人工知能の人工生命への接近

久木田水生

1. 序

人工知能(AI)は人工的なシステム（コンピュータやロボット）によって人間の認知や推論などの過程を模倣する試みである[1]。外界を認知し、それに基いて推論し、一定の判断を下すことは、およそあらゆる生物にとって不可欠な能力であるから、従ってある意味でAIは人工生命の一部と言うこともできる。しかしながらこの二つの分野は、誕生し確立され始めた当初から、顕著に異なった方法論を採用していた。AIは人間の知的振る舞いを入力と出力という観点で捉え、それと等価な振る舞いをするシステムを開発することを目指してきた。一方で人工生命は、出力としての振る舞いそのものよりも、そのような振る舞いを生み出すメカニズムの本質を明らかにすることを目指してきた。そのため伝統的にAIと人工生命は――その明白な親近性にもかかわらず――区別されてきたのである。

しかしながら従来の方法に基いた AI の限界が様々な点で認識され、さらにはコンピュータやロボティクスの技術が発展するにつれて、AI の方法も大きく変化してきた。そして現在の AI は人工生命にますます接近している。本稿では AI における近年の研究を取り上げ、それが人工生命と多くを共有していることを見る。

2. 記号的AI

19 世紀のイギリスの数学者チャールズ・バベッジは歯車や滑車などの複雑な組み合わせからなる計算機、「解析機関」を考案した。これにはCPU、メモリ、入力装置、出力装置という、現在のコンピュータの基本的な構成要素がすべて備わっていた。とりわけ注目するべきは、パンチカードにプログラムされた作動パターンを逐次的に実行することによって様々な計算に対応できるように設計されていた、という点である。それゆえに解析機関はプログラム可能な汎用計算機の元祖であると評価されている[2]。バベッジとその協力者であるエイダはしばしば「この機械は考えていると言えるのか？」という質問に煩わされた。エイダは解析機関についての論文の中で、この質問に対して否定的な見解を述べているが、同時に「最終的には機械を使った経験をもとにして、この問題に対する答えを出すべきだ」と留保している(Feigenbaum & McCorduck, 1984)。

バベッジが解析機関の着想を得てからほぼ 100 年後、推論や計算という概念の本質を捉える試みが数学と記号論理学の分野で盛んに行われる中で、アラン・チューリングは、計算とは与えられた記号に対し決まった規則に従って逐次的に変換を施すことによって一定

の解を得ることである、という考えに基いて、計算の過程を抽象的な機械の動作として表現する一つのモデルを作った。現在では「チューリング・マシン」として知られる彼の考案した理論的機械は、その構成、およびそのプログラムの文法において非常に単純であるが、しかしその計算能力、その汎用性は驚くべきものであった。実際、現在でも自然数上の計算に関して（従って整数や有理数上の計算に関しても）、チューリング・マシンよりも計算能力の高い計算モデルは作られていない。それゆえ、チューリング・マシンによって計算可能な関数が，人間にとって実効的に計算可能[3]な関数のすべてだということすら主張されている（チャーチ＝チューリングのテーゼ）。

チューリング・マシンの成功が余りにも顕著であったために、自然数上の演算のみならず人間の思考一般を何らかの計算とみなし、その過程を機械とプログラムによって実現しようという発想が生まれたのも無理のないことだった。必要なのは、人間の思考が従っている規則を厳密に定式化すること、そしてその規則に従った思考を機械によって実行するためのプログラムを組むことだ――そのように思われたのである。その際、入出力の値や、計算の過程を表現する言語は、機械の中ではあくまでも形式的な記号として扱われた。機械は入力された記号を一定の規則に従って処理し、その結果を出力する。記号の意味を解釈する役割はもっぱらプログラムを与え機械を操作する人間の側に任されていた。むしろ解釈や理解という側面を棚上げして、人間の思考をどこまで形式的・機械的な計算に置き換えることができるか、ということを追及したことが、初期の AI の成功の理由であったとも言えるだろう。

チューリングはさらに、今日「チューリング・テスト」として知られる、機械が思考しているか否かを判定するためのテストを提案し、機械の思考についての問題に火をつけた。彼は知性を、「知性的に振舞う能力」と捉え、外部から観察して知性的に振舞っているように思われる行為者は、実際に知性を持っていると見なすことができると主張したのである。彼は知性的な振る舞いの代表的なものとして、言語を使用してコミュニケーションを行う能力を挙げ、言語を使用して円滑にコミュニケーションを行っているように見えるものは、実際に知性を持っていると見なすことを提案した。

