

生命システム論を基にする心的状態と知能の説明

坂田純

2018年7月

JCSS-TR-78

[連絡先]

〒680-0462 鳥取県八頭郡八頭町福本49-14

js-dpr21@hal.ne.jp

© Jun Sakata, 2018

日本認知科学会

事務局

〒214-8571 神奈川県川崎市多摩区東三田1-1-1

明治大学 理工学部 電気電子生命学科内

jcss@jcss.gr.jp

生命システム論を基にする心的状態と知能の説明

要約

本論文において、生命システム論の一つとして形成されたオートポイエーシス論を基に、表象（知覚）や認識等の心的状態、そして知能の成立について試論する。現在、認識や知能を統合して理解可能にする理論が望まれているが、これらの事象を体系的に説明する理論の構築は困難な状況である。このような状況においては、大胆な仮説の設定と一つの思考枠組みによる一貫した体系化を試みるのが、有用であると考えられる。本論文では、一般システム論としても期待されているオートポイエーシス論を基軸に定め、表象や認識と、知能の関与する学習や知能行動に対する一貫した説明を試みる。

キーワード：知能行動, 欲求行動, 表象と認識, オートポイエーシス

1 はじめに

現在、人工知能（AI）分野等において、計算機に対し、人間と類似の形式で問題を処理する（つまり「思考」する）ことが期待されるようになってきている。実際には、限定された処理においては人間と同等かそれ以上の性能を持つのであって、AIと人間の思考が同質なわけではない。動物や人間において、表象、認識、情動等の言葉で表現される状態が成立しているのは疑い得ないが、現存のAIにこのような心的状態が成立しているとは考え難い。文字通りの人工の知能を生成するには、これら心的状態を人工的に生成する必要があると考えられる。そして心的状態や知能の生成を目指すにはまず、心的状態や知能とはどのようなものでどのような機構により立ち現れるかを、明らかにする必要がある。しかしこのことこそが歴史的な難問である。

心的状態の成立について、いくつもの思想が存在しており、それらの思想を分類して大まかにまとめることは可能である。本論文では、生命システム論の一つとして形成されたオートポイエーシス論を基に、表象や認識等の心的状態、そして知能の成立について試論する。実験的事実の蓄積のあとには、それらの知識としての体系化が必要である。この論文は、蓄積された知見の統一的体系化の一助となるよう、オートポイエーシス論による一貫した説明を試みる。

2 この論文での議論の展開の仕方

現在、認識や知能を統合して理解可能にする理論が望まれているが、その基盤となるべき生命現象と心的事象に関わる議論は多岐にわたり、これらの事象を体系的に説明する理論の構築は困難な状況となっている。この困難な状況により、知覚や認識そして知能に関する体系化された理論の構築は、当初期待されたほどには進んでいない。困難な状況を招いた要因の一つは、不確かな仮定による議論の展開であるが、それをわかったうえでなお、いくつもの仮説とその検証により理論的統合を進めざるを得ない。本論文でもいくつかの仮定を導入し、一つの理論を基軸に定めて議論を展開する。ただし、脳科学等ではほとんど参照されることのない、そ

れでいて重要な知見を提示し、それら知見を説明可能なように理論化を試みる。

2.1 基盤とする理論

生命現象の延長上での心的事象（一般に“認知”の言葉で示されることが多い）に関する議論は、脳科学を中心に行われている。生物学はまず観察、実験可能な事柄を積み重ね、それらの事象から生命の機構を考察するため、生命現象を作動させ続ける有機的機構の解明はどうしても観察より遅れることになる。さらに生命現象の自律性や一回性より、観察により事実とみなされている事項も、完全に確定された事実というわけではない。これらのことから現在の認知に関する理論は、まだ発展途上と言える。

生命現象を基にした認知理論は脳科学以外にも複数存在する。心理学、精神医学、神経生理学などの一見あまり共通性のないと思われる分野において、類似の認知理論が存在している。本論文はこれらの理論を参照して議論を展開していく。

多数の階層にまたがる事象であっても、できるだけ少数の原理で説明できるに越したことはない。しかし無理な単一化による原理を全ての事象にあてがうのでは、現実の事態に不適合な説明を与えることになる。一般に、階層的に立ち現れている事象を説明するには、各階層ごとに現れる原理の抽出を試みるのであるが、身体を規定する物理的制約が行動においても制約であるように、階層間においてもなお成立する原理（個々の階層の原理よりも抽象化されるであろう）、もしくは階層形成そのものの原理を描き出すことは可能であろう。本論文では、階層的なシステム全体に対し、できるだけ一般的な原理による記述を目指す。そして階層形成の原理としてオートポイエーシス論を利用する。この理論は、神経生理学者のマトゥラーナ・ヴァレラが生命現象の機構として考案した理論を、異なる分野でも適用できるよう様々に拡張したものである。オートポイエーシス論にも類似の理論が存在するが、非線形科学の延長上として議論が展開可能なことと、作動の機構の図示が試みられていることから、本論文ではオートポイエーシス論を主に利用する。

2.2 事象を説明すること

科学一般において、変数を用いて状態遷移を数式で表現可能としたときに、その状態遷移に関連する事柄を説明できたとみなす場合が多々見られる。しかし、例えば酸素呼吸における ATP 産生の機構において、化学式の複雑な連鎖を用い、最終的に細胞膜内外のプロトン濃度差で状態遷移を記述できたとして、そのことをもって ATP 産生を説明できたと判断する人間はほとんどいないであろう。実際には、プロトンチャンネルを通しての、膜内外のプロトン濃度差による浸透圧に従う、プロトン流入の力学的作用による膜タンパク質の構造変化が、ATP 合成の触媒反応を可能にすると考えられている。一般の生物学分野では、数式や化学式の提示は言語記述による説明の補助的なものである場合が多いが、非線形科学や情報科学を基にする生物学分野では、状態遷移の間に見られる量的な関係を数式で記述することをもって事象の説明とする場合が多い。数式で量的関係を記述できれば、その数式に沿う状況を作り上げることで、事象の再現が可能となる場合が多いため、機構は明らかでないが事象そのものは復元可能という事態も起こり得る。

本論文は心的事象の説明を目指すため、変数やパラメータを用いた記述でもって事象を説明する、という立場はとらない。しかし、議論の展開に非線形科学等の助けが有用なため、システムの生成において変数がどのように決定されるかを、議論の端緒とする。また数式の代わりにできるだけ図示による説明を試みる。

2.3 心的事象の生起の二分類

心的事象が身体という物理的構造物の作動の結果として生じることは確かである。心的事象の生起について、次の二つの考え方以外はみることがない。

1. 精神の世界が実在し、身体に精神が宿ることによる。
2. 身体の作動において、存在の位相を変換して物理的事象から心的事象が立ち現れる。

本論文は上記2の立場をとる。上記2の立場を受け入れれば、表象や意識を、「神経系全体の作動の反映（によって立ち現れた状態）」とみなすことができる^{*1}。そして物理的実体を持たない表象や意識をも構成素として認めることで、非実体的な抽象的要素によるシステムとして、言語現象等の様々な認知に関わる現象をモデル化可能と期待される。

2.4 対象とする現象

本論文では細胞システムから始め、表象や認識等の心的状態の議論を経て、知能について言及する。主に動物学の知見を用いて、知能行動の成立条件を示す。いずれの現象においても、オートポイエーシス論を用いて体系化を試みる。

まず3章でオートポイエーシス論の概要を示す。4章で、動物行動学における知能行動の理解の仕方を説明する。そして5章において、動物行動学で重要な概念である本能行動、学習、知能行動について、オートポイエーシス論による解釈を試みる。これらの概念の説明においては、できるだけ、物理的構造物である身体とそこに還元できない心的状態との関係を考慮して、議論を進める。最後6章で課題と応用を簡単にまとめる。

3 システム論（オートポイエーシス）

オートポイエーシス論で本論文に必要な部分だけを取り出して記述する^{*2}。以下、本論文において重要な考え方を箇条書きで示す。具体的説明は次節から各システム論の説明において行う。

1. 生成プロセスの連鎖の閉域が形成されたとき、システムが成立する。
2. システムの本体は生成プロセスの閉域であり、観察されるものはその構造である。
3. システムは構成素の産出を通じて、環境と自己の境界を区切り続ける。
4. 複数のシステムが新たな構成素を設定して作動するとき、この事態をシステムの構造的カップリングと呼び、新たな構成素による新たなシステムが成立する。
5. それまでのシステムの作動の反映を構成素とするとき、新たな構成素によるシステムを自己言及システムと呼ぶ。

^{*1} 脳科学でもこの立場を取る場合が多い

^{*2} 本論文のオートポイエーシス論の理解は、[1]と[2]の著作を元にした。

3.1 生命システム

3.1.1 細胞システム

オートポイエーシス論は神経生理学的知見を基に生命現象の理論として考案されたので、生命システム論がオートポイエーシス論の基盤となる。また非線形科学との接続も可能である。

まず図 1 に、非線形科学における触媒サイクルを示す^{*3}。この図はオートポイエーシス論のものではないが、オートポイエーシスシステムとみなして議論を展開することも可能である。細胞システムの場合は、 E_t と S_t は有機化合物であり、 E_n が触媒となって基体 S_n が変化し、次の触媒 E_{n+1} が産出される。この関係が順次触媒 E_t と基体 S_t で成り立ち、始まりの触媒へと回帰するとき、自己触媒による産出連環の閉域が形成される。図 1 を元に、細胞システムをオートポイエーシスシステムとしてみると以下のように記述できる。

