(1 - \alpha) \text{Sim}(\text{cat}(P_1), \dots, \text{cat}(P_n); [\text{cat}(P_1),$$

$$\dots, \text{cat}(P_n), \text{cat}(C)])$$

ここで、 $\text{cat}(P_n)$  は前提のカテゴリー、 $\text{cat}(C)$  は結論のカテゴリーである。 $\alpha$  は非負の重みである。また、 $[\text{cat}(p_1 \dots p_n)]$  は  $p_1$  から  $p_n$  を包括するもっとも低レベルのカテゴリーを表す。この式の第一行目は、前提中のカテゴリーと結論中のカテゴリーの類似度を表し、残りの行は、前提中のカテゴリーと、前提と結論中の対象全てを包含するもっとも低レベルのカテゴリーとの類似度の 2 つを表している。ただし、この種の帰納的論証には 3 つ以上の項目を含むものもあるため、類似度計算を若干拡張する必要がある。彼らは  $\text{Sim}(P_1, \dots, P_n; X)$  は  $P_1$  から  $P_n$  までの中でもっとも  $X$  に似ているものの類似度としている。また、網羅度は、前提中のカテゴリーと、前提と結論中の対象全てを包含するもっとも低レベルのカテゴリーの事例を上記の方法で一つずつ類似度を計算し、それを平均したのものとなっている。

このようなモデルから、先に挙げた例を含む 13 の現象に関わる帰納的論証のもっともらしさについての人間の判断をうまくモデル化できることが明らかになった。ただし、被験者の知識の程度、年齢などによっては、このモデルの当てはまりが悪い部分があることも明らかになり、いくつかの修正や洗練が行われている [岩男 99, López 92]。

### 3.3 特徴ベースモデル

一方、Slovan は類似・網羅度モデルのようなカテゴリー構造を仮定せずに、単に前提と結論に現れるカテゴリーが持つ特徴だけから帰納的論証の強さを予測する「特徴ベースモデル (feature-based model)」を提案している [Slovan 93]。このモデルは、前提に現れるカテゴリーの特徴ベクトルを入力として、結論部の述語(「血中の鉄分濃度が人間よりも高い」等々)の活性値を出力とする、単層パーセプトロンと同様のコネクショニストモデルである。このモデルによれば、帰納的論証のもっともらしさの判断は、(1) 前提中のカテゴリーと結論中のカテゴリーの特徴のオーバーラップにより高まり、(2) 結論部のカテゴリーの特徴が多ければ弱まる、ということになる。たとえば、(1) より「まぐろ かつお」の方が、「まぐろ ひらめ」よりももっともらしいと判断され、(2) より、仮にひらめ—まぐろ、ひらめ—かますとの特徴のオーバーラップが等しいとしても、「ひらめ かます」の方が「ひらめ まぐろ」よりももっともらしいと判断される。

類似・網羅度モデルでは、ある意味でカテゴリー間の類似性は固定されたものとして扱われていた。しかし、特徴ベースモデルでは、カテゴリーを構成する特徴をベースにしているため、帰納推論を行う状況に合わせて動的に類似性を計算することができる。また特徴ベースモデル

は類似・網羅度モデルよりもシンプルであるだけでなく、結論が前提に含まれるような特殊な場合における帰納的論証について、人間の判断とよりよい適合性を見せる。

## 4. 類 推

類推は類似した過去の経験=ベースをもとにして、現在の問題=ターゲットを理解、説明、解決する推論の形式を指す。類推のプロセスは、ターゲット表象の生成、ベースの検索、写像、正当化/適合、学習のサブプロセスからなるとされている。類推は、類似によって知識を拡張していく推論と見なすことができる。こうしたことから類推は人間に見られる柔軟な思考を解明するための鍵と考えられ、1980年代から現在に至るまで、思考の認知研究の中で重要な位置を占めてきた。