チューリングの提案に対しては様々な方面から激しい批判が向けられた。その多くは人間の機械に対する優越性に対する根拠のない信念（もしくは願望）から行われたものだとしても、人間の思考には単なる記号の計算以上のものがあるということはやはり明らかに思われた。それでも人間の思考を規則に従った記号処理過程としてシミュレートするという試み（記号的 AI）は、チューリング以降も続けられており、ギリースが「チューリング的伝統」(Gillies, 1996, chap. 2)と呼ぶ一つの大きな流れを形成している。この伝統において、

一階述語論理の定理の自動証明、対象の分類、データ・マイニング、帰納的推論や発想的推論(abduction)を実行するシステムなどが開発され、様々なエキスパート・システムや機械学習に応用されている。しかし現在ではこういった試みは通常、知性を持った機械を作るという目標からは切り離されたものとして追求されている。開発者たちは、何らかのチューリング・テストにパスする機械を作ることを目指している。しかし彼らは知性を持った機械を作ることを目指しているわけではない。彼らは彼らの作ったシステムが知性を持っていると言えるかという問題に関して積極的にコミットしようとしないのである。

彼らがその問題にコミットしない主たる理由は、彼らが AI をあくまでも人間の思考を補助する道具として捉えているからであろう(Feigenbaum & McCorduck, 1984; Rheingold, 1986)。彼らの研究が実用上、大きな意義を持っていることは明らかだし、さらにギリースが指摘するように、科学的方法や論理学についての私たちの理解に大きなインパクトを与える可能性も大きい(Gillies, 1996, chaps. 4f)。しかし彼らの研究が、機械が知性を持てるかという問題に対して何かしらの示唆を与えてくれると期待することは誤りである。

だからといって機械に知性を持たせるという課題が全面的に放棄されたわけではない。ただし現在この課題に取り組んでいる AI 研究者たちは、チューリング的伝統に連なる研究者たちとは異なる方法を取っているのである。

3. 「記号接地問題」

チューリングは言語を使用する能力をもっぱら構文論的操作の能力、すなわち入力された記号列を形式的に操作し、その結果として得られた記号列を出力する能力と捉えていた。しかし現実に言語を使用する人間の活動には、このような単純な入出力計算としては捉えられない側面がある。サールによる「中国語の部屋の議論」はそのことを明快に示している(Searle, 1980)。ある部屋に中国語をまったく知らない人間と、中国語の完全な応答表が入れられていると考えよう。この部屋に中国語で質問を書いた紙を入れる。すると中の人物は応答表に従って答えを書いた紙を返す。その答えは質問に対する完璧な答えになっており、それゆえ外部の観察者にとっては、中の人物は中国語を適切に使用しているように思われる。チューリング・テストと同様の基準からすれば、この人物は中国語を使用できる知性を持っていると判断してよいはずだが、しかしこれは明らかに不合理である。

サールによれば、チューリングの誤りは言語能力のテストを構文論的能力に求めたところにある。そもそも言語を適切に使用しているといえるためには、その使用者は言語の意味を理解していなければならず、そして意味の理解とは、単に構文論的に言語を操作できるという以上のことを含んでいる。人間は会話する際、相手の言葉に対して機械的に決ま

った返事を返しているわけではない。人間は発話のなされた状況に応じて相手の言葉を解釈・理解し、それに基いて適切と思われる返答を返しているのである。この解釈・理解という過程が、人間の言語使用にとっては本質的であるように思われる。従ってある AI システムが人間の言語使用を十分にシミュレートするべきものであるならば、そのシステムはそれが操っている記号の意味を、何らかの仕方で理解している必要がある。このように主張することで、サールは言語使用能力についての問題の焦点を、「意味の理解」の問題の上に置き換えた。

現在、AI 研究において、意味のある記号を使用する AI システムを作ることが一つの大きなテーマになっている。この問題に対するアプローチは、チューリング・テストをパスすることを目指すためのアプローチとは本質的に異なったものにならざるを得ない。というのもチューリング・テストは記号操作の能力だけをテストするものである一方、記号が意味を持つという作用は、記号、対象、そして記号の使用者（解釈者）という三項関係において初めて成立するものだからである。従ってある記号が物理的対象を指示するなら、その記号と使用者の双方が、その物理的対象に何らかの仕方で結びついていなければならない。それではそのような記号的関係はどのようにして成立するのだろうか。