- 細胞システム本体はあくまで E_t を生成するプロセス（図 1 の矢印）の閉域である。プロセスの閉域が形成され続ける限り細胞システムは成立する。プロセスの閉域を形成する E_t がシステムの構成素にあたる。
- 細胞膜などの細胞構造体は生成産物の排出の結果としてある。生成プロセスを細胞システムの本体とみなすと、細胞膜はプロセスの閉域を可能にする環境の一部であるとみなせる。生命システムは自己を維持する環境の一部を、排出物の総体による構造（図 1 の円）として、産み出し続ける。
- 細胞システムの本体は生成プロセスの閉域であるため、システムには構造（細胞膜など）の内部と外部の違いの区別は存在せず、構造の内部と外部のどちらも環境に属する。基体 S_t は環境に属し、環境の変動により S_t の種類が変われば構成素 E_t の種類が変わり、結果として構造も変わり得る。逆にシステムの自律性で E_t が変わっても構造が変わり、構造外部の環境との関係も変わり得る。細胞システムと環境に入力と出力の関係を問うことはできないが、細胞システムと環境は基体と構造を介して相互に浸透する。

生命システムの機構の解明が難しい理由の一つは、プロセスの循環が必ずしも始まりの構成素そのものへと回帰しなくても、生成プロセスが廻り続けて、その間は生命が存続することである。酸素呼吸におけるグルコース代謝の経路のように、関与する酵素と生成産物が固定的な場合もあるが、細胞形態や生体内機能等は構成素の変化等により大きく変異する。実際の細胞では余剰といえる多重の代謝経路が存在しており、環境との相互作用により細胞形態や細胞機能が変化し続けるにもかかわらず、生成プロセスが一貫して回り続ける限り同一の細胞である。したがって重要になるのは構成素の再帰的循環そのものよりも、類似の構成素において同型の産出プロセスが廻り続けることである。このような同型の産出プロセスが円環を成して循環し続けることは、冰山から流れ落ちる水流が偶然その経路に跡を付けることにより川の流れを成すことに、類比される [3]。一度ある型の産出プロセスが作動したとき、その産出プロセスを作動可能にした環境が一時的にでも存在したはずであり、産出物がその環境構築に関与することで、産出プロセスの閉域形成と共にその同型性を担保する環境が成立し続けることになる。細胞システムでは、細胞膜や細胞骨格が、産出プロセスの同型性を担う、構築された環境である。オートポイエーシス論では、構成素の型と産出の順序を決める規則をシステムの「コード」と呼ぶ。始原の生命が産まれた後は、親細胞から引き継がれた細胞内物質と細胞構造を持ち、そして化学物質は化学変化の、細胞構造は力学的、電気的な規則を持つ。これらがコードとなって有機化合物の産出プロ

^{*3} 著書 [1] の 131 ページの図を参考にした。

セスが同型性を保って循環する。

一般に生命システムの作動を、変数の値によって状態が決定される関数として記述しようとする場合、変数は特定の物質の濃度等で記述される。しかし構成素となる物質の種類が閉域形成がなされるまで決定されず、閉域が形成された後に異なる物質に変更される場合もあるので、関数を記述するための変数を不定とせざるを得ない [1]。また物質の濃度等が状態の決定要因となり得るのは、すでに成立し、かつ同型性を保ちながら変化し得る細胞構造が、化学変化の規則と物理的規則の元に、物質に対する挙動を決定するからである。以上のことにより、生命現象を関数によって記述することが困難となっている。

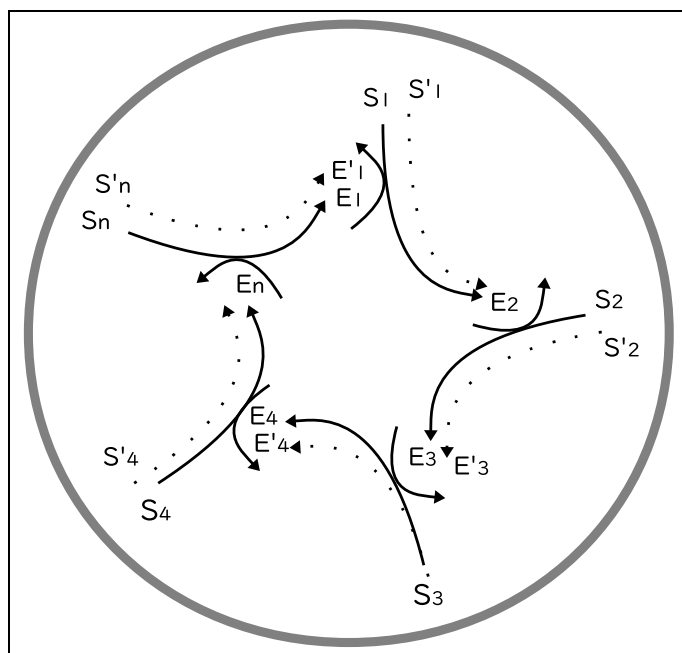


図1 自己触媒サイクル

3.1.2 身体（多細胞生命体）システム

まず、非線形科学におけるシステムの階層形成を示す、自己触媒サイクルによるハイパーサイクルを図2に示す。I₃を例にとると、I₃は自己複製を行うと共にG₁とE₃を生成する。I₃はGおよびNのサイクルの構成素であると共にIおよびEのサイクルの構成素である^{*4}。一つの構成素が複数階層で共有されることで階層性が保たれている。しかし図2では階層間で同質の構成素を共有する場合に限定されるので、オートポイエーシスではこの図式はとっていない。各階層の複数システムが新たな構成素を設定して、その閉域による新たなシステムが作動し、そのシステムが一段階上の階層を構成すると考える。このようなオートポイエーシスシステム間の新たなシステム形成を「構造的カップリング」と呼ぶ。図3(a)にオートポイエーシスによる多重的システムの概念図を示す^{*5}。生命システムでは例えば、図3のAからFが細胞システムで、G、Hが細胞システムのカップリングによる器官システム、Iが器官システムのカップリングによる身体システムとみなすこと

^{*4} 構成素に自己触媒が含まれているのは、自己を触媒することで、早急に進行して消失してしまう反応を遅延したり、逆に加速せたりすることによって、安定してサイクルを循環し続けるためである。

^{*5} 著書 [4] の 368 ページの図を参考にした。

ができる^{*6}。図3 (a) はあまりに煩雑なため、以後、プロセスのみを図示した (b) の形式で記述する。低階層のシステムと高階層のシステム共に自律性を持ち、一度システムが成立してしまえば、お互いに環境として相互に浸透する（相互浸透）と考えられている。ただし低階層のシステムが完全に瓦解してしまえば、高階層のシステムは存続が不能となるため、高階層システムは低階層システムに依存している。また各階層の自律性は、システムによって程度が異なると考えられる。

多細胞生命体での細胞間の協同は、細胞形態の同質性や、隣接などの空間的な布置などにより保たれており、これらの性質や布置が、細胞間伝達物質等による細胞間関係の構築を可能にしている。そこで細胞システムのカップリングで生成する身体システムにおいて、細胞の同質性や空間的な布置による細胞間関係が構成素となって閉域形成が成立していると仮定する。受精卵からの発生に伴う、細胞分裂や細胞の成長という、時間的、歴史的な身体形成により、一つの生命体の同一性が担われている。さらに個体発生は、種の系統進化で得られた歴史的経緯を反映しており、発生過程に従うことで外部環境に適応した身体が構築される。身体という多細胞間の高度な協調による構造体は、身体システムの構成素の閉域が形成され続けなければ崩壊してしまう（死とその後の腐敗）が、個体発生とその後の成長が系統発生を反映しているおかげで、閉域形成の持続性が高度に高まっている。

定型化された個体発生を可能にしているのが遺伝子であり、遺伝子の働き方をさらに指定しているのがエピジェネティクスと呼ばれる分子機構である。遺伝子とその発現機構が決められているのはタンパク質の種類であり、遺伝子発現の順序を指定してタンパク質の生成順序を制御するような、塩基配列以外の分子機構がエピジェネティクスである。エピジェネティクスの機構の解明は現在進行中であり、詳細はまだわかっていないが、卵細胞由来のタンパク質やRNA、発生に従って行われる塩基への修飾が、身体の順序だった構築に大きく貢献していると考えられている。エピジェネティクスの機構は非常に複雑で、発生を厳格に縛っているのではなく、発生や成長段階での環境への適応を担っているとも考えられている。オートポイエーシス論での器官や身体は、元となるシステム（器官システムであれば細胞システム、身体システムであれば器官システム）間の関係（細胞の性質や分布）が構成素となることで成立するとした。遺伝子発現やエピジェネティクスは、細胞の形態や機能を規定する（一意に決定するわけではない）ことで、コードとして細胞間の関係の順序だった変遷を規定していると考えられる。ただしこれらのコードは細胞の自律性や外部環境の変動によって変化し続けており、身体の同型性と個体の唯一性をもたらしている。

3.2 表象システム

我々の知覚において、音や像などの形式における現れを、本論文では「表象」とする。知覚器を持つ生命体には表象が成立する可能性があるが、神経系による知覚と運動の連鎖を持つ動物に限定して議論を進める。神経系を持つ動物において、神経ネットワークの状態と表象の成立に、何かしらの対応関係が存在することは疑い得ない。しかし、神経ネットワークという物理的基体の状態がどのように表象と連関するのかはほとんど明らかになっていない。人間の思考の形式においては、記述の形式そのものを変えるような変換を理解することができない可能性が高いためである。オートポイエーシス論による表象の生成の説明も、身体や神経系の作動において、物理的事象に相関して心的事象が立ち現れる、とする立場を取らざるを得ない。なお神経系も身体の一部であるが、神経系を除く知覚器や作動器およびそれらの支持構造体をまとめて、本論文では身体と表現する。