ただし、類推には計算論的に見てさまざまな問題が含まれている。たとえば、類推において用いられるベースは、ターゲットとは別領域であることもある。したがって類推におけるベースの検索は、ルールベースのモデルのようにタスクに固有なルール群の中から適当なものを選択するというほど簡単なものではない。また、写像も計算論的に見てきわめて難しい問題を含んでいる。たとえば、仮にベースの要素が  $n$  個あれば、写像可能な候補は  $2^n$  だけ存在する。また、写像する要素が決まりそれが  $m$  個であったとすると、写像の仕方は  $m!$  だけ存在する。一般に  $m, n$  は数個ということはないので、何の制約もなしに写像を行うことは不可能である。こうした困難にも関わらず、類推は日常的であり [Lakoff 80]、また幼児にすら可能である [Goswami 92]。このなぞを解くために、さまざまな研究が行われた。

### 4.1 構造写像理論

認知科学の類推研究をその初期からリードしてきたのは、Gentner とそのグループである。彼女が提案した構造写像理論は、写像の過程で生じる膨大な数の無意味な仮説を排除するために、構造という概念を持ち出す。この理論に従えば、妥当な類推写像は以下の3つの基準を満たすものである [Gentner 83]。

1. 属性の非写像: 対象の属性は写像されない。
2. 構造的な一貫性: 構造的に整合しているものが一対一に写像される。ベース中の対象は、ターゲット中の対象に一対一対応する。また、関係は項の数が等しいもの同士が一対一に対応づけられる。また対応づけられた関係が束縛する対象同士も必ず対応づけられる (parallel connectivity)。
3. システム性原理: 高階の関係が優先的に写像される。また一階の関係の中でも高次の関係の引数であるものが写像される (systematicity principle)。

この理論は、類推は個別の対象の類似に基づいているのではなく、構造 (=関係のシステム) における類似、あ

るいは一致に基づいていることを主張する。写像が構造に基づくという定式化により、類推の過程で生じる不要な仮説を大幅に削減することができる。もう一つの重要な点は、構造は主観的なものでも、また内容に依存したものでなく、ベースとターゲットの表象の統語的、形式的な性質によって決定されるということである。

Gentner らはこの理論を支持する証拠をさまざまな実験や分析を通して挙げている。たとえば、同様に重要な事柄であっても、写像の候補とされるのは高階の関係に支配された事柄であること [Clement 91]、類推の場面になると構造的な一致に基づく判断が働くこと [Gentner 93]、構造の認識が発達によって異なること [Gentner 86] などである。

また、Falkenhainer らは構造写像理論を用いた類推のコンピュータモデルを構築している [Falkenhainer 89]。このモデルでは、まず主に上記の1の基準を用いて暫定的な対応仮説を生成する。次に、一対一対応や parallel connectivity の基準を用いて、Gmap と呼ばれる一貫した対応仮説の集合が作り出される。次の段階では、システム性原理により、ベース中の要素がターゲットに写像されることになる。生成される Gmap は一つとは限らないので、最後の段階では対応の良さを、構造の深さ、対応する要素の数、写像候補の数などによって評価する。このモデルの計算結果は、Gentner らが行った実験の結果と一致していることが明らかにされている。

### 4.2 多重制約理論

Gentner は類推における類似を構造的な側面に限定して理論を組み立てた。しかし、人間の類推に関するものは構造だけではない。たとえば、対象の構造は単一ではなく、同じものであっても、類推を行う目的に応じて異なった構造が現れる。こうしたことから、類推にはプラグマティックな類似性も関与している [Holyoak 87, Spellman 96]。

また、人間は未知のターゲットを理解するために類推を行うのであるから、ターゲットの構造はほとんど分からない場合も少なくない。このような場合は、人間は属性や見かけの一致などに基づく類似性を頼りに類推を行う [Ross 87]。むしろ、これらのレベルの類似性は類推の妥当性を保証するものではないが、現実世界では非常によく似た2つのものは同じ構造を持つことも多い [Medin 89]。

