

5.12 ■ 情報処理技術 — 過去十年そして今後の十年 —

# 人工知能の身体性と社会性

中島 秀之 電子技術総合研究所

## ■ 知能とは

まずは“知能”的定義をしなければなるまい。しかしながら、これは現状では不可能である。定義できるほど知能のことが分かっていない。しかし、知能をある程度分類してみせることは可能である。

かなり原始的な生命であるアーバでも、環境からの入力を元に栄養のある方向を探知し、そちらへ移動し、栄養を補給することによって自己を保存する能力を持っている。おそらくこれが最低限の知能と呼べるものではなかろうか？（実際、そのような生存を目的としたロボットの研究もある<sup>4)</sup>。）アーバはおそらく外界からの刺激に受動的に反応するだけであるが、蟻や鳥などは巣を作つて、環境を自分に有利に変える能力を持っている。道具を使う鳥もいる。しかし、彼らは先祖代々伝わる（遺伝）手法を繰り返すばかりで、進歩がない。これに対し人間は環境から学び、新しい手法を産み出し、それを伝えていくことが可能である。このような観点で知能を分類してみる

と以下のようになるのではなかろうか：

1. 外界に反応し自己を保存する能力
2. 環境を自己に有利に変更する能力
3. 学習能力
4. 予測能力
5. 道具を使う能力
6. 抽象的記号操作を行う能力（コミュニケーションを含む）

これらの能力は排他的ではない（たとえば道具を使うためには環境を変え、結果を予測（プラン）する能力が必要である）。しかし、あえて分類してみると、反応・環境変更・道具使用の能力は身体性とかなり密接な関係があるのに対し、学習・予測・抽象的記号操作は身体を第一義的に必要とはしない。

上記のように、外界接続性の定義すらできないのではないか？ その意味で、知能と環境とは密接に関連していると言わねばならない。このような観点から人工知能の歴史を振り返ってみたい。

# 特集 Special Features

## ■人工知能の歴史

少し歴史を紐解いて、コンピュータと人工知能の発展を振り返ってみよう。

1936	Turing : チューリングマシン
1946	ENIAC (世界初の電子計算機)
1948	Shannon : “通信の数学的理論”
	Wiener : “サイバネティクス”
1950	Turing : チューリングテスト
1952	IBM 701 コンパイラ (瑞)
1956	IBM FORTRAN 電気試験所 : トランジスタ式 コンピュータ (日) ダートマス会議 (人工知能の始まり)
1959	Algol
1960	COBOL Lisp パーセプトロン
1963	対話型 Lisp Minsky : “人工知能への道” Feigenbaum : “コンピュータと思考”
1965	エキスパートシステム DEN-DRAL ファジー理論
1970	自然言語理解システム SHRDLU
1972	Prolog Dreyfus : 人工知能批判
1975	エキスパートシステム MYCIN Minsky : フレーム理論
1977	McCarthy, Hayes : フレーム問題
1986	Rumelhart, Hinton ら : PDP
1991	Brooks : “表象なしの知能”
1995	Russell : 限定合理エージェント

人工知能の歴史は計算機やプログラミング言語の歴史とほぼ同時に始まっていることが分かると思う。人工知能研究の始まりとされるダートマス会議は、世界初の電子計算機の誕生から10年を経てはいるが、FORTRAN (実用化された最も古いプログラミング言語) 誕生の次の年である。それ以前には一般の研究者が計算機を使うのは困難であったことを

考えれば、計算機の普及と人工知能の誕生はほぼ同時期であるといってよい。

1950年にAlan Turingはチューリングテスト<sup>7)</sup>と呼ばれる、知能を持つことに関する仮想実験を提唱した。これは、簡単にいうと、次のようなものである。テレタイプを2台用意する。1台は他のテレタイプにつながっており、他の人間が座っている。もう1台はコンピュータに接続されている。このコンピュータのプログラムは人間の反応をシミュレートするようにできている。そして、この2台のテレタイプのどちらが人間でどちらがコンピュータか分からなければ、このコンピュータプログラムは知能を持っているといつてよいというものである。

