

# 認知科学

—心の哲学へ至る潮流—

石川 幹人\*

## 1 認知科学とは

認知科学 (Cognitive Science) とは、「心とは何か、心はどのように働くのか」という疑問を追求する学問分野のひとつであり、1970年代から学問分野としてのアイデンティティを確立し始めた比較的若い学問である。歴史的には心理学、情報科学、神経生理学、言語学、人類学などの諸学問の学際領域から発展してきたものであり、その萌芽は1930年代にまで遡ることができる。

認知科学が基盤とする方法論は「モデルによる理解」である。我々人間が行い得る「知的行為」が、これこれの心的表象をこれこれの形式で計算操作すると、そうした行動が生み出されると説明できる「認知モデル」を構築することが、認知科学の中心課題である。このモデルの妥当性は、心理学的な実験との整合性、神経生理学的な知見との整合性、モデルに基づいて構成される情報システムの実効性の観点から評価される。

しかし「心とは何か」という疑問は、哲学が永らく相手にしてきた問題であり、認知科学がモデル理解という方法論を確立したからといって、そうした問題ににわかに決着がつく見通しが立ったわけではない。問題点を整理する有力な観点が提供されたのであり、むしろ問題の根深さはより顕

---

\*いしかわ・まさと / 文学部助教授 / 知能情報学

わになったとも言えよう。

本稿では、諸学問から認知科学が生まれ発展してきた歴史を、三領域からそれぞれ簡単に振り返ることにする。そして最後に、哲学の分野との最近の結びつき、すなわち、現代の「心の哲学」を語るうえにおいて認知科学の研究活動がいかなる位置を占めているか、を解説する。

## 2 情報科学からの展開

認知科学の成立には、情報科学の計算にかかわる理論と、それを実行する機械であるコンピュータの技術が不可欠であった。この分野の発展は、20世紀初頭の記号論理学に基礎をおいている。記号論理学から、論理体系で世界を明瞭に記述するという発想が生まれたからである。この発想に加えて、論理規則による演繹を機械仕掛けで普遍的に実現できることが、チューリングマシン (A.M.Turing 1936) の概念で示され、「考える機械」の研究の端緒となった。

40年代になると、情報システムの制御理論であるサイバネティクス (N.Wiener 1948) や、デジタル信号の符号化と通信理論 (C.E.Shannon 1948) が発表された。同時に実用的なコンピュータが開発され、その後、徐々に性能をあげていった (J.von Neumann 1958)。

「心とは何か」という疑問は、コンピュータ技術の発展の前では、「機械は心をもつか」あるいは「人間は機械なのか」という形で現れた。機械が心をもつ判定基準にチューリングテスト (A.M.Turing 1950) が提案され、それを目標に、いわゆる「人工知能 Artificial Intelligence」の研究がスタートするのである。「人工知能」という名称自体は、1956年のダートマス会議において命名された。

人工知能の初期の成功には、論理命題を証明するプログラム (A.Newell and H.A.Simon 1956) や、対連合学習システム (E.A.Figenbaum and H.A.Simon 1962) があげられる。その後、複雑な知識の表現方法やその処理方法として、関連語を有向グラフで結ぶ意味ネットワーク (M.R.Quillian 1968)、推論式を基本にした一般問題解決器 (A.Newell and H.A.Simon 1972)、概念階層の枠組み記述 (M.Minsky 1975) などが提案された。一

方で、状態推移に関する記述の研究 (J.McCarthy and P.J.Hayse 1969) のなかからは、計算のうえでの量的な問題としてフレーム問題が指摘された。

こうした基本技術は、80年代に入ると、実用的なエキスパートシステムの開発へとつながっていった (N.J.Nilsson 1980)。膨大な知識をコンピュータに蓄えようという巨大プロジェクトも生まれた (D.B.Lenat 1985)。しかし、先のフレーム問題を始めとした論理的な記号表現の問題点が表面化し、一転して停滞への歴史を歩むのである。ここで、記号表現に代わって、神経回路を模した結合表現が脚光を浴びることとなる。この原理は、コネクショニストモデルあるいは並列分散処理モデル (D.E.Rumelhart and J.L.McClelland 1986) と呼ばれた。エキスパートシステムのなかには、記号表現に加えて結合表現を採用するハイブリッドシステムも多く現れたが、それも停滞の救世主とはならなかった。