こうしたことから、構造に加えて、プラグマティックな類似性や属性レベルの類似性を制約として組み込んだ多重制約理論が Holyoak と Thagard らのグループによって提案された [Holyoak 95]。この理論にしたがい、ベース検索と写像の各々についてコンピュータモデルが構築されている [Holyoak 89, Thagard 90]。2つのモデルはともにベースとターゲットの要素間の対応をノードとする、ローカリスト・コネクションネットワークによっ

て実現されている。このネットワークは構造的な制約に従い、要素間の対応関係を作り出すとともに、それらの間に興奮性、抑制性のリンクを張る。たとえば、ベースの中のある関係  $rel_1^b(x, y)$  がターゲットの関係  $rel_1^t(p, q)$  と対応づくとする、 $rel_1^b$  と  $rel_1^t$  の対応を表すノード、及び  $x$  と  $p$ 、 $y$  と  $q$  の対応を表すノードが作られる。そして、関係の対応ノードと対象の対応ノードの間には興奮性のリンクが張られることになる。また、対立する対応仮説を表すノード（たとえば  $rel_1^b$  と  $rel_2^t$  の対応仮説ノード）からは抑制性のリンクが張られる。一方、プラグマティックな類似性や属性レベルの類似性については、特別なノードを用意し、これらの類似性を満たしているノードに常に強い興奮性の信号が送られることになる。こうしたネットワークに対して、一般的な緩和アルゴリズムを適用することで、最終的にベースの候補や写像の候補が見つげ出されることになる。

彼らのモデルも、心理実験から得られたデータをかなりうまく説明できることが示されている。また、ベース検索において、候補となるベースの数が大きくなっても、収束までに必要なサイクルには変化がないことが示されている。これらのモデルでは、ローカリストの表現を用いているため、ベースやターゲットの構成要素が一つのノードとされていたが、はノードの活性パターンの同期関係に基づく dynamic binding を採用した、分散表現の類推モデルも提案されている [Hummel 97]。

#### 4.3 抽象化に基づく類推

類推には他の制約も関与している可能性がある。特に、抽象化 (abstraction) が類推に関与しているという主張は、人工知能研究の中では比較的ポピュラーなものである [Russell 89, 桜井 93]。実際に、抽象化は類推の成否に決定的な重要性を持っていることは、心理学の研究でも明らかにされている。たとえば、解決がきわめて困難な洞察問題である放射線問題を類推的に解決させる際に、事前に複数のベースを与え、その共通点を挙げさせると、この問題の解決が相当程度促進される [Gick 83]。これは複数のベースが与えられたことにより、これらを抽象化した知識が獲得されたためであると考えられている。

ただし、人間は一般に抽象的な知識を利用することは得意ではない。むしろコンピュータにおいても同様であるが、これは利用の際のコストが高いためである。知識が抽象的になればなるほど、現実の具体的な情報とマッチをとることが難しくなるためである。そこで人間はどのような抽象化を行い、利用しているのかが問題となる。鈴木は類推研究の成果に加えて、概念研究、問題解決、転移の研究の成果から、人間の用いる抽象化は、

- 一般化された目標の達成に向けたものになっている、
- 抽象化内の対象や関係はその目標の達成という観点から、意味的、機能的に結びついている、
- またそこに関与する対象は目的を実行するための条

件を満たしている、という性質を持っていることが指摘し、これらの性質を持つ抽象化を準抽象化 (quasi-abstraction) と呼んでいる [鈴木 96]。ここでは、任意の抽象化ではないという意味で「準」という限定が付き、一般化された目標の達成に関わるものであるという意味において「抽象化」という用語が用いられる。また、理論的な要請ではないが、準抽象化は小さな知識単位であり、その構造は単純である。こうした特徴を持つ抽象化が関与することにより、人間の類推においては検索や写像がきわめて迅速に行われることが示されている [Suzuki 98]。