チューリングテストでは人間はどんな質問をしてもよい。詩を作らせててもよいし、文学作品の感想を聞いてもよい (ディックの“アンドロイドは電気羊の夢を見るか”で、このチューリングテストをアンドロイドに対して行うシーンが興味深い)。プログラムの方も、人間をまねるためにあらゆる努力をする。たとえば計算問題に関しては、時間をかけたり、時々計算を間違えたりするわけである。ありとあらゆることが可能であるが、テレタイプの交信に限定されているところがみそである。そうでなければ見かけや行動能力が効いてくる。

この例でも分かるように、初期の、つまり第一世代の人工知能はシンボル処理の時代であったといえる。知能の本質はシンボル処理だけで表現できるという“物理記号仮説”が提唱されている。Herbert Simonは1995年のIJCAIでもその立場を再強調した。

この第一世代人工知能はある意味でSHRDLU<sup>8)</sup>で最盛期を迎えることになる。これはシミュレーションの積木の世界でロボットに自然言語による対話を通じて指令を与えることのできるシステムで、ドメインを限定すれば計算機システムにも言語理解が可能であることを示した。SHRDLUにおける言語の理解とは、積木の世界に関するさまざまな質問への応答や動作指令の実行として表現される。さまざまな言い回し、指示代名詞などに潜在する曖昧性を命令の実行可能性の観点から選別し、可能な解釈が一意であれば良しとする点が注目された。SHRDLUはシンボル処理人工知能の代表的成果ではあるが、言語理解を、言語による応答のみではなく、シミュレーションながらも行動 (外界の操作) で示している点にも注目されたい。

続くエキスパートシステムの時代は、人工知能の夏とも表現されるほど、人工知能が期待された時期であった。限定された範囲であるが実用に供されたシステムも少なくない<sup>☆1</sup>。しかしながらこの期待はそう長く続かなかった。システムが高度になればなるほどそれが失敗したときのダメージも大きい。飛行機の操縦に例えれば、人が操縦していた時代に、個々の操作に間違いがあってもすぐ修正できた。しかし、自動航行システムへの入力に間違いがあれば大きな事故につながりかねない。いかに高度になっても間違いの可能性のあるエキスパートシステムは実用にならなかった。この理由の1つはシステムが自分の出した結論の帰結を評価できないという点にある。自分の出力のフィードバックが得られないでのある。ある意味で体が必要なのである。

☆1 ここでも、実用化されたものは人工知能と呼ばれることはないという規則が生きている。ユナイテッド航空の自動予約システムは、旅行代理店の通常の人間の能力を上回るエキスパートの手法を実装したもので、まぎれもなくエキスパートシステムであったが、今では誰もそう認識していないようである。

このあたりでニューラルネットの再来(PDP)があるので、本論の主旨とは関連しないので省略する。

エキスパートシステム熱が冷め始めた頃にBrooksに代表される身体性の時代(ロボットの時代)がくる。彼は“表象なしの知能”<sup>1)</sup>を主張し、服属アーキテクチャによるロボットを作った。この設計思想はRosen-scheinらの状況依存オートマトンの考え方に基づくもので、内部表象を持つ代わりに外部の状況を利用しようとするものである<sup>☆2)</sup>。

そして、現在は複雑系の時代である。この考え方については後で詳しく述べるが、従来の意味で工学的に設計・製作できないようなシステムをいかに構築するのかという問題である。もちろん知能もそういった性格を持つものである。

## ■不完全情報

歴史からも明らかなように、人工知能は情報処理の一分野であるが、典型的な情報処理とはかなり異なる性格を持つ分野である。典型的な情報処理とは、OS、コンパイラ、文書処理、会計処理、通信制御、運転制御、通信などで、誤りを含まない正確な処理が要請される。これに対し人工知能研究では“ヒューリスティクス”という用語に代表されるように、大部分の場合は正確な処理よりうまくいくが、ときどきは失敗するかもしれないような処理が研究の中心である。これは“柔軟性”として表現されることもある。

通常の情報処理が完全性と最適性の軸で評価されるとしたら、知能は

柔軟性と効率の軸で評価されるといつてよい。人間は、コンピュータと比較すると多くの分野でより柔軟にうまく振る舞えるが、同時に誤りも多い。誤りは効率のよい処理のために支払わねばならない代償である。以下ではその理由を考えてみたい。