90年代を迎えると、膨大な情報を蓄えたうえでよく考えて解を出す、という設計思想を根本的に考え直す動きが現れた。サブサンクションアーキテクチャ (R.A.Brooks 1990) がそれであり、外界の部分的な情報に即応するモジュールの連合体として知的行動を実現する思想である。こうした潮流から、人工生命 (C.Langton 1989)、エージェントシステム (J.C.Brustolini 1991)、といった研究も発展してきている。

### 3 心理学からの展開

心理学の分野では、1920年頃から、内的な心理状態を除外して、もっぱら動物の反応と条件付けを研究対象とする行動主義が主流となっていた。行動主義は30年以上にわたり心理学研究を支配したが、50年代に研究方法の革命的転換が起きるのである。「認知革命」として知られるこの転換を後押しした研究には、短期記憶のチャンク構造の研究 (G.A.Miller 1956)、概念形成にかかわる認知過程の研究 (J.S.Bruner 1956)、言語の文法構造の情報表現 (N.Chomsky 1957)、認知における注意の役割を示すフィルター理論 (D.E.Broadbent 1958) などがあげられる。

認知革命後の心理学は、心的状態を表すモデルを積極的に認め、そのモ

デルから知的な行動がいかに機能的に説明できるかを問題にした。ここではすでに認知科学の方法論が確立したといえるが、当時は「認知心理学 Cognitive Psychology」(U.Nisser 1967)と呼ばれていた。もちろん、これらのモデル構築には、次々と提案される情報科学の先端理論が影響を与える。それによって、心理学分野ではそれまで十分に研究されてこなかった部分の、知覚・記憶・思考・言語研究が大きく開花した(P.H.Lindsay and D.A.Norman 1977)。

認知心理学の興味を中心を一口で言うと、言語行為を代表とする人間の知的活動が、人間が記憶する知識体系からいかなる思考プロセスで生じるか、となろう。ここで重要な位置を占めるのは知識の記述表現であり、場面に応じた知識活用法の研究(R.C.Schank 1975)、概念知識の構造カテゴリーの研究(E.Rosch 1978)などが相次いだ。視覚イメージの心的回転実験(R.Shepard 1971)からは、心の中のイメージ情報は、アナログ表現(S.Kosslyn 1980)かそれとも離散的な記号表現(Z.W.Pylyshyn 1984)か、という論争が巻き起こった。

70年代後半から認知心理学は、大脳の神経生理学との関連性を徐々に深めていき、認知心理学から認知科学と呼ばれることも増えてきた。1977年に学術誌「認知科学」が発刊され、1979年にはアメリカで認知科学会 The Cognitive Science Societyが発足した。

80年代は、究極の知識表現や単一の中心判断機構といった考え方よりも、多様な記述表現や分散的な判断機構といった考え方が優勢となってきた。人間の記憶は、手続き記憶、エピソード記憶などの多種の機能の複合体として類別された(E.Tulving 1983)。不確実な状況での判断の研究(A.Tversky and D.Kahneman 1982)からは、人間のもつ多様な思考方略が明らかにされた。そして知的な行動とは、状況についてのメンタルモデル(P.N.Johnson-Laird 1983)をもつことから実現されるとみなされた。脳の中には部分的な機能を担うエージェントモジュールが多数存在し、全体として「心の社会」(Minsky 1986)を形成しているのだという主張も、広く受け入れられた。

また、霊長類サルを使った比較認知科学の研究も多くなされ(T.Matsuzawa 1991)、霊長類の認知機構と人間のそれとの類似性・連続性も指摘される

ところとなった。こうした研究は90年代に入り、心の働きを生物進化の文脈に位置づける研究分野の成立につながる。この分野は「進化心理学 Evolutionary Psychology」(L.Cosmides and J.Tooby 1992)と呼ばれ、ある特定の認知方略が確立された根拠を、その方略の進化上の優位性から説明するのである。

## 4 神経生理学からの展開

神経生理学は、人間の知的活動の源である脳の解明・理解という目標を目指して発展してきた。40年代に、単一の神経細胞の数理モデル(W.S.McCulloch and W.H.Pitts 1947)と、神経細胞相互の学習則(D.O.Hebb 1949)が提案されたのが、この分野の発端と言えよう。その後、神経細胞の信号伝達モデル(A.L.Hodgkin and A.F.Huxley 1952)、細胞間のシナプス結合の可塑性や興奮・抑制結合(J.C.Eccles 1957)などの重要な発見が続いた。

脳への展開としては、神経細胞の集団によって判別学習装置を実現するパーセプトロン(F.Rosenblatt 1959)の発見があげられよう。これと同等の仕組みが後に小脳で見つかることとなる(M.Ito 1984)。また、特殊環境で生育したネコの視覚一次野から神経細胞の反応選択性(D.H.Hubel and T.N.Wiesel 1962)が、脳梁切断の症例から大脳半球の機能差(R.W.Sperry 1966)が判明した。