#### 4.4 構造の生成: Copycat

構造写像理論においてはもちろん、多重制約理論においても構造の制約は中心的な制約である。しかし、どちらの理論でも構造は与えられたものであった。これは、かなり問題を含む仮定と言えよう。実際の類推においては、同じベースであっても、類推を行う場面に依りて異なる構造を見せることがあるだろうし、類推の過程で、今まで気づかなかった構造が産み出されることもあるだろう。Hofstadter と Mitchell が中心となって行われた Copycat と呼ばれるプロジェクトでは、文字列の 4 項類推を課題として、この問題にアプローチしている [Hofstadter 95, Mitchell 93]。たとえば、 $abc \quad abd$  ならば  $ijk \quad ?$  という形の問題が与えられる。この場合は、多くの人は  $ijl$  という解を出す、人によっては  $ijd$  という解を出す場合もある。また、もしターゲットとして  $xyz$  が与えられると、ある場合には  $xya$  が解とされることもあるし、 $wyz$  が解とされることもある。

このように 1 つの問題から様々な解が導き出される流動的で、柔軟な類推を実現するために、Copycat では階層的な概念構造を用いてトップダウンな処理を行いつつも、多数の局所的でボトムアップなプロセスが並列に動作する、マルチエージェント的なコンピュータモデルを提案している。Copycat は、Slipnet、Codelet、Workspace という 3 つの構成要素を持つ。Slipnet は、この課題の表象を生成する時に用いられる様々な概念の貯蔵庫である。これが通常概念ネットワークと異なるのは、滑り (slip) と呼ばれる仕組みが組み込まれており、ノードの活性値に応じてノード間の距離 (活性値の伝播の度合) が可変になっているという点である。Codelet は、課題達成のためのオペレータであり、Slipnet 内の関連するノードの活性値に応じて、非決定的に選択される。Copycat では様々な Codelet が並列的に作用し、異なる構造を作り出し、競合することが認められている。Workspace はこれらの Codelet が作用し、表象が生成される場所である。ここでは作り出された表象がどの程度構造的に一貫しているかが評価される。この評価の値によって、システムの温度が変化し、それによって Codelet の適用のランダム性が変化する。Hofstadter らはこの理論によって、

文字列四項類推課題における人間のパフォーマンスをその分散まで含めた形で説明可能であることを示した。

#### 4.5 生態学的アプローチ

さて従来の類推研究の中で繰り返し強調されてきたのは、人の類推（特にベースの検索）は表層に惑わされやすく、なかなか自発的に適切な類推を行うことはできないというものであった [Gick 80, Gentner 93, Ross 89]。しかし、こうした知見は、現実的な場面においても成り立つものなのだろうか。

Dunbar は実験室場面を離れたより日常的な状況における類推の研究を行った [Dunbar 95]。彼は、分子生物学で国際的に評価の高い4つの研究室に一年間密着し、ミーティングの記録やインタビューを通して、現実の科学的な発見の過程についての on-line な分析を行った。すると、科学者たちは仮説の生成、実験の計画、及びその修正など、さまざまな場面において類推を自発的にかつ頻繁に用いることが明らかになった。また、説明のような場面を除けば、用いられる類推は同じ領域に属するベースを用いたものであり、構造的にも表層的にも類似したベースが用いられていた。また、彼らはケベック州の独立について新聞に掲載された400の記事を分析したところ、そのうちの200以上もの記事に類推が用いられていたことがわかった。そして、それらの類推の3/4は政治や経済以外のいわゆる領域をまたぐ類推 (cross-domain analogy) であること、そして大半が構造に基づくものであることが明らかになった [Dunbar 01]。

これらの結果は、人間は日常的な状況では、従来考えられていたよりもはるかに類推をうまく用いていること、かつ類推の大半は表層レベルの一致だけではなく、構造に即したものであることを示している。実験室状況と異なり、領域についての豊富な知識が利用可能で、類推を行うゴールや必然性がある状況下では、人間は適切な類推を行うことができるのである。Dunbar らはさらに、被験者に実際に類推を生成させるというゴールを与えれば、実験室的な状況下でも、生成される類推の大半は構造に準拠したものとなることが示した [Dunbar 01]。