^{*6} 間に組織等を挿入することで、より高階層のシステムを想定することも可能であり、逆に器官を削除して細胞システムのカップリングによる身体システムを考えることも可能である。

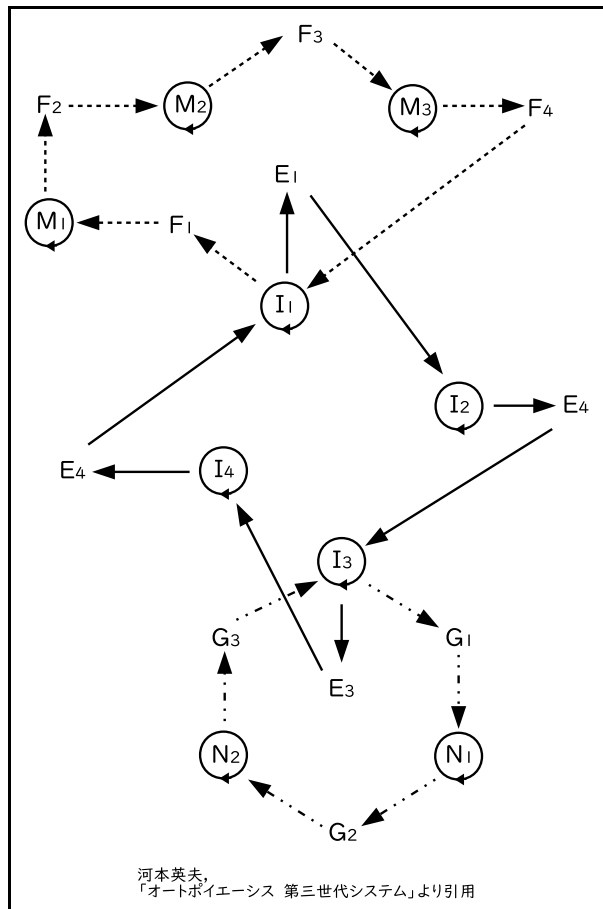


図2 ハイパーサイクル

議論の対象とする動物を、複数種の知覚器と、末梢神経系と中枢神経系の区別をつけられる種とする。中枢神経系と末梢神経系の区別は、機能と形態の違いに起因し、神経システムの機構を説明するための便宜的な区別という側面を持つ。神経系の状態は、3次元空間中の電位の分布状態や神経細胞の発火の分布等で考えることができるだろう。神経系の状態は、知覚器からの神経インパルス列を受けて直接に変化するため、神経系全体の状態そのものが身体の作動の反映とみなせる。同時に、神経系全体の状態を反映して身体は作動する。中枢神経系と末梢神経系の区別を持つような動物種においては、個体全体は卵細胞から発生するので、神経系の作動は個体発生における身体と神経系の発達に相関する。神経領域内の区分は最初から存在するのではなく、身体発生に伴い各神経領域が自他の領域の区分を区切り続けることで、区切られた個別領域として外部の観察者に認識可能となると考えられる。

発生に伴い神経領域も分化していくので、発生初期には個別の表象は存在せず、原表象とも呼ぶべき未分化の表象の状態であるだろう。変化し続ける身体の状態と、自他の区分を区切り続ける個々の神経領域の状態の間に、対応関係の閉域形成が成立したとき、個々の表象が成立すると仮定する^{*7}。このとき、ある種類の刺激（例えば空気振動の波や特定の波長の電磁波）とその刺激を受容可能な受容器（耳や目）と神経領域の受容野

^{*7} この表象の成立は、身体の発生と発達を、オートポイエーシス論に組み込むために必要となった定義である。一般的な表象の成立の考え方と若干の違いはあるが、神経系の状態と相関して表象が成立することを前提とする点は、一般的な考え方と同じである

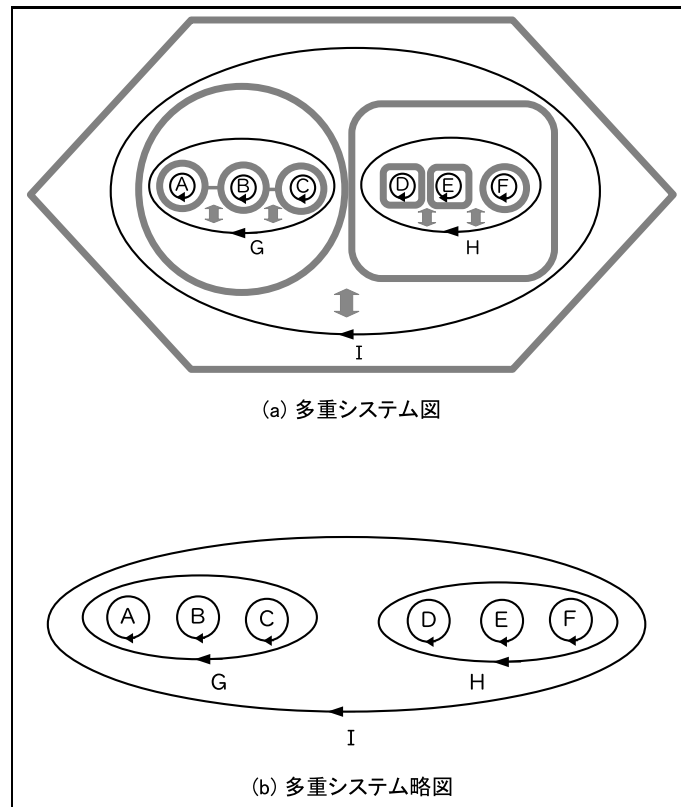


図3 多重的オートポイエーシスシステム

(聴覚野と視覚野)が対応付けられているので、特定の外部刺激に特定の表象(音や視覚像)が生じる。神経領域全体が分化して個々の神経領域に区切られたのであるから、それぞれの神経領域は周りの神経領域との関わりの中で、特定の受容野として機能を受け持つことができる*8。ある身体の状態に対する神経領域の応答を構成素とするシステムの構造が、オートポイエーシス論による表象であると定義する。ここでオートポイエーシス論で使用される「自己言及」の言葉を導入する。自己言及の定義は、「オートポイエーシスシステムのそれまでの作動が、新たな構成素の産出に際し、攪乱を与え、その構成素の状態がシステムそのものの作動を反映すること」である [2]。身体の一部である神経系が身体全体の作動を反映しているのので、表象は、自己自身を自己に反映することで成立する自己言及システムとみなせる。神経領域は他の領域から神経接続を受けており、身体の状態を反映して神経インパルス列が流れ込み続けるので、ある神経領域の状態はその時刻の自己の状態に周囲の興奮を引き受けることで、次の自己の状態を産出する(産出基礎付け)。

表象システムの模式図を図4に示す。細胞システムの図1では、構成素 E_n は E_{n-1} の触媒作用により基体を変形させられた、 E_{n-1} とは異なる構成素であった。表象システムにおける構成素は、ある身体の状態に対する神経領域の応答であるので、時間経過に伴うその領域の状態が遷移していく。そのため図4では、構成素 E_n は $t = n$ におけるある神経領域の状態を示し、この神経領域の状態の遷移による構造が、対応する個々の表象を示している。細胞システムでは、構成素である化合物は触媒であるとともに細胞構造の一部になり、結果

*8 この考え方はいわゆる全体論に基づいている。全体論も個々の知覚野や運動野が局所化されていることを認めている。全体論と局所論の違いはむしろ、生体の側の自律性を認めるか否か、各神経領域が神経系全体の中で機能を分担するかそれとも単独で機能を持つかである [5]。

として構成素の総体が、細胞システムの構造である細胞構造体を作り出し維持し続ける。表象システムにおいては、構成素 E_t にはそれまでの E の状態が全て畳み込まれているので、それぞれの構成素 E_t 自体がそれまでの E の総体とみなせる。一度閉域が形成され維持される限りで、それぞれの構成素 E_t にそのときの表象が対応するが、それまでの E の総体としてあるので、表象を神経領域の状態の遷移による構造と定義している。

コードは構成素の型と産出順序を規定するので、表象システムのコードは、身体の状態を自己の状態として反映する神経系の性質そのものと、神経系の状態を引き受けながらも自律的に作動する身体による攪乱である。 D_n は $t = n$ におけるコードを成す攪乱であり、神経系の自己言及性および自律性と身体の状態変化である^{*9}。個々の神経領域は神経細胞の集合であるため、神経細胞の自律性により、その神経領域全体の応答特性も自律的に変化する。神経領域が明確な分画を持たない連続的な構造の場合（大脳皮質の感覚野など）は、応答する領域の分画が変遷可能であり、分画がある程度一定して保たれることが個別の表象の成立条件である。また個々の神経領域は複数の神経領域との接続を持つ場合がほとんどであり、知覚野の場合は知覚器の興奮に選択的に応答することで、特定の刺激に対する知覚野として存立する。 E_0 はあらかじめ決めておくことはできず、 E_0 のときの身体の状態が再び現れた際に、 E_0 もまた再現されて閉域が形成されたとき、はじめて E_0 が確定する。細胞システムの場合と同様に、次の構成素が産出されて産出プロセスが継続されることが重要なので、構成素 E は閉域を維持可能な限りで変異が可能である。細胞システムや身体システムは、閉域形成が途絶えた時点でシステムが消滅すること（死を意味する）がほとんどであるが、表象システムは身体システムが成立している限りで再生可能である。