ハーナッドは形式的な記号体系にとっては意味が外在的(extrinsic)である一方、人間の認知は単なる記号の操作ではなく、私たちの頭の中では記号的表象の意味が内在的(intrinsic)であるので、形式的な記号体系が私たちの頭の中の意味に対するモデルとして成功する見込みはない、と結論する(Harnad, 1990)。彼は形式的体系を実装した記号的 AI が置かれた状況を、中国語を学ぶのに中国語－中国語辞書だけを与えられた人間にたとえる。その人物はある言葉を意味を知るために別の意味の分からない言葉を参照するばかりで、どこまでいっても言葉を有意味なものとして理解することはないだろう。あるシステムの操る記号が単に形式的なものではなく、有意味な記号であるためには、記号の対象が存在している世界に、記号が「接地」させられなければならない、とハーナッドは言う。それではどうすればある記号的 AI システムの扱う記号を記号以外の何かに「接地」させることができるのだろうか？　彼はこれを「記号接地問題(the symbol grounding problem)」と呼ぶ。

記号接地が単なる意味論的規約の問題ではなく、対象の認知の問題であることには注意する必要がある。例えば記号「○」がボールを指すものと取り決めようと考え、

○　＝　ボール

という表現によってその規約を表したとする。しかしこの方法によって○を接地させるこ

とが出来るのは「ボール」というメタ言語における表現をすでに接地させることが出来ているシステムだけである。最終的には実際のボールを当のシステムに提示することによってしか、〇を接地させることは出来ないだろう。そのためにはシステムが対象を認知する能力を持っていなければならない。

ハーナッドは人間が外界の対象を認知し、そして表象を形成する過程を分析しながら、この問題に対する解決の指針を提案する。まず彼は人間が外界の対象を区別したり同定したりする能力に注目する。個々の入力を区別する能力は、「イコン的表象(iconic representations)」，すなわち「離れた対象から感覚表面上への投射の内的な類比的変形(internal analog transforms of the projections of distal objects on our sensory surfaces)」を持つ能力に依存するが、しかしある範疇に属する対象を同定するためにはイコン的表象では十分ではない。なぜならば同定のためのイコンが数多く存在しそれらが連続的に混ざり合っているために、どのイコンがその範疇の成員のイコンなのかを特定するという別な問題が生じるからである。従って対象がある範疇の成員であると同定するためには、ある範疇の成員をそうでない対象から区別するのに十分な感覚的投射の「不変的特徴(invariant features)」が選択されなければならない。このように選択された特徴をハーナッドは「範疇的表象(categorical representations)」と呼ぶ。

このように対象を分類するのに十分なイコン的表象と範疇的表象が得られたならば、ある名前にその表象を結びつけることが可能である。この名前は表象を介して接地しているということが出来る。さらにそこから通常の形式的体系と同様に複合的な表象を作っていくことが出来る。ハーナッドは次のような例を挙げている。「馬」と「縞」という記号が、イコン的表象と範疇的表象によってそれぞれ馬と縞模様という範疇に接地されているとする。このとき

「縞馬」　＝　「馬」＆「縞」

と定める。このとき記号「縞馬」が指すのは「馬」＆「縞」という記号列である。しかし「馬」と「縞」がそれぞれ接地できているので、「縞馬」もそれらの接地を継承することが出来る。このようにして縞馬を一度も見たことのない人間でも縞馬の「象徴的表象(symbolic representation)」を持つことができる。

上のような過程を人工的システムにおいて実現することが出来るならば、確かにそのシステムの扱う記号を接地させることが出来るかもしれない。しかし範疇的表象を作る過程を AI で実現することはかなりの困難を伴うだろうことは容易に想像できる。ハーナッド

自身の提案は、コネクショニスト・ネットワークによる学習を利用するというものである。そのような「ハイブリッド・システム」は記号的 AI とコネクショニスト・ネットワークのそれぞれの欠点を補うことが出来るだろうとハーナッドは言う。