こうした知見は、類推研究にとって単なるデモンストラーションを越えた意味を持つ。というも、構造写像理論、多重制約理論とも、非常にかげはなれた領域間の類推や、既有的知識が最小限でも可能な類推を説明するために、実証的に受け入れがたい仮定をモデルの中に持ち込んでいることがあるからである [鈴木 96]。今後、Dunbar らのような生態学的なアプローチに基づく研究が蓄積され、より現実的な妥当性の高いモデルの構築が望まれる。

## 5. おわりに

本論文では、認知心理学における演繹、帰納、類推研究の基本的な考え方と近年の動向を紹介した。ここで触れ

ることのできなかった話題は多いので一つずつあげていくわけにはいかないが、特に確率推論については一言触れておきたい。この分野では [Gilovich 91, 市川 97, 印南 97] など、日本語で読める優れた書籍が出版されている。また、近年 Gigerenzer を中心した Max Plunck 研究所の ABC (Adaptive Behavior and Cognition) Research Group が、不確定状況下の人間の推論について生態学的、進化論的アプローチを用いて、興味深い結果を導き出している [Gigerenzer 99, Gigerenzer 00]。

人間の思考は Fodor が述べるように、等方性とクワイン性を持つきわめて複雑なシステムとなっている。しかし、本論文で述べてきたように、そこには一定の傾向性、法則性が存在しており、科学的な探求の対象とすることは十分に可能である。ただし、本当にトータルに人間の思考を解明するためには、認知心理学の用いる実験的手法だけでは不完全であることは明白である。モデル化から始まり、インプリメンテーション、検証に至る、人工知能コミュニティーが持つ設計的、トップダウンなアプローチは不可欠である。私見かもしれないが、近年、人工知能における思考や学習の研究者の関心は Web やロボットなどの現実的な問題の方向に向かい、人間知能の理論的研究との接点は減ってきているように思われる。しかし、これは何も悲しむべき事態ではないだろう。そこからの知見でさらに強固で、洗練されたアプローチを獲得した人工知能研究者との新たな出会いが（近い将来）あるはずだから。

## 謝 辞

本論文の執筆に当たってメディア教育開発センターの大西仁氏から有益な助言をいただいた。

## ◇ 参 考 文 献 ◇

- [Braine 91] Braine, M. D. S. and O'Brien, D. P.: A theory of if: A lexical entry, reasoning program, and pragmatic principles, *Psychological Review*, Vol. 98, pp. 182 – 203 (1991).
- [Cheng 85] Cheng, P. W. and Holyoak, K. J.: Pragmatic reasoning schemas, *Cognitive Psychology*, Vol. 17, pp. 391 – 416 (1985).
- [Clement 91] Clement, C. A. and Gentner, D.: Systematicity as a selection constraint in analogical mapping, *Cognitive Science*, Vol. 15, pp. 89 – 132 (1991).
- [Cosmides 89] Cosmides, L.: The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task, *Cognition*, Vol. 31, pp. 187 – 276 (1989).
- [Dunbar 95] Dunbar, K.: How scientists really reason: Scientific reasoning in real-world laboratories, in Sternberg, R. J. and Davidson, J. E. eds., *The Nature of Insight*, MIT Press, Cambridge, MA (1995).
- [Dunbar 01] Dunbar, K. and Blanchette, I.: The in vivo / in vitro approach to cognition: the case of analogy, *Trends in Cognitive Science*, Vol. 5, pp. 334 – 339 (2001).
- [Falkenhainer 89] Falkenhainer, B., Forbus, K. D., and Gentner, D.: Structure mapping engine: Algorithm and examples, *Artificial Intelligence*, Vol. 41, pp. 1 – 63 (1989).
- [Fodor 83] Fodor, J. A.: *The Modularity of Mind: An Essay*