知能が扱う問題の性質の1つとして、情報と処理の部分性が指摘されている。これは原理的に避けようのない問題である。フレーム問題<sup>☆3)</sup>が代表しているように、完全な記述や完全な処理は原理的に不可能である。それを何とかするのが知能である。すなわち、完全な情報を得ようとする努力や完全な処理の手法を探す努力は無駄であり、部分情報の部分処理を効率よく行う手法を探すべきである。Russellはこの考え方を限定合理エージェントという言い方で、エージェントという限られた情報や計算資源しか持たないものの視点からの問題の定式化を主張している。

## ■状況依存性

動物は、部分的特徴だけを使うことによってうまくやっている例が多く報告されている。動物行動学の成果として、一見複雑に見える魚や鳥の行為も、実は状況の一部(自然環境の他に仲間や敵の生物を含む)に対する単純な反応の連鎖にすぎないとする発見がTinbergenらによって報告されている<sup>6)</sup>。たとえば、トゲウオの雄は生殖期の別のトゲウオの雄のみに対して飛び跳ね闘争を行うが、これは、下半分が赤い物体に対する反応であることが分かっている。

模型で実験したところ、色のない正確な模型より、目と赤い色だけがついた円錐体に対してより激しく攻撃を加えた。これは魚が赤色しか見えていないということではない。別の実験で他の色や形態的特徴の細部も見えることは確認されている。しかし、この特定の場合には、そうした他の情報より色に反応するのである。トゲウオは赤い色を持つ魚が多い環境ではうまくやっていけないだろう。しかし、生殖の競争相手となる他の雄のみがその色を示す環境(おそらくこの色は雌にとって誘惑刺激なのであろう)では、うまく振る舞えるのである。

もう1つだけ例を示しておこう。セグロカモメの雛は、餌を求めるときに親カモメのくちばしの先をつつく。親カモメは餌を胃から吐き戻し雛に与える。雛は親鳥を認識しているのではなく、黄色いケチバシの端の赤い斑点に反応している(トゲウオの場合と同様にさまざまな色の模型でこれが確認されている)。このような雛のつつき行動を起こさせる刺激を解発刺激という。雛のつつきが、今度は親カモメが餌を与える行動の解発刺激になっている。このように生物間では互いの行動が相手の次の行動の解発刺激になっており、その結果全体として複雑な行動様式が保たれる例が多く知られている。

処理の大部分を環境の特殊性にゆだねる。これを状況依存性の活用と呼ぶ。つまり、知識の表現や推論規則を状況依存の形で記述する、より柔軟で効率のよい理論や行動が可能となる。このようにして、情報処理方式を研究するが人工知能の特徴である。そして、状況依存性を考えるときには、さまざまな環境との相互作用つまり入出力が必要となる。それを与えてくれるのが体である。

<sup>☆2</sup> 実は、この考え方自体はSimonがすでに記述<sup>5)</sup>しているものである。彼は蟻の歩いた軌跡は複雑に見えるが、その理由は蟻の内部にあるのではなく、地形の複雑さにあると説明している。これはさらに、動物行動学の主張である、一見複雑に見える動物の行動も、実は単純反射の状況に依存したうまい連続にすぎないという考え方<sup>6)</sup>に遡ることが可能である。

<sup>☆3</sup> 元々は、ある行為の影響を記述しようとしたとき、その行為によって変化することがらと変化しないことがらをいちいち明示的に記述するのは(記述、推論において)煩わしい(計算量が指数関数的に増大する)という問題。ここでは、その発展形としての行為の前提条件や帰結の完全記述が不可能であるという問題を指している。

## ■複雑系と身体性

ロボットを物理シミュレータの中で歩かせるより、実機を歩かせる方が簡単だという研究者もいる。構造物と環境がある程度複雑になると、シミュレーションも大変である。むしろ、環境をそのまま使う方が容易であるのもうなづける。

トゲウオやセグロカモメのように、環境に埋め込まれた指標を上手に使えば効率のよい情報処理が可能である。Brooksはこのようにして昆虫型ロボットを作った。しかしながら、このロボットは外界に反応し、自己を保存する能力は持っているが、それ以外の環境を自己に有利に変更する能力、学習能力、予測能力、道具を使う能力、抽象的記号操作を行う能力はいっさい持ち合っていない。知能レベルとしては大変低いと言わねばなるまい。人間が頑張って作り込めば、Brooksのロボットに、昆虫のように巣を作る（環境を自己に有利に変更する）能力を持たせることは可能かもしれない。さらに限定された学習能力も持たせられるかもしれない。しかしながら、そこから先はシンボル処理抜きには困難なのではあるまい。