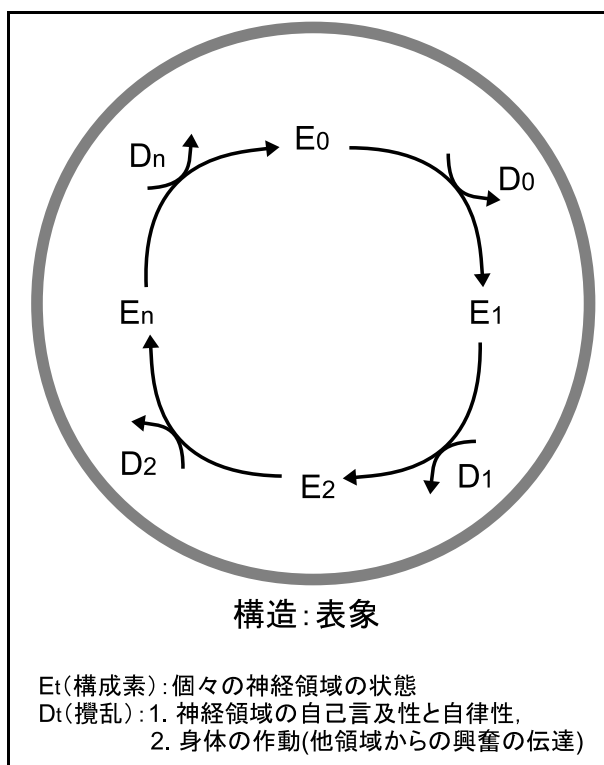


図4 表象システム

^{*9} 身体の一部である神経系が身体全体の状態を反映すること、神経系の接続が複雑な自己ループを形成していることから、神経系は二重の自己言及性を持つ。

3.3 社会システム

人間の活動において社会が成立、維持されていることは間違いないので、人間を構成素とした社会システムをまず想起することができる。しかしオートポイエーシスシステムにおいては、産出される構成素自体の自律性は低いはずである。人間と社会の関係を考えると、個々の人間の社会からの自律性は、身体システムに対する細胞の関係と比べ、非常に大きいと考えられる。そこでオートポイエーシス論での社会システムは、人間の間で行われるコミュニケーションを構成素として閉域を形成する高階層のシステムと考えられている。この場合、人間と社会の双方にとって、お互いが環境として浸透している（相互浸透）と考えられる。コミュニケーションの連鎖の閉域が社会システムとすると、社会は村落レベルであり、実際の社会は多重のオートポイエーシスシステムと考えられる。一度成立した社会システムは自律性を持つので、その自律性を通して次の状態を再生産し続けるが、環境との相互浸透を通して異なる状態を生成することもあり得る。例えば、原始社会は再生産の結果が同じ状態に収束するように閉域形成されたシステムで、近代は再生産の結果で生じた差異をより促進するように閉域が変化したとみなすことも可能である。言語体系とそれを習得する個々の人間の関係は、社会システムと個々の人間の関係と類比することができる。

4 知能と言語の動物学での議論

そもそも「知能」を定義すること自体が困難であるが、生得的な行動だけでは理解できない、獲得された行動や洞察による「知能行動」が、動物に存在することは確かである。次章からの知能のオートポイエーシス論的議論のために、まず動物学における知能の理論をここで簡単にまとめておく。

動物行動は身体の状態を反映する神経系と、神経系の状態をさらに反映して作動する身体、共同的作動で成立していることに疑いはない。同時に、脊椎動物のようなある程度神経系の発達した動物では、各行動において、知覚による表象の成立がそれらの行動を導いていることも、疑うことが難しい。人間の知覚と他の動物の知覚は異なるであろうが、動物においても、音や像として表象が現れ、快や不快のような情動が現れているはずである。何かを「感じる」ということは、物理的過程で説明することが極めて困難であり、せいぜい神経系の状態と表象や感情のような心的状態の間に密接な関連性があると指摘できるだけである。動物学や心理学において、このような考え方を「平行論」と呼ぶ。

動物は系統発生による知覚器、作動器、神経系を持ち、各動物種ごとに適応的な形で外部世界を知覚している。動物行動は多様であるが、脊髄反射の場合は知覚器と作動器の間にただ一つの介在神経を介した、求心性感覚線維と遠心性運動線維による電位伝達で、表象を持ち出すまでもなく行動を説明できる。しかし逃避や摂食行動など多くの行動が、その作動において、視覚、聴覚、嗅覚などの頭部受容器から中枢神経系を経て、脊髄を下降する運動神経による作動器への興奮伝達を必要とする。知覚器と作動器の間に複雑な神経接続を持つため、反射行動だけで動物行動を説明することが難しい。例えば餌を取るために迂回する原初的な知能行動でさえ、反射だけでは到達不能である。神経系の発達による行動の遅延と反射の抑制により、餌を求めて迂回するという行動を、一連の行動連環の中に差し挟むことができる。上記の迂回行動のように、動物行動の多くは目標志向的行動とみなすことができ、このことは表象の成立のみならず、「目標を志向する」という擬人的な状態が、遠心性神経線維の始点（人間であれば脳の運動皮質）を物理的に興奮させることを意味している。したがって動物行動は、神経系という物理的基体に平行する表象の成立と、その逆に表象や志向性のような心的状態による物理的基体への影響の、二つの難問によって初めて可能になる。残念ながらこの二つの難問を解

決した理論は存在せず、そしてこの論文でも前提条件とせざるを得ない。ただしこれらを認めることで、動物行動から人間の言語までを同一の論理で記述が可能と期待される。

動物の知能行動は行動の可塑性によるところが大きい。動物行動学での行動の可塑性に関する議論を先に導入する。動物行動は、種に固定的な刺激応答事象である「本能行動」と、後天的に可変である「欲求行動（目標志向的行動）」の連環として成り立つ [6, 7]。コンラート・ローレンツによると一連の目標志向的行動は、その目標となる本能行動の遂行により完了する。目標となる本能行動の遂行を妨げる障害に直面したときに、可能だが先験的には決定されていない行動を洞察により選択することを、知能行動と呼ぶ。ただし学習や欲求行動そのものも知能でありうる。本能行動は種に固定的であり、後天的に可変である学習や知能行動は、欲求行動の一部である。本能行動と欲求行動共に神経生理学的な機構はほとんど明らかになっていない。

動物における高位の知能行動に、類人猿の道具使用が存在する。その道具使用に対する行動観察の記述から、類人猿では人間に類似する感情と思考が成立していると考えられる。しかしながら、類人猿が箱や棒を道具として使用できるのは、目標となるバナナと道具となる箱や棒が同時に視野に入るようなときのみであり、箱や棒をどこかほかの場所に探しに行って手にして戻ってからバナナを取ったりはできない。人間以外の動物においては、現在現れている表象への固着が強く、現在から未来へ、知覚されていないが知覚された世界の外部へとつながりを持った世界が、まだ存在していない。人間の言語は、時間と空間の次元を含む世界像（人間にとっての環世界^{*10}）を、他者と共有するためのツールであり、それと同時に、認識と思考の形式を与えるものでもある。

5 動物の知能

5.1 本能行動

本能行動は種固定的な行動形式のことを指し、欲求行動は目標への志向を伴う可塑性を持つ行動形式を指す。学習の成立の余地もなく現れる種固定的行動か否かで、両者を明確に区別できる場合も多いが、それを除くと両者の行動の間に本質的な差異を見分けることは容易ではない。

本能行動は遺伝的に固定されていることは間違いないが、遺伝子とその発現が、どのように固定的な動物行動として現れるかは明らかになっていない。神経系と身体（知覚器と作動器）の協調に行動形式がプログラムされているはずであり、正常な身体発生（遺伝的に規定された発生）に伴う協調がプログラムを成立させるはずである。最も単純な本能行動は、痛覚により手足を引っ込めるような脊髄反射であるが、このような単純な行動は多様な本能行動の中のごく一部である。またネコの歩行のように、CPG（セントラルパターンジェネレータ）と呼ばれる複数の脊髄神経の自律的発火が、知覚器および作動器の間の協調として、特定の運動形式をもたらしていると考えられている [9]。

本能行動は身体発生に伴う、脊髄反射や CPG による身体動作の作動と、知覚（表象）との間の対応によって成立すると推測される。特定の鍵刺激により一意に結びつけられた本能行動が解発されることから、脳中枢を経過することで成立する視覚や聴覚のような表象も、鍵刺激の一つとなるはずである。3.2 節において、身体の状態に対する神経領域の応答を構成素とするシステムの構造を、表象と定義した。身体の作動と神経系の作動は対応関係を維持しながら循環し続けているので、表象とは逆に、神経領域の状態に対する身体の作動の閉域形成が本能行動と定義可能である。表象は一度閉域が形成されると、知覚器の興奮とその知覚野の応答が継続されるが、本能行動では一連の身体の作動が循環したところで、一つの本能行動として完了する。

^{*10} 生物種ごとに知覚される、個体を取り巻く世界のこと [8]。

本能行動のオートポイエーシス論による模式図を図5に示す。図4と概形は同じであるが、構成素 E は身体の作動である。コードは神経系の自己言及性及び自律性と、身体と外部環境の関係の変化による攪乱である。 E_0 はやはりあらかじめ決定することはできず、 E_0 のときの神経領域の状態（特定の知覚野の興奮）が再び現れたとき、 E_0 もまた再現されて閉域が形成されたとき、はじめて始まりの身体の作動 E_0 が確定する。この閉域が一つの本能行動を示しており、身体の作動 E_0 を引き起こすように対応が形成された知覚野の興奮により、一連の身体の作動が連鎖していく。実際の本能行動の解発では、攪乱 D が途中で生起しない場合もあり、その場合はその本能行動が中断される。本能行動が遂行された度合いにより神経領域全体の状態（CPGの興奮の度合いなど）は異なり、その状態により次の本能行動の解発へとつながっていく。本能行動と表象は、正常な身体の発生と成長によりそれぞれの閉域を形成しているため、何かしらの要因（病気やケガなど）で身体の健康状態が悪化すると、一連の本能行動の消失や部分的な脱落が起こり得る。