- on Faculty Psychology, The MIT Press, Cambridge, MA (1983), 伊藤笏康・信原幸弘 (訳) 『精神のモジュール形式』・産業図書 (1985).
- [Gentner 83] Gentner, D.: Structure-mapping: Theoretical framework for analogy, *Cognitive Science*, Vol. 7, No. 2, pp. 155 – 170 (1983).
- [Gentner 86] Gentner, D. and Toupin, C.: Systematicity and surface similarity in the development of analogy, *Cognitive Science*, Vol. 10, pp. 277 – 300 (1986).
- [Gentner 93] Gentner, D., Rattermann, M. J., and Forbus, K. D.: The roles of similarity in transfer: Separating retrievability from inferential soundness, *Cognitive Psychology*, Vol. 25, pp. 524 – 575 (1993).
- [Gick 80] Gick, M. L. and Holyoak, K. J.: Analogical problem solving, *Cognitive Psychology*, Vol. 12, pp. 306 – 355 (1980).
- [Gick 83] Gick, M. L. and Holyoak, K. J.: Schema induction and analogical transfer, *Cognitive Psychology*, Vol. 14, pp. 1 – 38 (1983).
- [Gigerenzer 99] Gigerenzer, G., Todd, P. M., and ABC Research Group, the: *Simple Heuristics that Makes us Smart*, Oxford University Press, New York (1999).
- [Gigerenzer 00] Gigerenzer, G.: *Adaptive Thinking*, Oxford University Press, New York (2000).
- [Gilovich 91] Gilovich, T.: *How We Know What Isn't So*, The Free Press (1991), 守一雄・守秀子 (訳) (1993) 「人間この信じやすきもの: 迷信・誤信はどうしてうまれるか」新曜社.
- [Goswami 92] Goswami, U.: *Analogical Reasoning in Children*, Lawrence Erlbaum Associates, Hillsdale, NJ (1992).
- [Griggs 82] Griggs, R. A. and Cox, J. R.: The elusive thematic-materials effect in Wason's selection task, *British Journal of Psychology*, Vol. 73, pp. 407 – 420 (1982).
- [Hofstadter 95] Hofstadter, D.: *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*, Basic Books, New York (1995).
- [Holyoak 87] Holyoak, K. J. and Koh, K.: Surface and structural similarity in analogical transfer, *Memory & Cognition*, Vol. 15, pp. 332 – 340 (1987).
- [Holyoak 89] Holyoak, K. J. and Thagard, P.: Analogical mapping by constraint satisfaction, *Cognitive Science*, Vol. 13, pp. 295 – 355 (1989).
- [Holyoak 95] Holyoak, K. J. and Thagard, P.: *Mental Leaps: Analogy in Creative Thought*, The MIT Press, Cambridge, MA (1995), 鈴木宏昭・河原哲雄 監訳 (1998). 『アナロジーの力: 認知科学の新しい探求』・新曜社.
- [Hummel 97] Hummel, J. E. and Holyoak, K. J.: Distributed representations of structure: A theory of analogical access and mapping, *Psychological Review*, Vol. 104, pp. 427 – 466 (1997).
- [市川 97] 市川伸一: 考えることの科学, 中公新書 (1997).
- [印南 97] 印南一路: すぐれた意思決定: 判断と選択の心理学, 中央公論社 (1997).
- [岩男 99] 岩男卓実: カテゴリに基づく帰納推論における専門性の影響, *教育心理学研究*, Vol. 47, pp. 59 – 67 (1999).
- [Johnson-Laird 72] Johnson-Laird, P. N., Legrenzi, P., and Legrenzi, S. M.: Reasoning and a sense of reality, *British Journal of Psychology*, Vol. 63, pp. 395 – 400 (1972).
- [Johnson-Laird 83] Johnson-Laird, P. N.: *Mental Models*, Cambridge University Press, Cambridge, MA (1983), (海保博之 (監訳) (1988) 『メンタルモデル』産業図書).
- [Johnson-Laird 93] Johnson-Laird, P. N.: *Human and Machine Thinking*, Lawrence Erlbaum Associates, Hillsdale, NJ (1993).