Harnad (1990)は、記号接地問題に対するアプローチのごく大まかなスケッチを提示しているに過ぎないが、しかし知性の実現という点では否定的に見られていた記号的 AI が有意味な思考を持つ可能性があり，そしてそのような AI にとっては、世界を認知・表象し、学習によって概念を形成する能力が本質的である、ということを示した点で重要である。ここでの一つの教訓は、知性のモデルを構築するためには、知的振る舞いを生み出すメカニズムに注目するべきだ、ということである。20 世紀前半には人間の思考を生み出すメカニズムがどのようなものであるかは、ほとんど知られていなかった。それゆえ初期の AI が、知的活動の入力と出力のみ、あるいはせいぜいその言語的思考過程のみに注目し、その過程を生み出すメカニズムを無視せざるを得なかったことは理解できる。しかし現在、私たちはそのメカニズムについて、以前に比べてかなり多くの利用できる知識を持っている。のみならず私たちはそのメカニズムをシミュレートするための技術も持っているのである。次節では脳のメカニズムを「ニューラル・ネットワーク」によってモデル化するコネクショニストの試みを検討しよう。現在多くの分野で異なる目的のために多種多様なニューラル・ネットワークが作られている。ここでその全体を網羅的に概観することは不可能である。従ってここではニューラル・ネットワークの基本的な発想と構造についてのみ触れる。

4. ニューラル・ネットワークによる認知のシミュレーション

認知や思考にとって最も重要な器官が脳だということは明らかであり、従って認知や思考のモデルを構築するための最も有望な方法は、脳の構造とその働きのメカニズムを人工的に再現することである――そのようにコネクショニストたちは考える。私たちの脳は、記号的 AI が記号を操作する方法とはまったく異なる方法で、情報を処理している。第一に、脳の中で情報は記号的 AI が扱う分節化された記号の組み合わせという形を取っていない。脳の中で起こっているのは、あくまでも生理学的な過程であり、そしてその過程が脳による情報処理の過程である。第二に、脳は一つの入力に対して一つの出力を返して処理を終了するのではない。脳は絶えず流れこむ情報に対してリアルタイムに状態を遷移させることによって情報を処理している。ここにあるのは外部環境と脳という二つのシステムの間のダイナミックな作用である。コネクショニストたちが捉えようとするのは、このような脳の活動のメカニズムである。そしてこのメカニズムを模したモデルによって認知

や思考のシミュレーションができるならば、脳の活動こそが思考であるという主張を支持する根拠があるだろう。コネクショニストはこの発想に基き、脳のモデルを構築しようと試みる。

生物の脳は多数のニューロンが集まり、それらが複雑に結合することによって構成されている。そこでコネクショニストはコンピュータの中に、ニューロンの働きを抽象化したもの（以後はこれをニューロンと呼ぶ）を単位とするネットワークを構築して脳をモデル化しようと試みる。このようなシステムは「ニューラル・ネットワーク」、または「コネクショニスト・ネットワーク」と呼ばれる。ニューラル・ネットワークでは入出力されるのは記号ではなく、複数の数値からなるベクトルである。入力されたベクトルは、入力ユニットを構成するニューロンに対する刺激として受け取られ、その刺激は複雑なネットワークの中でニューロンからニューロンへと変形を施されながら伝達され、その結果が出力ユニットから再び一つのベクトルとして出力される。

ニューラル・ネットワークの振る舞い方は、ニューロンの結合の構造、および個々の結合（シナプス）に対する「重み付け」によって決定される。これは次のようなものである。一つのニューロンはシナプスによって結ばれた多数の他のニューロンから正（興奮性）または負（抑制性）の信号を受け取る。正の信号と負の信号の差が一定の値（閾値）を超えるとニューロンが「発火」し、そのニューロンから信号が発信される。ネットワークの中の各シナプスには異なる「重み」が設定されており、重みの強いニューロンからの信号はそれだけ増幅して受け取られる。この重み付けがニューロンの振る舞いを決定するプログラムになるのである。

ただしニューラル・ネットワークにおいては、個々のシナプスに与える重みがシステム全体の振る舞いにどのように影響するかということを予測することはほとんど不可能である。そこでコネクショニストたちは次のようなアプローチを取る。最初に、ネットワーク中の各シナプスに対してランダムに（ただしあまり極端な値にならないように）重みを割り当てる。それからシステムに入力を与え、さらにそれに対する望ましい出力を教えておく（これを「教師信号」と呼ぶ)。システムを作動させると、システムはその入力に対して自らが出力する値と、教師信号との差を計算し、その差が縮まるように、重み付けを少しずつ変化させる。