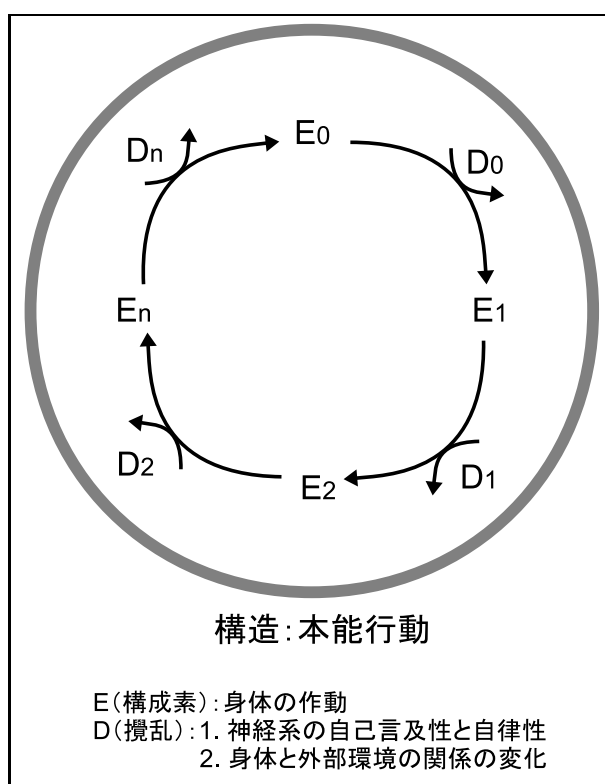


図5 本能行動

5.2 欲求行動（目標志向的行動）

最も単純な欲求行動の一つに、定位の方向転換がある。例えば、カエルが餌であるハエに気付いたとき、まずその方向に眼を向け、次に小刻みな歩行運動を行い、自分の体がハエに対して相対的な向きをとるようにする [10]。この一連の行動に遺伝的に決定された協調運動が対応しているわけではないので、本能的生得運動ではなく、また試行錯誤など介する余地なく定位の方向転換を行うので、学習行動でもない。「餌をとるため」に生得的でない行動が生起するのであるから、「目標を志向する」という心的状態が定位の方向転換という運動協調をもたらしていることになる。実際のところ、構造体の複合的協同の中から、成立している機能を検出し

て実証することは非常に難しいため、すでに定位の方向転換を可能にする運動協調が神経回路によって成立しているが検出できていない可能性も存在する。現時点では欲求行動に次の二つの可能性を考慮すべきである。

1. 行動を駆動しているのは中枢神経系の興奮であって、欲求や認識はその状態によって遅れて現れてくるだけで、欲求や認識が行動を駆動していると勘違いしているだけである*11。
2. 表象そのもの（元になる神経領域の状態でなく）、認識、意識、欲求、意志等の心的状態が実際に身体の作動を駆動する。

これら二つの考え方を受け入れるには、それぞれ次の難問を解決しなければならない。

1. 多様で複雑な欲求行動が、遺伝子のみで決定されるのではなく、身体の発生と発達における環境の作用との協調によって、神経回路網と身体の作動の協調として成立する。
2. 心的状態が物理的基体である神経領域の状態を変化させる。

本論文では主に2の考え方にそって欲求行動を説明するが、1の考え方で説明できる場合はその説明も行う。

5.2.1 認識

動物行動が認識により駆動されると考えると、複雑な動物行動も説明が容易になる。そこでまず、表象のカップリングによる認識の成立について考察する。

人間においては個別の表象（視覚、聴覚、触覚など）が癒合と分離を繰り返しながら、意識の流れの中に変性しながら現れる。動物と人間の知覚の差異がどのようなものか明らかではないが、少なくとも脊椎動物では、個別の表象が癒合と分離を繰り返しながら一体となって意識に現れているであろう*12。そこで本論文では、癒合と分離を繰り返す表象全体の意識への現れを、認識と定義する。人間の場合は個々の表象は比較的分離して認識可能であるが、動物種によっては人間にとっての複数種の表象が癒合して一つの表象であったり、逆に、より分割された表象であり得る。動物種ごとに、各表象間の癒合や分離のしやすさが異なり、さらに同種であっても個体ごとにその程度が異なる。図6に認識システムの模式図を示す。AからEまでが個別の表象システムで、Fが認識システムであり、その構造が意識に現れる認識である。システムのカップリングによる新たなシステムの成立で最も重要なものは、新たなシステムで閉域を形成する構成素と、新たなシステムのコード（構成素の型や産出順序を定める制約）である。社会システムの構成素が個々の人間の間で行われるコミュニケーションであったように（3.3節）、認識システムの構成素は、そのときの表象全体の癒合と分離そのものであるだろう。図6は、表象の癒合と分離の状態が閉域を形成した後の、ある瞬間の図である。表象と表象が直接に関係を結ぶ場合、各表象を振動子とみて引力と斥力による相転移として表象の癒合と分離を解釈したり、その他非線形科学の知見を利用することができるであろう。そして表象の癒合と分離による表象全体の関係が、次の癒合と分離の基礎付けとなって循環していくと考えることができる。表象の癒合と分離による関係が次の関係の基礎付けとなって循環していくのであるから、コードは表象の癒合と分離が構成素の種類となることと、癒合と分離の度合いがどのように変遷して循環していくかを規定している。したがって表象システムのコードと同様に、身体の状態を自己の状態として反映するという神経系の性質そのものと、神経系の状態を引き受けながらも自律的に作動する身体による攪乱がコードを成すと考えられる。神経系と身体にはそれらの周りの環境が浸透してくるので、さらに神経系と身体が攪乱として浸透する認識も、結果として個体の環世

*11 すでにリベットにより提案された考え方である [11]。

*12 人間との差異は明らかでないが魚類にも意識があると考えられている。

界に対応して現れる。

さらに認識が行動に影響を与え得る（運動野を興奮させ得る）と仮定すれば，認識と行動の閉域によって動物行動を説明することも可能になる。例えば「餌」の認識とその対象への方向転換運動が対応するのであれば，捕食行動の待機とともに，視覚像内の差異が壁に止まるハエのような認識となって，ハエに向かって方向転換が引き起こされる。ただし，認識が常に環世界に適応するわけではないので，認識との閉域による動物行動も，場合によっては環世界に対応しない。鍵刺激となる対象なしに本能行動が突発する事態である真空行動を，このような場合の例とみなすことができる。以上が認識と行動が直接対応し得ると仮定したときの，定位の方向転換の説明である。

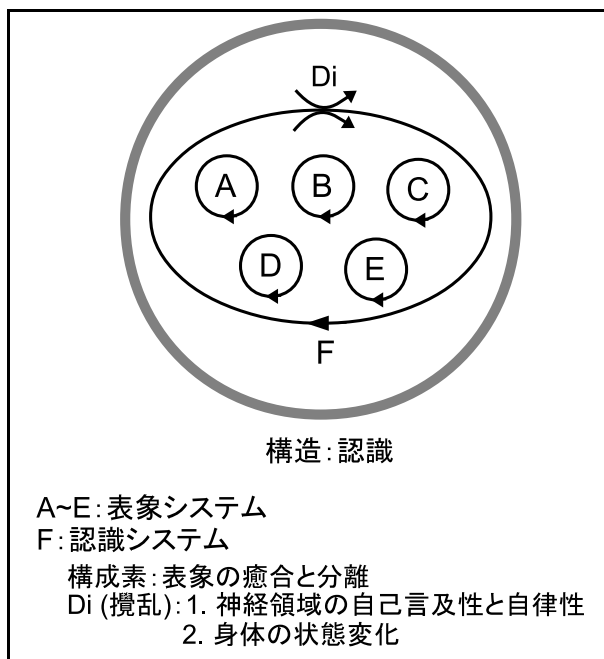


図6 認識システム

5.2.2 高次システムから低次システムへの浸透

5.2.1 節の終わりに，認識と行動の閉域による定位の方向転換の説明を示した。壁に止まるハエを認識したためにハエに向かって方向転換が起こるのであるから，ハエの認識が方向転換という行動を駆動していることになる。認識や欲求が身体を作動させることを，オートポイエーシス論の言葉で言い換えると高次システム（認識システム）の低次システム（身体システム）への浸透である。高次システムと低次システムはお互いが環境になって相互に浸透することで影響を及ぼしあう。高次システムから低次システムへの浸透で参照にすべきなのは，社会システム（3.3 節）での社会から個々の人間への浸透である。社会システムは，人間の間で行われるコミュニケーションを構成素として閉域形成する高階層のシステムであった。個々の人間の社会システムへの自律的応答は，目の前の相手とのコミュニケーションそのものよりも，その総体である社会を想像的に創造してそれに応答している側面が強い。社会システムと個々の人間の関係を認識システムに適用すると，認識システムがそれより低階層の表象システムや身体システムへ浸透するとき，表象や身体自身が自律的に認識に応答しなければならない。