70年代に入ると脳の研究がいっそう盛んになった。大脳辺縁系における海馬では、信号の長期増強(T.V.Bliss and T.Lemo 1973)が見つかり、心理学的な記憶研究と接続した。また、脳外科手術の多くの症例にもとづき、大脳の部位と心的機能との関係がまとめられた(W.Penfield 1975)。視覚系においては、計算論的モデル(D.Marr 1982)が提案されるなか、色判別の神経回路網の生理学的解明(S.Zeki 1983)が進んだ。

80年代後半は、神経回路網の理論が進み、先のパーセプトロンが任意の判別関数の学習まで拡張(D.E.Rumelhart and J.L.McClelland 1986)され、コネクショニストモデルとして情報科学と密に結合することとなった。同時に、PET、F-MRI、SQUIDなど、人間の脳の活動を非侵襲的

に測定する技術が進み、心的活動における大脳部位の活性化の状況が時間を追って測定されるようになってきた。こうした研究動向から、大脳全体として思考がどのように実現されているかという情報モデル (B.J.Baars 1988) も提案され始めた。

90年代以降になってようやく「心がいかにして脳から生まれるか」という、神経生理学本来の目標が射程に入ってきたようである。だが、神経回路の集合体でどのように心的機能が実現できるか、あるいはそれが進化の過程でいかにして自然に発生してきたのか、という問題は依然として難問である。細胞群の同期発火 (F.Crick 1994) やカオス的挙動 (W.J.Freeman 1994) にその手がかりを求める試みはあるが、複雑系物理学 (M.Gell-Mann 1994) や量子物理学 (M.Jibu and K.Yasue 1995) のさらなる発展が必要に思われる。

## 5 哲学への展開

認知科学とその発展をもたらした諸科学は、哲学との接点を歴史上いくつももってきた。それらには、助け合う関係もあれば反目し合う関係もあるが、終始刺激的なものであった。その刺激の度合いは最高潮に達しながら 21 世紀にもち込まれようとしている。

20 世紀初頭にウィーンから発した論理実証主義の運動は、20 世紀前半における科学の哲学的基盤を形づくった。情報科学の基礎となった記号論理学もこの運動から生まれた。実証的事実の記述とそれらの間の論理的関係を明白にすることが、科学の営みであるとするのだ。心理学における行動主義もこの運動に支えられたと言えよう。

しかし、論理実証主義に基づく科学論は、50 年代から批判にさらされる。言語はゲーム規則 (L.Wittgenstein 1953) のように恣意的であり、客観的と思われる科学的観測は我々の理論に依存する (N.R.Hanson 1958) というのだ。こうした指摘により、世界は神から与えられたという「所与の神話」が崩壊していく (W.Sellars 1963)。この傾向から、実証が容易ではない心的構造が容認され、「認知革命」へと至ったとみることもできる。

その後実証主義は、反証可能性 (K.Popper 1959) の概念を取り入れ、

モデル構築による科学的方法論として整備された。だが、科学が対象とする事象の「実在性」は揺らいだままであった。「科学革命」は観測事実でなく理論が起こすというパラダイム論 (T.Kohn 1962, 1970) は広く受け入れられるものの、実在性を否定した相対主義 (P.Feyerabend 1978) を招くのである。こうした動向は、科学は社会状況によってつくられるのだとする社会構成主義 (K.J.Gergen 1985) にもつながっていく。

哲学の分野で科学的方法論が議論されるなかで、認知科学はモデル構築による方法論で着実な歩みを進めていた。しかし、いざ「心をもつ機械」などと理想を語ったときには、哲学的な議論は避けて通れないのであった。「心」とは何である(とする)かということは、「心身問題」と呼ばれる哲学上の積年の課題であるからだ。古くはデカルトの「心身(心脳)二元論」まで遡れる。

認知科学者が、心を研究対象とする前提は、外界との入出力関係と心的状態同士の関係という機能的プロセスで心をつまようとするところにあり、この立場は機能主義 (H.Putnum 1961) と呼ばれる。心と脳とを一元的に扱うひとつの試みであるが、「心」を重視する立場からは異議のあるところである。