コネクショニストの方法と従来の記号的 AI の方法との違いは際立っている。記号的 AI においては、プログラムはシステムの働きを制御するために与えられるものである。従ってプログラマは、システムの振る舞いに対してすべての責任を負う。またシステムは最初に与えられたプログラムに忠実に処理を実行するだけで、その結果によってプログラムを

変化させるようなことはしない。コンピュータは人間によって命じられたこと以外に何も出来ない、という AI に対するおなじみの反論の根拠はここにある。しかしニューラル・ネットワークにおいては、プログラムを組んだ時点でそれがどのように動作するのかはプログラマにも予想できていない。プログラマは重み付けの初期値をランダムに設定するだけである。その後はシステムが、入力に対する出力と正解との間の誤差を自動的に修正するように重み付けをし直していく。ここでの人間の役割は、ある問題に対する正しい答えを示して、システムが自動的に学習をしていくのを助けることである。

ニューラル・ネットワークには、信号は入力ユニットから出力ユニットに向けていくつかの階層を通過しながら一方向に流れていくだけのものと、あるニューロンから発信された信号が同じニューロンにもどる（フィードバックされる）経路を持つネットワークとがある。前者は「フィードフォワード型ニューラル・ネットワーク」と呼ばれる。後者にはさらに「相互結合型ニューラル・ネットワーク」と「再帰型ニューラル・ネットワーク (recurrent neural network; RNN)」に分類される。フィードフォワード型ネットワークでは、例えば音声言語のような時系列データを適切に処理することが出来ない。しかし RNN ではシステムの以前の状態が現在の状態に反映されるので、時系列データを扱うことが可能になる。エルマンは RNN を利用して、時系列データとして与えられる文の学習のシミュレーションを行っている(Elman, Bates, Johnson, Karmiloff-Simth, Parisi & Plunkett, 1997; 守・都築・楠見, 2001)。エルマンの開発した単純再帰型ネットワークは、文法の学習や語の予測に関して優れた能力を持つことが分かっている。

再帰型ニューラル・ネットワークに関して興味深いのは、その状態の時間的遷移がカオス的な振る舞いをすることである。エルマンらは、RNN の非線形のダイナミクスと、そこに現れるカオスの重要性を強調する(Elman et al., *ibid*.)。カオスは、軟体動物の神経細胞、ラットの嗅球など、実際の生物の神経現象において現れることが知られている。津田は力学システムのカオス的振る舞いにおける情報の保存、呼び出しなどのメカニズムを、記憶や想起といった心的過程と関連付け、そこから脳の機能をカオス的なモデルで説明することを提案している(津田, 1990)。またチャーチランドは、再帰型ネットワークが認知のみならず運動の制御の面においても有用であることを指摘する(Churchland, 1996, chap. 5)。

ニューラル・ネットワークは人工知能や人工認知システム、発達心理学、言語学、ロボティクスなど、様々な分野で応用されている。認知モデルとしての成功は、記号的 AI を圧倒していると言ってよいだろう。このニューラル・ネットワークの成功は一体に何によるのだろうか。一つの明らかな要因は、ニューラル・ネットワークの持つ可塑性であろう。ニューラル・ネットワークにおいては比較的単純なユニットが多数集まり、まとまって一

つの仕事をしている。一つのユニットの全体に対する貢献はさほど大きくない。だからこそ個々のユニットに対する操作によって少しずつシステムの振る舞いを変化させることが容易なのである。ニューラル・ネットワークにおいては、例えば一つのニューロンを取り除いたとしてもその影響はたかがしれている。これを有限状態機械と比較すると違いは明らかである。ある有限状態機械から、一つのノードを取り除いてみることを想像してみればよい。ほぼ確実にその機械はもはや正常に動作しないだろう。それゆえ有限状態機械のような AI のプログラムを、求める結果に近づけるために徐々に変化させるということは困難であり、そのプログラムを自動的に修正するようなアルゴリズムを作ることも難しい。どんなチューリング・マシンの動作でもシミュレートできるチューリング・マシンを作ることは出来る。しかし求める関数を計算するチューリング・マシンのプログラムを自動的に書くようなチューリング・マシンを作ることは不可能である。有限状態機械を知性のモデルと呼ぶことに対する私たちの躊躇は、ピアジェの「知性とは、方法が分からないときに使うものだ」という言葉に端的に示される。知性とは学習や問題解決の過程において典型的に現れるものであり、すでに何かを知っている状態を指すのではないと私たちは考えている。そしてそれゆえに有限状態機械を知性のモデルと見なすことに抵抗を感じるのである。その点、ニューラル・ネットワークは、可塑性や柔軟性、学習能力において知性のテストのひとつをクリアしているといってよい。そしてまた、そのように「経験」に即して、学習していくからこそ、ニューラル・ネットワークの状態が何かしら現実の世界とつながりを持っており、記号的 AI の扱う記号のような意味のない形式とは異なっているように私たちは感じる。