- [河原 01] 河原哲雄: 概念の構造と処理, *人工知能学会誌*, Vol. 16, pp. 435 – 440 (2001).
- [Lakoff 80] Lakoff, G. and Johnson, M.: *Metaphors We Live by*, The University of Chicago Press, Chicago, IL (1980).
- [López 92] López, A., Gelman, A. S., and Guthrie, E. E., G. 19and Smith: The development of category-based induction, *Child Development*, Vol. 63, pp. 1070 – 1090 (1992).
- [Medin 89] Medin, D. L. and Ortony, A.: Comments on Part I: Psychological essentialism, in Vosniadou, S. and Ortony, A. eds., *Similarity and Analogical Reasoning*, Cambridge University Press, Cambridge, MA (1989).
- [Mitchell 93] Mitchell, M.: *Analogy-Making as Perception: A Computer Model*, MIT Press, Cambridge, MA (1993).
- [Oaksford 94] Oaksford, M. and Chater, N.: A rational analysis of the selection task as optimal data selection, *Psychological Review*, Vol. 101, pp. 608 – 631 (1994).
- [Oaksford 99] Oaksford, M., Chater, N., and Grainger, B.: Probabilistic effect in data selection, *Thinking and Reasoning*, Vol. 5, pp. 193 – 243 (1999).
- [Osherson 90] Osherson, D. N., Smith, E. E., Wilkie, O., López, A., and Shafir, E.: Category-based induction, *Psychological Review*, Vol. 97, pp. 185–200 (1990).
- [Rips 94] Rips, L. J.: *The Psychology of Proof*, MIT Press, Cambridge, MA (1994).
- [Ross 87] Ross, B. H.: This is like that: The use of earlier problems and the separation of similarity effects, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 13, pp. 629 – 639 (1987).
- [Ross 89] Ross, B. H.: Reminders in learning and instruction, in Vosniadou, S. and Ortony, A. eds., *Similarity and Analogical Reasoning*, pp. 438–469, Cambridge University Press, Cambridge, MA (1989).
- [Russell 89] Russell, S. J.: *The Use of Knowledge in Analogy and Induction*, Pitman Publishing, London (1989).
- [桜井 93] 桜井, 脇園, 原尾: 抽象化に基づく類推, *情報処理*, Vol. 34, pp. 558–565 (1993).
- [Sloman 93] Sloman, S. A.: Feature-based induction, *Cognitive Psychology*, Vol. 25, pp. 231 – 380 (1993).
- [Spellman 96] Spellman, B. A. and Holyoak, K. J.: Pragmatics in analogical mapping, *Cognitive Psychology*, Vol. 19, pp. 307 – 346 (1996).
- [鈴木 96] 鈴木宏昭: 類似と思考, 認知科学モノグラフシリーズ 1, 共立出版 (1996).
- [Suzuki 98] Suzuki, H.: Justification of analogy by abstraction, in Holyoak, K. J., Gentner, D., and Kokinov, B. eds., *Advances in Analogy Research: Integration of Theory and Data from Cognitive, Computational, and Neural Sciences* (1998).
- [鈴木 01a] 鈴木宏昭: 思考と相互作用, 認知科学の新展開 2 「コミュニケーションと思考」, 岩波書店 (2001).
- [鈴木 01b] 鈴木宏昭: 思考のダイナミックな性質の解明に向けて, 認知科学, Vol. 8, (2001).
- [高橋 96] 高橋, 服部: 演繹推論, 市川伸一 (編), 認知心理学 4 思考, 東京大学出版会 (1996).
- [Thagard 90] Thagard, P., Holyoak, K. J., Nelson, G., and Gochfeld, D.: Analog retrieval by constraint satisfaction, *Artificial Intelligence*, Vol. 46, pp. 259–310 (1990).

{ 担当委員: × × }

19YY 年 MM 月 DD 日 受理

## 著者紹介

鈴木 宏昭 (正会員)

1988 年東京大学大学院教育学研究科博士課程単位取得退学 (学校教育学専攻) . 博士 (教育学) . 現在, 青山学院大学文学部助教授 . 類推, 類似判断, 問題解決, インタフェースなど, 高次認知機能の研究に従事する . 著書に「類似と思考」(共立出版)「アナロジーの力」(監訳, 新曜社)「コミュニケーションと思考」(共著, 岩波書店)がある .