ここでは表象の神経系（身体の一部）への浸透を考える。表象は、ある身体の状態に対する神経領域の応答を構成素とするシステムの構造と定義した（3.2節）。身体の状態と神経領域の状態の間で、対応関係が閉域を形成したときに、神経領域の状態と平行して表象が立ち現れると仮定している。神経領域が身体の状態を反映するのと同様に、身体も神経領域（運動野）の状態を反映して作動するので、表象から神経領域への浸透、つまり表象への神経領域の自律的応答によって身体が作動する。表象から神経領域への浸透の仕方にも二通りが考えられる。一つ目は、表象からの浸透を神経領域の自己言及的応答に置き換える考え方であり、二つ目は表象そのものから浸透するとする考え方である。

まず一つ目の考え方から考える。個々の神経細胞が応答しているのは、神経細胞間の電位伝達だけでなく、脳波のような全体の指標となるもの（その領域の電位変化の振動リズム等）でもあるだろう。神経細胞の集団としてみれば、周辺領域の間に成立している複数のリズム的変動の中から、自らが応答する変動を自律的に選択しているであろう。本能行動が鍵刺激により駆動されるためには、運動野が、運動野に接続する複数の知覚野の中から、鍵刺激を受容する知覚野に選択的に応答する必要がある^{*13}。各領域間の複雑な接続において、神経領域全体では脳波のような全体の指標が立ち現れるので、それに対して知覚野や運動野が自律的に自己の応答を決める。これにより表象の元となる知覚野の興奮と神経領域全体の状態が、運動野の興奮（状態変化）として浸透する。もし認識も神経領域の状態に直接対応するのであれば、認識も表象と同様に考えることができる。この浸透の仕方は、「表象や認識の元になっている神経領域の状態に反応しているのを、表象や認識に直接身体が反応していると勘違いしている」場合に当たる。この場合は高階層から低階層へではなく、同階層から同階層への浸透とみなせる。

二つ目の考え方は、表象そのもの、つまり見えているものや聞こえているもの、見えていることや聞こえていることが、神経領域の状態へ浸透することである。表象が成立している場合に、すでに意識が存在するかはよくわからない。少なくとも人間の場合は、見えているものや聞こえていることは意識に現れることで初めて観察可能となるのであり、脊椎動物一般において同様であると思われる。表象としての現れはその個体の意識において現れるのであり、意識の元の状態であるとはいえ神経領域が表象を現れとして持つわけではない^{*14}。人間が人と人との関わり合いを見て、そこに社会を投射して自律的に応答することを適用すると、自律的に神経領域が、神経領域と神経領域の関わり合いに表象を投射して反応していることになり、神経領域にここまでの能力はないのに、この考え方は一つ目の考え方に帰着することになる。神経領域の状態と並行して表象が立ち現れることを説明するのが困難であると同程度に、表象が神経領域に影響を与えることの説明が困難である。一方、表象が立ち現れることを認めるのであれば、その逆に表象が影響を与えることを認めてもよい、という考え方もあり得る。

ここで神経領域（運動野でなくてもよい）が表象そのものに自律的に応答できると仮定する。この仮定により、表象だけでなく認識、「体験」や「文脈」（5.2.4節）、「世界像」（5.3節）が神経領域の興奮に浸透することを認めることができるであろう。表象よりも高階層のシステムが神経領域に浸透するとき、一つ上の階層からその階層に浸透することを繰り返して、結果として数階層上から浸透する場合と、高階層からそのまま神経領域に浸透する場合の二通りが考えられる。表象や認識だけでなく、体験や世界像も意識への現れであるので、高階層から直接神経領域に浸透すると考えてよいであろう。

*13 知覚野から運動野への接続は脳基底核等を介する。

*14 いわゆる「デカルト劇場」が同じ問題を表している。

5.2.3 認識と身体の構造化

学習と知能行動の説明のために、先に身体と認識の階層化について述べる。神経系は身体の操作のためにあるので、身体の複雑さと神経系の複雑さは相関して発達したとされる [14]。例えば同じ四足歩行でも爬虫類と哺乳類では関節構造が大きく異なり、哺乳類の四肢の方が多様な使い方が可能である。また神経系の発達と共にかつての高次領域が一段階前の機能に特殊化し、新たに発達した領域が低次領域を制御する役割を担うようになったと考えられている。そして神経系の増大と共に運動野への連絡が増加し、知覚と運動の直接的な対応が減少すると共に新たな対応を構成する余地が生じる。

脊椎動物の系統進化により、認識と行動は多重の階層化がなされる。認識と行動の階層化は、新たな領域が加わることで単純にその領域が機能を追加したのではなく、それまでの領域と新たな領域とで再統合が行われ、全体として一つの階層を新たに形成したと考えられる。高次の階層が生成されても、低次の階層が消滅するわけではなく、隠匿される形で高階層の基盤であり続ける。脳の障害における行動の構造の消失と原始的な構造の現れや、動物実験に見られる自然状態では見られない環境との対応値をもたない行動は、脳基質の障害や動物の自律的な実験環境への応答による、高次の階層の消失によるある意味で病的な構造として出現したと考えられる [12]。

5.2.4 学習

動物行動における学習は、特定の外部環境に対する行為の最適化と再行動化と定義できる。獲得された行動を再現することから、記憶が学習に関与すると考えられている。人間の記憶領域についてはまだ統一的意见があるわけではないが、多構造間の経路がその統合機能として記憶の機能をもたらすと考えられている点は一致している [5]。また人間の記憶は性質の違いから複数種類に分類されているが、動物の記憶が人間と同様に分類できるか定かではない。この節では動物実験に見られる、外部環境と行動の間の獲得された対応を可能にするものに限定する。

動物の学習において再現されるのは、構造化された行動であり、記憶されたものは知覚と行動の一対一対応ではなく状況に対応する一つの行為である。例えば紐を引けば餌が得られるように学習したネコは、最初は偶然手が紐にかかって餌を得られたとしても、次に紐を引くときは口で啜って引くのであるから [12]、行動（手で紐を引く）と結果（餌が得られる）の直接の対応ではなく、「紐を引くことで餌が得られること」を学習しているはずである。そのため条件反射のみで学習を説明することは困難であろう。有名な「パブロフのイヌ」の実験では、実験室の中という限られた場所の中で、イヌの方のある種の能動性によって、条件反射が成立していることが示唆されている。ベルの音という条件刺激そのものというよりも、「実験室の中、実験という条件」でのベルの音に反射を構成しているとみなすべきであろう。本節ではネコの「紐を引くことで餌が得られること」の学習を例にとる。

ここで認識と行動が直接対応すると仮定する。知覚と行動の階層化がなされているので、ネコには本来意味のない紐や箱もそれとして認識される。また記憶領域は周りの神経領域を賦活してかつての認識を引き起こす。記憶領域が賦活するとき、ネコにおいては、過去と現在の認識が畳み込まれた一つの認識が現れていると考えられる。また過去の認識も、ある瞬間の認識ではなく、時間の継起に従って連続する認識であろう。この例では紐を引いて、それに連続して餌が得られる、それぞれの認識が畳み込まれて一つの認識となる^{*15}。継起

^{*15} 体験という認識が、それまでに連続して生じた認識を反映してあるので、体験を認識の自己言及による新たなシステムの構成要素とみなせる。

して連続する認識の畳み込みにより、対象の認識に加えて「体験」の認識が可能になったといえる。記憶の曖昧さにより、想起される体験はかつての体験そのものではないはずである。そして体験の認識と行動が対応することで、記憶領域の賦活により、かつての行動が、そのままではなく自分により親しみのある動作を用いて再現される。

記憶領域の賦活によって、連続する認識が体験として一つの認識となったと考えれば、本能行動（5.1節）と類似の模式図で学習を図示できる。動物行動は本能行動と欲求行動の連環として成り立つのであった（4章）。本能行動と欲求行動も共に身体の作動の連環であり、種に固定的か可変的かで、それら身体の作動の連環に区分を設けている。図5では、図示された身体の作動の連環が一つの本能行動であったが、学習を、本能行動（これ自身が身体の作動の連環）の連環の中に過去に行った行動を安定的に挿入した行動連環全体とみなせる。実際に、最後餌を取ったり食べたりするその行動そのものは、種に固定的な身体の作動による本能行動である。図7に学習の模式図を示す。学習により挿入された行動は点線で示してある。攪乱として記憶領域による賦活（体験）が加わっており、本能行動の部分であっても、学習で得られた体験は浸透している。

以上の説明は、認識と行動が直接対応しないとしても説明できる。一度上の事例を体験したネコは、次に同じ場面に遭遇したときに、思考錯誤することなしに紐を引くことができるのであるから、過去と現在を繋ぐ記憶が関与していることは確かである。記憶領域は多構造間の統合機能によって記憶の機能を担っているのだから、ある知覚野（この例では視覚野）の状態を受けて、知覚野や運動野の全体を賦活してかつての状態を再現していると考えられる^{*16}。ネコは偶然手で紐を引いたのちに、次に同じ場面に遭遇したときには、口で紐を引くのであるから、記憶されているのは学習されたときの行動や知覚そのものではないであろう。学習においても知覚（知覚野の状態変化）と行動（身体の作動）の間に対応が取られて連鎖すると考えられる。この対応において、「手で紐を引く」代わりに、「口で紐を引く」という種（または個体）としてより行いやすい行動が挿入されている。記憶領域の賦活により再現されるのは、かつての神経領域の状態そのものではなく、行動の連鎖が引き起こされる程度に類似の状態であるから、結果として、記憶されているのはその瞬間に可能な行為の束である。神経領域の増大による階層化とそれによる行動の遅延が生じているので、単純な反射が抑制されているうちに、行為の束の中から現在可能でかつその種（より高等な動物ではその個体）にとって最も親しみ深い行為が選択される（自然と引き起こされる）。記憶の曖昧さが、引き起こされる行動の多様性とその原因となる知覚の多様性をもたらしているといえる。上記の説明で気を付けなければならないことは、ネコが紐や箱を単純に表象として知覚していることを前提していることである。本能行動や定位の方向転換は、生得的に身体の作動と結びつけられていたり、可能な行動が限られていることから、網膜像から直接生起可能なような表象（視覚野の興奮）において説明が可能であった。しかし、紐や箱はネコにとって生得的に行動に結びつけられた対象ではないため、二次元の網膜像の中から、本来はネコにとって意味を持たない紐や箱を、次の行動の対象としなくてはならない。ネコにおいても人間と同様に、高次の視覚野において抽象的な形態に反応するように階層化されている可能性があり、その場合は視覚野全体の状態として、より複雑化された視覚と行動の対応が可能であろう。また生後7週間ほどは、非生得的な運動協調を成立させる期間であるため[13]、その間に視覚野の変移に敏感に応答するように行動連環が形成されたと考えられる^{*17}。