ホンダの二足歩行ロボットを代表として、最近ではヒューマノイド研究が始まっている。しかし、ヒューマノイドには知能が必要である。外見だけが人間に似ているのではない。

ヒューマノイドの体は、工学的な意味では超多自由度系という特徴を有している。特定のタスク遂行のために必要十分な自由度を与え、それを最適制御する工業用ロボットとは異なり、多目的で柔軟であることを良とする。そのような超多自由度の複雑系を我々はどのように設計・構築・制御するのかが問題である。ここで考える複雑系とは多層システムのことと言い換えてよい。



ヒューマノイドの知能も同じではなかろうか。あり余る自由度を必要に応じて限定して使うような知能のアーキテクチャが必要である。我々はそのような可能性の1つとして動的服属アーキテクチャ<sup>3)</sup>を提案し、サッカーエージェントに組み込んだ。

## ■知能の社会性

知能と環境の相互作用の重要性について述べたが、知能間の相互作用も忘れてはならない。たとえば、人類の文明や文化は遺伝的にではなく社会的に進化し、継承されている。親の世代の知識が子の世代に、学校などのさまざまな機会を通して伝えられるのである。このような場面では単一知能の学習機能だけを考えていたのでは問題の本質を見逃す恐れがある。問題解決においても同様である。1人では解けない問題が集団では解決できたりする。このような社会知能は蟻や蜂にも見られる。蜜蜂も使う抽象的記号はこのような社会知能としての必要性（つまり、個体を超えた情報伝達）から生まれたのかもしれない。社会学でも進化的観点は古くから持たれているようである。

このような集団による問題解決を扱う研究分野は分散人工知能として始まり、現在ではマルチエージェントと呼ばれることが多い。マルチエージェント研究には2つの意味がある。1つはマルチ、つまり複数知能の相互作用の研究である。もう1つはエージェント、つまり限定された知識と計算資源のもとに環境との相互

作用を通じて自律的に行動する主体の研究である。最近ではソフトウェアエージェントの他に、複数のロボットによる協調行動の研究も始まっている。その1つとしてロボカップ<sup>2)</sup>という研究の枠組みも日本発で構築された。

## ■まとめ

空を飛ぼうとする試みが、鳥のはばたきをまねることから始まったよう、人工知能の研究は人間の知能をまねることから始まった。いまだに飛行機の原理には到達していない（まだ人間を真似ている）。

外界から切り離された脳を考えることが困難であるように、知能の働きは、閉じたシステムとしてではなく、外界に開いたシステムとして捉える必要がある。そのためには自律システムとしての知能の研究が重要であり、そのシンボル版がエージェント、物理版がロボットである。そして両者を統合した究極がヒューマノイド研究である。さらには、単一の知能ではなく社会的知能を考える必要もある。

### 参考文献

- 1) Rodney, A.: Brooks, Intelligence without Representation, Artificial Intelligence, Vol.47, pp.139-160 (1991).  
(柴田正良訳: 表象なしの知能, 現代思想, 18 (3), pp.85-105).
- 2) Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I., Osawa, E. and Matsubara, H.: Robocup - A Challenge Problem for AI -, AI Magazine, Vol.18, No.1, pp.73-85, spring (1997).
- 3) Nakashima, H. and Noda, I.: Dynamic Subsumption Architecture for Programming Intelligent Agents, In Proc. of International Conf. on Multi-Agent Systems 98, pp.190-197, AAAI Press (1998).
- 4) Pfeifer, R. and Scheier, C.: Understanding Intelligence, MIT Press (1999).
- 5) Simon, H.A.: The Sciences of the Artificial, MIT Press, Cambridge, Massachusetts, Second Edition (1981).
- 6) Tinbergen, N.: The Study of Instinct, Clarendon Press, Oxford (1951).  
(邦訳: 本能の研究, 三共出版 (1975)).
- 7) Turing, A.M.: Computing Machinery and Intelligence, Mind, Vol.LIX, No.236 (1950).
- 8) Winograd, T.: Understanding Natural Language, Academic Press (1972).

（平成12年3月6日受付）