一方、コネクショニズムに困難がないわけではない。ニューラル・ネットワークはおよそどんな関数でも、学習を通じて近似的にシミュレートすることが出来るが、しかし正確に求める関数そのものが得られる保証はない。厳密な計算や論理的推論を行うには、やはり形式的な記号体系を用いることが必要であるか、あるいは少なくともそちらのほうが遥かに効率が良いように思われる。前節では記号体系を接地させるためにニューラル・ネットワークを利用するというハーナッドのアイディアを紹介したが、コネクショニストの側からも、ニューラル・ネットワークと記号処理モデルを融合させたハイブリッド・モデルによる解決が試みられている(守・都築・楠見, 2001, chap. 2)。このようなモデルにおいては、いかにして二つの異なるレベルの処理を結びつけるかということが問題になるが、この点に関しては現在盛んに研究が行われている[4]。

私たちの思考が記号の計算のようなものを実際に含んでいるのかどうかはともかくとして、少なくともそれだけでは知性を説明することは出来ない。あるシステムの働きが知性

と呼ばれるためには、それが当のシステムに経験と学習とを可能にするようなものでなければならないということが、現在では大方の AI 研究者の見解になっているように思われる。

5. 進化のメカニズムと創造性

デネットは意識の創造を促すものとして、三種類の進化的過程を挙げる。すなわち（1）遺伝による種の進化、(2) 神経様式の進化、そして (3) ミーム[5]の進化である (Dennett, 1990)。これらはすべて、変異、遺伝、選択という、古典的ダーウィニズムが説明する進化のメカニズムに従っている。一般に人工知能による学習や概念の獲得は、(3) の過程を人工的に実現させる試みとして捉えられる。実際、機械学習における最適解の探索方法として、現在、「遺伝的アルゴリズム」が広く利用されている。遺伝的アルゴリズムとは、生物進化の過程を模した手続きによって、与えられた問題に対するより良い解を探索する（あるいは創造する）方法である。生物進化の過程と同様に、遺伝的アルゴリズムにおいても、より適応度の高い解が選択され、より多くの「子孫」を残すことになる。その際に親である解の「DNA」が混合されたり突然変異を起こしたりすることによって、次世代において新しい解が生み出される。

出力と教師信号との誤差を計算し、それに基いてシナプスの重み付けを変化させるというニューラル・ネットワークの学習方法も、試行錯誤によってわずかな変化を蓄積することでより良い解（重み付け）を探索するという意味で、進化的過程の一種と考えることが出来る。さらにはより明示的な遺伝的アルゴリズムを採用したニューラル・ネットワークもある。そこではシナプスの重み付けではなく、ネットワークの構造自体が遺伝的アルゴリズムによって決定されている(守・都築・楠見, 2001, 2.4)。これはデネットの言う (2) のタイプの進化的過程と見なせるかもしれない。遺伝的アルゴリズムはまた、複数のエージェントの間での記号によるコミュニケーションの進化のシミュレーションにも利用され、認知言語学や進化言語学の重要な道具になっている(Grim, 2007; MacLennan, 2007)。

遺伝的アルゴリズムは人工知能や認知科学の分野だけでなく、人工生命の分野においても、生物進化のメカニズムの「論理的な形式を抽出する」ために利用されている(Langton, 1996)。例えばドーキンスは遺伝的アルゴリズムに似た方法でコンピューターによる簡単な形態発生のプログラムを作り、多様な「生物」（「バイオモルフ」）たちを誕生させている (Dawkins, 2006, chap. 3)。