このように認識と行動の対応を仮定するよりも大幅に記述が複雑になるが、神経領域の状態と身体の作動の対応だけでも学習の説明は可能である。ただしより複雑な学習の場合でも同様に説明可能か明らかではない。

^{*16} 記憶領域による周りの神経領域への賦活は、神経回路網の計算モデルで説明は可能と思われる。

^{*17} 生後の体験により行動連環が成立しているのだが、専門用語としての学習にはこの事態は含まれない。

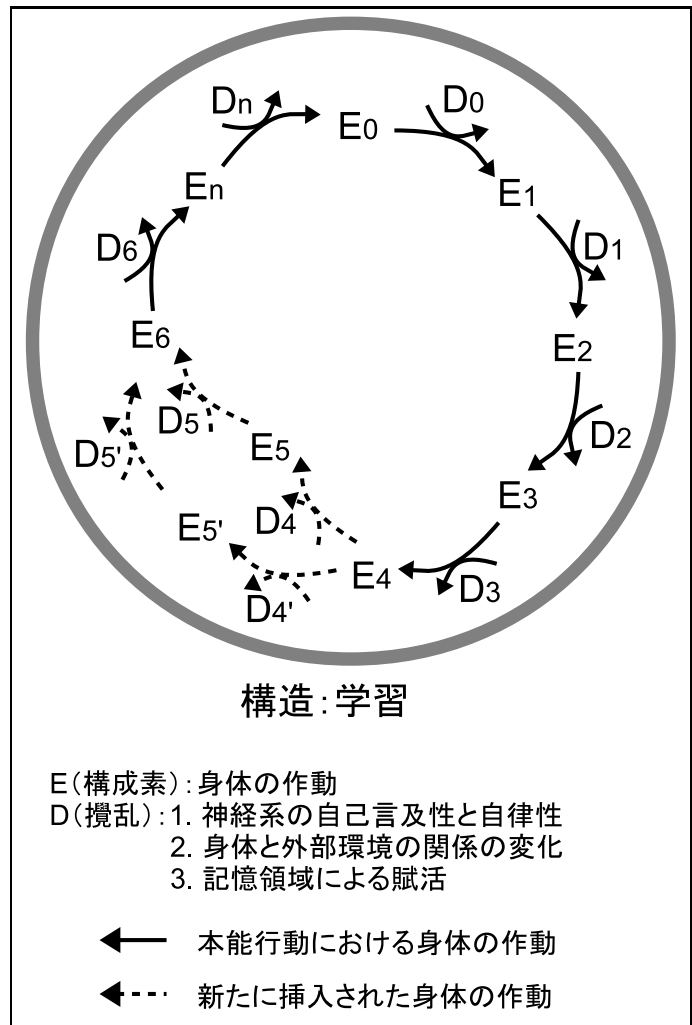


図7 学習

5.2.5 類人猿の道具使用 (知能行動)

オランウータンによる、箱を道具として用いて吊るされたバナナをとる場面が、コンラート・ローレンツにより詳細に記述されている [7]。当該箇所を下に引用する。

部屋の天井にバナナが紐でつるされており、それはオランウータンが床からとどかないほどの高さであった。部屋のすみにはオランウータンが梯子として用いるのに十分なだけの高さの箱がひとつ置かれている。洞察的な行動に関するさまざまな試験をかつてうまくやりとげたことがあるがこの新しい問題は知らないオランウータンは、まずバナナの方をちらっと見て、それから箱の方に眼をやり、しばらくとまどって両方を見くらべる。このとき彼は、ちょうど深く考えこんでいる人間のように頭-や体の他の部分-をかきむしる。それからかんしゃくを起こして手足をばたばたさせ、大股に歩いて侮蔑したようにバナナや箱に背をむける。けれどもそれでも彼は落ちつかず、再び問題にたち返り、餌と箱の間をくり返し見くらべる。今まで不きげんだった彼の表情は急に「明るくなり」、眼は今やバナナからバナ

ナの下の方の床の空間へ、そして箱へ、それから再び床の空間へ、またそこからバナナへと移る、としか私にはいえない。次の瞬間に彼は歓喜の叫びをあげて得意のトンぼ返りをしながら箱の方へ行き、今度はバナナが手にはいることをすっかり確信してただちに箱をバナナの下に引いてくる。

このオランウータンは初めてこの問題に直面しているのであるから、学習ではなく、洞察によりこの問題を解決している。

上記の行動が生起している際の心的状態をまとめる。初めての問題に直面しているのであるから、すでに対応付けられているのではないいくつかの行動を、一連の行動連環として結び付ける必要がある。もちろん新たな行動は、この問題の解決を可能にする行動連環の一部でなくてはならない。「餌と箱の間をくり返し見比べ」たのちに、確信をもって箱をバナナの下に引いてくるのであるから、連続して現れる認識が一つの統合された構造として現れ、到来するであろう未来の体験に向かって、認識と行動が引き付けられて連鎖していったはずである。人間による行動に照らし合わせれば、あるべき未来の状態を、今見えている認識像と共に思い浮かべて、想起した未来の像に向けて箱を移動させて状態を一致させている。ネコの学習の場合(5.2.4節)には、過去と現在の認識が畳み込まれた一つの認識が体験となり、そして記憶領域の賦活により体験の想起とともに行動が再現される。類人猿の道具使用においては、いくつかの体験を結び付けた架空の体験を未来として、現在の認識と分離して保持している。一瞬ではあるが、類人猿において初めて未来と現在の位相が分離して成立している。上記の知能行動を模式図で表すと、図7(学習)とほぼ同じになる。しかし知能行動では、すでに閉域が形成された行動の連鎖の中に新たな行動が挿入されるのではなく、順番も挿入されるかも確定されていないいくつかの行動がある瞬間に結合して一つの行動連環を形成する。

ここで認識を用いずに一連の道具使用の記述を試みる。類人猿の道具使用においても、かつて形成された行動と表象(知覚野の興奮)の閉域を元にする。神経系の高度な階層化により、個々の表象の階層化と本能行動の抑止と発現の遅延がすでに成立している。バナナや箱の間で視線を往復させる行動が、どのように生じするか定かではないが、ネコの場合と同様に生後数か月間で、一定のリズムに従い視点を移動させて視覚像を遷移させることは、運動協調としてすでに成立していると推測できる。そのため本能行動の発現が抑止されている間に、バナナと箱の間の視線の往復運動が生起し、かつて行ったことはあるが一連の運動協調に組み込まれていないいわば架空の行動および表象の対応が、すでに成立している行動連環と結合し、拡張された新たな行動連環が成立する。この行動連環の拡張には記憶領域の賦活が関与しているであろう。また類人猿は人間と同様に非特殊皮質(連合皮質)の比率が高く、これら広範な神経領域による統合機能も関与しているであろう。結合可能な表象と行動は多数存在するが、バナナや箱とその間の床の空間という連続する表象に結びついて生起可能な行動は限られるため、視線の移動の連続のある瞬間にその限られた行動と表象の対応が結合される。一時的に行動連環が成立したため、始めの身体の作動が開始されれば行動が連環していくが、本能行動や学習の連環に比べ身体の作動の連鎖は不安定であり、容易に行動連環が瓦解し停止する。人間にとっては、現在の知覚を行動により変化させることで、来るべき未来の表象の実現を達成しているように思えても、架空の行動と表象が行動連環に結合して新たな行動連環が一瞬成立しただけである。

以上より、道具使用においても認識と行動の対応なしに説明が可能である。しかし人間の知能行動(言語の使用や、箱をどこか別のところに行って取ってくる)までも認識や自由意志なしに可能かどうかはわからない。単純な学習や道具使用ですでに説明が煩雑になっており、認識や意志による行動の駆動を仮定しなければ、人間の知能の説明は困難と思われる。議論を進めるには言語とそれを介した行為に対する詳細な考察が必要である。また人間を含む脊椎動物一般において、認識、情動、意志(思い込みであるとしても)と呼ばれる心的状態の成立を疑うことが難しいので、これら心的状態が結果として現れてくるだけであれば、その生物学的