「心をもつ機械」と称する「心」に、初めて積極的に異議を唱えたのは、現象学哲学者 (H.L.Dreyfus 1965) であった。ハイデガーが言うように、我々人間は世界内存在であるはずであるのに、世界から切り離されたコンピュータが心をもつとは何を意味するのかといった具合である。さらに、我々がもつ「心」という感覚に訴える批判も現れた。他者の心はいったい理解できるものなのか (T.Nagel 1974)、言葉を理解するとはどういうことか (J.Searle 1980) といった議論である。こうした批判を受けて、人工知能研究から方向転換する研究者 (T.Winograd 1986) も現れた。

一方で、認知科学の成果を取り入れ哲学的な議論を展開する認知哲学者も現れている。信念や欲求といった通俗心理学の用語を記述していくことで、心の記号表現への還元がなされるとする機能主義をつきつめた立場 (J.Fodor 1981) と、脳のなかの神経回路網における活動を記述することで、心の結合表現への還元(あるいは心自体の消去)がなされるとする消去主義の立場 (P.M.Churchland 1985) での論争も続いている。

哲学者が問題にしているような「心」を認知科学の枠内に含めるには、「志向性」と「感覚質（クオリア）」の自然化を行わねばならないと言われる（N.Nobuhara 1999）。志向性とは、意味が何らかの対象に対して現れるという性質であり、感覚質とは、我々が「赤い」とか「痛い」とかに伴って感じる純粋な感覚である。両者とも、「意識主体」に不可欠な要素であり、かつ何か物的な事象に帰着させることが不可能に思われる。けれども認知哲学者の一部は、意識の多元的草稿理論（D.Dennett 1991）などで、こうした難題に果敢に挑んでいる。近年、「心」にまつわる哲学的議論はとくに盛んになっており、「心の哲学 Philosophy of Mind」（J.Kim 1998）と呼ばれる分野を形成しつつある。

心の哲学のなかで、認知科学の動向に照らして優勢な状況にあると言える立場は、生態学的実在論（J.J.Gibson 1979）である。人工知能は、外界から独立した知性を実現するのではなく、環境に埋め込まれて行動する知性の実現に向かっている。またそれは、進化のシミュレーションによって実現されるとも、しばしば主張される（S.A.Kauffman 1993）。認知的学習も状況に埋め込まれて達成されると捉える理論（J.Lave and E.Wenger 1991）が注目される。意識の心理学的理解と生理学的理解はともに、生物進化の歴史を背景にして成立する（N.Humphrey 1986）とみなされてきている。生態学的実在論は、ひとつの心、あるいはひとつの主体を論じる際には、それを含む環境とそれをもたらした歴史を空間的・時間的な全体として、合わせて捉えようとする。それは相対主義の脅威をかわし、実在性を維持する立場ともなるのである。そんななかで、ダーウィンの進化論が再び強力な哲学思想として浮上している（D.Dennett 1995）。

現在、心の哲学は興味深い状況を呈している。哲学者が科学と接近し、心の自然化や、心の一元的理解（D.J.Chalmers 1996）を試みる一方、神経生理学者（J.C.Eccles 1989）や数理物理学者（R.Penrose 1994）などの科学者が、二元論的立場で心の独自性を主張しているのだ。攻守ところをかえた状況である。また、我々の思想を語る営み自体が自然化の枠内に入ってしまうのならば、我々は心を適切に語るどころまで進化していない（C.McGinn 1996）という可能性も出てくる。考えてみると、その指摘もなかなか的を射ているような気がしてくる。今ますます、心の哲学がおも

しろい。

## 参考文献

- ポール・サガード「マインド：認知科学入門」共立出版, 1996
- 橋田浩一他「講座認知科学(1) 認知科学の基礎」岩波書店, 1995
- N.A. スティリングス他「認知科学通論」新曜社, 1987
- スタン・フランクリン「心をもつ機械」三田出版会, 1995
- ジョン・アンダーソン「認知心理学概論」誠信書房, 1980
- 松本修文編「脳と心のバイオフィジックス」共立出版, 1997
- 苧阪直行編「脳と意識」朝倉書店, 1997
- 門脇俊介「現代哲学」産業図書, 1996
- S.Bem and H.L.deJong: Theoretical Issues in Psychology, SAGE, 1997
- 石川幹人他「量子論と心」人工知能学会誌, Vol.14, No.3, pp.413-424, 1999
- 石川幹人「生物進化と人工知能設計における構造の役割」明治大学教養論集, No.318, pp.1-28, 1999
- 石川幹人「構成論の心理学考 - 心の科学の臨床的視座」明治大学教養論集, No.327, pp.59-79, 2000
- 石川幹人他編著「心とは何か - 心理学と諸学の対話」北大路書房, 2001