ドーキンスは多様なバイオモルフを生み出すプログラムの「創造性」（あるいは$\boldsymbol{Z}^9$の探索空間に存在するバイオモルフたちを見つけ出す「探索能力」）を、方向付けられた小さな変

化を累積する力（すなわち遺伝的アルゴリズム）に帰している。ランダムで小さな変化が非常に複雑な形態や機能を生むのは、それが何らかの基準によって方向付けられ、そしてその変化が世代を経て累積していくからである。この方向付けは生物進化という局面では生存に適さないものは取り除かれるという仕方で働く。

ドーキンスはさらにこのようなメカニズムが、何らかの複雑さを持ったあらゆるシステムの発展のメカニズムとして有効であることを主張する。彼は、創造とはビルディング・ブロックとして与えられた素材を有益な仕方で組み合わせることに他ならない、と考えている。そしてその意味で生態系、人間の知能、そして人工知能は連続している。その組み合わせの仕方は、数学的な意味では、あらかじめ決められたものであるが、その組み合わせの数が膨大になるとき、適切な組み合わせを「発見」することは「創造」と同義語になる、とドーキンスは考える。彼は次のように述べる。

> コンピューター・モデルにおける人為淘汰であれ、現実の外部世界における自然淘汰であれ、累積淘汰というものは効率の良い探索の手順であり、それによってもたらされる結果はまさしく創造的な知性であるかのように見える。（Dawkins, 2004, 邦訳, p. 121）

ドーキンスに従えば、生物進化のメカニズムも，そして人間の認識や知識――社会的なものはもちろん、個人的なものまで――の発展もまた、何らかの形の遺伝的アルゴリズムによって説明されることになる。従ってそのメカニズムを模倣することによって、生物や知性のモデル化が可能になるのである。

6. 結び

従来の方法による AI では真に人間の知性をシミュレートすることはできないという見解が一般的になると、知性のモデルについての考え方が改められ、それに伴い AI の設計の思想も大幅に見直されるようになった。望ましい振る舞いを得るためにプログラムを複雑化、洗練させていくことよりも、複数のシステム間の相互作用やフィードバックによって、比較的単純なプログラムによって制御されたシステムから複雑な振る舞いを生み出すメカニズムを研究していくことに焦点が置かれるようになってきたのである。この傾向は、脳の構造とそのメカニズムをコンピュータでモデル化することを試みるコネクショニズムにおいて、とりわけ顕著である。

このような発想は伝統的に人工生命において重視されてきた方法論であった。生命を理解しそれを創造するためには、生命を創造するメカニズムの本質を取り出す必要がある―

―これが人工生命の当初からのスローガンであった。そして今、人工知能もまた、知性的振る舞いを生み出すメカニズムに注目している。そればかりではなく、認知科学や進化生物学の領域における成果は、生命と知性はどちらも進化のメカニズムによって生み出される、という認識を浸透させている。

本稿は特に記号的 AI とコネクショニズムを取り上げ、近年の AI における人工生命への接近というテーマを論じた。しかしこのテーマを扱う際には、近年のロボティクスの成果や、そこから生じる記号的表象の必要性についての議論に触れることが適切であったろう。これらについては今後の課題としたい。

註

1 「人工知能」という言葉は二つの意味に使われる。一つは、知能をシミュレートする人工的なシステムという意味であり、もう一つは、そのようなシステムの開発・研究という意味である。「人工生命」についても同様。

2 当時の工作技術の限界から、この機械が実際に製作されることはなかった。

3 「実効的に計算可能」とは、決まったアルゴリズムに従って解が求められる、ということである。

4 これに対して、例えば谷は力学系に基づく認知のメカニズムを利用したモデルにおいては記号的表象は不要であるという立場を取り、従ってそもそも記号接地問題は生じないという（谷, 2005; Cf. 中村, 2004; 2006)。一方で、アンディ・クラークやスティーヴ・グランドなどは、やはり力学的なアプローチを支持するが、それでも表象の重要性を主張する（Clark, 1998; Grand, 2005）。

5 ドーキンスの造語。ある文化において伝播、継承される情報や概念、制度など。

文献

Churchland, Paul M. (1996). *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*, Bradford Books. （1997, 信原幸弘・宮島昭二訳, 『認知哲学――脳科学から心の哲学へ』, 産業図書.）

Clark, Andy (1998). *Being There: Putting Brain, Body, and World Together Again*, The MIT Press / A Bradford Book.