価値がよくわからなくなる。大多数の種に現れるこれら心的状態が無価値であるとは考え難い。考え方の一つは、これら心的状態が自身の行動を駆動しなくても、それに伴って身体の作動（表情やしぐさ等）として現れるならば、これに他個体が応答しさらに自分が応答することで、社会の形成が可能になる、とすることである。しかしこの場合も、必要となるのは身体の作動の連鎖であるので、最終的に知覚野の興奮が運動野に神経インパルス列を送ればよいだけであり、種として複数個体間で行動の連鎖を可能にする程度に知覚と運動の対応が成立すればよいはずである（社会性昆虫のように）。実際には心的状態そのものが必要なのではなく、身体と神経領域の対応が成立した際に偶然並行して現れた心的状態が、系統進化によって引き継がれただけという可能性がある。心的状態が身体の作動に影響を及ぼしえないのであれば、当然自然淘汰の対象にはならないので、何の機能も持たない心的状態が高等動物に残り続けても不思議ではない。これ以上議論を続ける準備ができていないため、本論文ではこれ以上言及しないこととする。

5.3 言語現象のオートポイエーシス論による解釈に向けて

人間に特殊な行為の代表として言語使用が挙げられる。類人猿の道具使用が可能な条件から、人間の言語の使用に潜在していることが推測できる [15]。類人猿の道具使用は、餌と箱や棒が同じ視野に入るときに限られており、一度学習した個体であっても、箱や棒をどこか他の場所に行って取ってきたりはできない。人間がどこかに行って箱や棒を取って戻ってくるができるのは、世界の時間的かつ空間的な恒常性が成立しているからである。人間が「台」や「棒」として名指しているのは、目の前の対象物を含む範疇であり、「台」や「棒」の役割を持つ対象物が同時にそして永続して存在することを前提にしている。人間が認識している環世界は、時間と空間の次元を前提として包含する「世界像」であり、言葉で名指すその対象は、物自体としての対象そのものではなく、「世界像」の一要素である。

オートポイエーシス論としての言語の考察には、動物学の延長としての議論のほかに、発達心理学や精神医学による言語の理解を参照する必要がある。またこれらの理論と言語学とのすり合わせも必要であろう。多岐にわたる理論を統合する必要があり、言語現象のオートポイエーシス論による解釈は今後の課題とする。

6 課題と応用

本論文では、オートポイエーシス論による、心的状態と動物行動の体系化を試みた。表象や認識をオートポイエーシス論を用いて定義し、動物の本能行動と欲求行動をオートポイエーシス論で説明を与えることができた。しかし体系化しきれなかった事項も多々存在する。

オートポイエーシス論は元々、循環と変遷の中で同一であり続ける対象を記述するための理論であったので、一回性の事柄を説明するには不向きである。そのため体験や知能行動のような一回性の出来事を、オートポイエーシス論によって図式化することが困難となっている。またオートポイエーシス論で重要な概念である自己言及システムを上手く図式化することができていない。体験や記憶を、認識の自己言及による新たな認識とみることができるので、自己言及システムの概念整理と図式化により、認識論へのオートポイエーシス論の適用可能性が向上すると見込まれる。一方、神経領域の状態に平行して心的状態が析出することを、前提として議論を進めざるを得なかった。この前提を説明可能な理論が存在するのであれば、この前提の解明を優先すべきであろう。いずれにせよオートポイエーシス論自体が発展途上の理論であるため、さらなる概念整理と具体的事例への適用が必要である。

6.1 各種理論とのすり合わせの可能性

オートポイエーシス論を用いた知能や認識に関する議論に対し、類似の思考を持ちながら独自の展開がなされた議論が存在する。本論文で展開した仮説の妥当性を考慮するためには、これらの議論との突き合わせが必要である。また一般的な生物学や脳科学の知見との突き合わせも必要である。本論文の内容の検証やさらなる発展のためには、これらの理論を体系的に理解する必要があり、困難な仕事であるため、理論の突き合わせのための指標を示すにとどめる。以下に箇条書きで重要な議論を示す。

1. 非線形科学による生命論.
2. 臨床の知見を元にした神経生理学および神経心理学.
3. ヴァイツェッカーによるゲシュタルトクライス論.
4. バルンシュタインに代表される運動神経生理学.
5. ローレンツに代表される動物行動学.
- 6.フロイト, ラカン派の精神医学.
7. ギブソンやメルロ=ポンティらによる心理学を中心とした考察.
8. 認識論の哲学.
9. 各種言語理論.

6.2 人工知能の可能性

知能の定義は非常に難しいが、「問題解決のための合目的行動を、先天的に決められていないにもかかわらず選択できること」と仮定する。そう仮定すれば動物学の言葉では、学習や知能行動などの欲求行動を知能に含めることができる。例えば人間にとっての計算は、数の概念や演算の規則を習得して初めて可能になる高度な知能であるが、計算機の計算はあらかじめ設計された通りに動作した結果なので、一般に知能とは呼ばれない^{*18}。同様に機械学習も、内部動作の仕組みを考慮に入れば、知能とみなすことは難しい。世間一般において人工知能という言葉で期待されているのは、人間や高等動物と同じように、世界を認識し思考して周りの環境に適切に応答することであろう。現在の人工知能は、限られた機能や状況において人間と同等かそれ以上の能力を持つが、人間や動物と同様の能力を持つのではない。

動物の身体は物理的基体であるので、原理的には機械や計算機も、人間や動物と同様の、心的状態（精神）を伴う知能を持ち得るはずである。身体を構成するのは自律的に作動し続ける細胞であるから、人工の「知能」を持つ機械を作るには、自律性を持つ素子が必要かもしれない^{*19}。身体は受精卵から発生に伴って形成されるので、表象や認識が、身体が徐々に形作られていく過程で成立するのであれば、発生する身体にあたる構造を機械に持たせる必要がある。表象や認識が成立する過程を何かしらの形式にコードすることができれば、身体との協調で作られるこれらの過程を、異なる機構に写し取って擬制して与えることは可能であろう。その際、ニューラルネットワークによる学習等で、ネットワーク全体の入力出力応答の形で写し取ることも可能であるかもしれない。このような写し取りが可能であれば、中枢神経系にあたる構造のみを持つ、知能を持つ機

*18 一部の動物の計算能力も計算機と同様である。

*19 このような素子を持つ計算機は、もちろん計算機としては役に立たない。

械も作成可能なはずである*20。

また物理的素子を持たないプログラム上での知能（精神を伴う）も、構築できる可能性がある [2]。計算機でのアプリケーション上の計算というプロセスが、生命システムのプロセスとあまりに異なるため、現時点では判断できない。またたとえ生成できたとして、身体を持つ機械以上に確認が難しい。

7 おわりに

本論文では、表象や認識および動物学における知能行動に対し、オートポイエーシス論による説明を試みた。表象や知能の議論で問題となるのは、物理的構造体とそこに還元されえない心的状態を、どのように結び付けるかである。本論文でも、表象の成立においては、神経系の状態に平行して現れることを前提にして議論を行った。まず表象を、身体発生に伴う、神経領域の状態の閉域が形成されたときの、その状態の遷移による構造と定義した。次に本能行動を、表象が成立すると同時に身体の作動の閉域が形成された際の、その構造と定義した。そして順次、表象と本能行動を元に、定位の方向転換、学習、道具使用（知能行動）のオートポイエーシス論による解釈を示した。

学習の際に、記憶領域の賦活によって、時間的に連続する認識が、「体験」として一つの認識にまとめ上げられたと定義した。これによって学習を、記憶領域による賦活（体験）を攪乱として、本能行動連環の中に過去に行った行動を安定的に挿入した、新たな行動連環全体として定義できた。また動物が知能行動（道具使用）を行うとき、いくつかの体験を結び付けた架空の体験を未来として、現在の認識と分離して保持しているとみなした。そして順番も挿入されるかも確定されていないいくつかの行動が、ある瞬間に結合して一つの行動連環を形成することを、知能行動として定義できた。一方これらの議論から、認識や意志による行動の駆動を仮定しなくても、学習や知能行動を説明可能であることが示唆された。本論文での仮説の検証のため、様々な理論との突き合わせが必要である。

参考文献

- [1] 河本英夫; オートポイエーシス 第三世代システム, 青土社, 1995.
- [2] 山下和也; オートポイエーシス論入門, ミネルヴァ書房, 2010.
- [3] ウンベルト・マトゥラーナ, フランシスコ・ヴァレラ; 知恵の樹-生きている世界はどのようにして生まれるのか, ちくま学芸文庫, 1997(管啓次郎訳).
- [4] 斎藤環; 文脈病, 青土社, 2001.
- [5] 山鳥重; 神経心理学入門, 医学書院, 1985.
- [6] コンラート・ローレンツ; 動物行動学 I 再送版, 新思索社, 2005(丘直道・日高敏隆訳).
- [7] コンラート・ローレンツ; 動物行動学 II 再送版, 新思索社, 2005(丘直道・日高敏隆訳).
- [8] ユクスキュル・クリサート; 生物から見た世界, 岩波文庫, 2005(日高敏隆・羽田節子訳).
- [9] 多賀徹太郎; 脳と身体の動的デザイン-運動・知覚の非線形力学と発達, 金子書房, 2002.
- [10] 大村敏輔; 本能行動とゲシュタルト知覚, 九州大学出版会, 1991.
- [11] 山口裕之; 認知哲学 心と脳のエピステモロジー, 新曜社, 2009.
- [12] モーリス・メルロ=ポンティ; 行動の構造, みすず書房, 1964(滝浦静雄・木田元訳).

*20 筆者の個人的感情をいうと、身体を持たない知能機械の作成は不可能に近いと考えている。

- [13] 大石孝雄; ネコの動物学, 東京大学出版会, 2013.
- [14] ニコライ・A・ベルンシュタイン; デクステリティ 巧みさとその発達, 金子書房, 2003(工藤和俊訳, 佐々木正人監訳).
- [15] 坂田純; 動物学から見た人間の言語に関する考察, 信学技報 117(82), pp. 37-42, 2017.