Dawkins, Richard (2006). *Blind Watchmaker*, Penguin Books Ltd. （2004, 日高敏隆監修、中島康裕・遠藤彰・遠藤知二・疋田努訳、『盲目の時計職人――自然淘汰は偶然か？』、早川書房.）

Dennett, Daniel C. (1990). 'Evolution of Consciousness', in Brockman (Ed.), *Speculations: The Reality Club*, Prentice Hall Trade, 85-108.（1992, 斉藤健二訳, 「意識の進化とコンピュータの進化」, 長尾力他訳, 『〈意識〉の進化論』, 青土社, pp.33-63.）.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Simth, A., Parisi, D. & Plunkett, K. (1997). *Rethinking Innateness: A Connectionist Perspective on Development*, Bradford Books.（1998, 乾・今井・山下訳, 『認知発達と生得性――心はどこから来るのか』, 共立出版.）

Feigenbaum, Edward A. & McCorduck, Pamela (1984). *The 5th generation*, M Joseph.（1983, 木村茂訳, 『第五世代コンピュータ――日本の挑戦』, TBS ブリタニカ.）

Gillies, Donald (1996). *Artificial Intelligence and Scientific Method*, Oxford University Press.

Grand, Steve (2004). *Growing up with Lucy: How to Build an Android in Twenty Easy Steps*, Orion.（2005, 高橋則明訳, 『アンドロイドの脳――人工知能ロボット“ルーシー”を誕生させるまでの簡単な 20 のステップ』, アスペクト.）

Grim, Patrick & Kokalis, Trina (2007). 'Environmental Variability and the Emergence of Meaning: Simulational Studies Across Imitation, Genetic Algorithms, and Neural Networks', in Loula et. al. (2007), 284-325.

Grim, Patrick, Mar, Gary & St. Denis, Paul (1998). *The Philosophical Computer: Exploratory Essays in Philosophical*

Computer Modeling, The MIT Press / A Bradford Book.
Harnad, S. (1990). 'The Symbol Grounding Problem', *Physica D*, *42*, 335-346.
久木田水生 (2006).「脳のカオス的モデルと心の哲学」,『実証段階におけるカオス研究の哲学的考察』, 平成 16－17 年度科学研究費補助金（基盤研究 C）研究成果報告書, 93-100, 掲載予定. http://www.geocities.jp/minao_kukita/files/chaos.pdf
Langton, C. G. (1996). 'Artificial Life', in Boden (Ed.) *The Philosophy of Artificial Life*, Oxford University Press, 39-94.
Loula, Angelo, Gudwin, Ricardo & Queiroz, João (Eds.) (2007). *Artificial Cognition Systems*, IDEA Group Publishing.
MacLennan, Bruce (2007). 'Making Meaning in Computers: Synthetic Ethology Revisited', in Loula et. al. (2007), 252-283.
守一雄・都築誉史・楠見孝編著, (2001). 『コネクショニストモデルと心理学――脳のシミュレーションによる心の理解』, 北大路書房.
中村雅之 (2004).「表象なき認知」,『シリーズ　心の哲学 II――ロボット篇』, 信原幸弘編, 勁草書房.
中村雅之 (2006).「力学系理論に基く身体性認知科学」,『科学基礎論研究』, 第 105 号, 11-17.
Rheingold, Howard (1986). *Tools for Thought: The History and Future of Mind-Expanding Technology*, Prentice Hall Computer Pub.（2006, 日暮雅通訳,『新・思考のための道具 : 知性を拡張するためのテクノロジー――その歴史と未来』, 日暮雅通訳, パーソナルメディア.）
Searle, J. R. (1980). 'Minds, Brains and Programs', *Behavioral and Brain Sciences, 1*, 417-424.
谷淳 (2005).「認知力学系とロボティクス」,『岩波講座　ロボット学 6　ロボットフロンティア』, 岩波書店, 127-155.
月本洋 (2002).『ロボットのこころ――想像力を持つロボットを目指して』, 森北出版.
津田一郎 (1990).『カオス的脳観――脳の新しいモデルを目指して』, サイエンス叢書 24, サイエンス社.
津田一郎 (2002).『ダイナミックな脳――カオス的解釈』, 双書　科学／技術のゆくえ, 岩波書店.

〔龍谷大学非常勤講師〕