AIやロボット工学に寄与するかもしれないハイデガーの洞察

――人間と動物の連続性と断絶――

君嶋泰明

はじめに

本稿の目的は、ハイデガーから、AIやロボット工学に寄与するかもしれない、ある洞察を引き出すことである。もちろん、ドレイファス(Dreyfus, 1979)以来、ハイデガーが、それらの分野の枠組みを見直す際の一つの参照点となってきたことは、よく知られている。だが、それらの分野における近年のいくつかの研究は、また新たな局面を迎えつつあり、異なる角度からの検討を必要としている。本稿が提示するハイデガーの洞察は、その際の検討材料の一つになりうると私は思う。

本稿は、AIやロボット工学やその関連分野が提示してきた知能モデルを発展史的に追う、という仕方で議論を進める。その際、典型的な三つのモデルが順に検討されることとなる。簡便のために、ここでそれらに名前を付けておこう。一つ目は、しばしばGOFAI（Good Old Fashioned Artificial Intelligence）と呼ばれるものである。二つ目は、ブルックスにならってカンブリア紀の知能（Cambrian Intelligence, 以下CIと略記）と呼ぶことにする(Brooks, 1999)。三つ目は、アンダーソンにならって新石器時代の知能（Neolithic Intelligence, 以下NIと略記）と呼ぶことにする(Anderson, 2003)。これらの名前で何が意図されているのかも含めて、それぞれ1節、2節、3節で検討していくことなるが、その際、それらに通底する、人間と動物の連続性と断絶というテーマをも追跡していくこととなる。以下で詳しく論じるように、以上の三つのモデルは、それぞれこのテーマに対して異なる態度を表明している。だが、本稿が提示するハイデガーの洞察は、それらの三つの態度を統一的に扱える可能性を示唆している。このことは3節の最後で示されるだろう。

1 GOFAI

1.1 GOFAIの性格づけ

アンダーソンによれば、GOFAIは、デカルト主義的かつ認知主義的と形容されうる(Anderson, 2003, p. 94)。デカルト主義とは、煎じ詰めれば、「経験によって影響されるが、因果ではなく理性によって支配される、自由な表象の内的領域を前提すること」(ibid., p. 93)である。アンダーソンはさらに、デカルト主義の特徴として、人間と動物の間の不連続性

を挙げる。

> デカルトにとって、動物は、確かに複雑で興味深いが、単なるメカニズムであって、物理的なオートマトンである。動物は確かに感覚を持つ。なぜなら、感覚には固有の器官さえあれば事足りるのだから。しかしながら、動物は思考を欠いている。おそらくその最も重要な証拠は、動物が言語を欠いているということだろう。こうした、世界における感覚と行為は思考を必要とするということの否定、そしてそれに付随する、思考と、言語使用に典型的に見られるより高次の推論と抽象との同一視は、デカルト主義的態度の真の核心かもしれない。(ibid.)

アンダーソンによれば、認知主義の中心的態度とアプローチは、こうしたデカルト主義的世界観を受け継ぐことに由来している。認知主義とは、端的に言って、「心の―思考の―中心的機能は、明示的な規則に従った記号操作という観点から説明されうる」という仮説である(ibid.)。この仮説は、「区別された、同定可能な、各種の内的状態ないしプロセス」(Clark, 1996, p. 43)の存在にコミットする。この内的状態ないしプロセスとは、要するに記号操作のことである。そして、そうした記号操作の「体系的ないし機能的役割は、特定の特徴ないし事態の代わりをすること」(Clark, 1996, p .43)、つまり表象することである。とはいえ、認知主義においては、記号はその意味からは切り離され、純粋に形式的なものとして扱われる。そして、その必然的帰結として、認知主義は、明示的に特定できる思考の規則にコミットする。というのも、実際、記号と意味の関係は難しい問題であり、この問題を避けて記号の形式的側面にのみ着目するなら、記号操作、つまりある認知状態から別の認知状態への変換は、それを支配する形式的な規則を必然的に必要とするからである。

GOFAI は、以上のように、形式的な規則に基づいて記号を操作することによって特定の事態を表象するものとして性格づけられる。さらに、こうした記号操作としての知能は、人間だけが持つものであり、動物と連続的なものではない、と GOFAI は考える。

1.2 GOFAI の問題点

ドレイファスの一連の仕事(Dreyfus, 1979, 1981, Dreyfus & Dreyfus, 1986)は、GOFAI の限界を最初に指摘したものとして評価できるだろう(Anderson, 2003, p. 95)。プリストンは、ドレイファスの GOFAI 批判を比較的簡潔にまとめている(Preston, 1993)。以下ではそれに従って、GOFAI の限界がどのようなものなのかを見ておきたい。

プリストンはまず、ドレイファスが「背景」と呼ぶものについて解説する。

あなたが家を出るときにドアに鍵をかけることを例にとろう。[…]あなたは、鍵を回すことができ、その作業がうまくいったかどうかを言うことができさえすればよい。残りは錠そのものがやってくれる。したがって、ある程度まで、あなたがことをなすことができるのは、世界の働き方によってであって、世界の働き方についてあなたが明示的に知っていることによってではない。(Preston, 1993, p. 44)

しかしながら、そうした実践の複雑さを過小評価してはならない。あなたは一般規則として鍵をかけるかもしれないが、日常的なもの（あなたはちょうど外の郵便受けに走っている）から非日常的なもの（家が火事だ）まで、たくさんの例外が存在する。[…]加えて、鍵かけが適用される他の（あなたのオフィス、車、オフィスの建物などへの）ドアも全て存在する。[…]そのうえ、あなたの生活には、鍵かけの実践が適用されない[…]無数の種類のドアも全て存在しており、あなたは進行中の日常的活動において、容易にそれらを、鍵かけが適用されるものから区別できるように思われる。(ibid., pp. 44-45)

このように、鍵をかけること一つとっても、それは、ドア、鍵の使用、その規則、その例外に関する無数の事実に依存していることが分かる。ドレイファスはこうした一群の事実や規則と思われるものを「背景」と呼び、それらは日常的活動が円滑に進行しているときの背景であるかぎりにおいて、明示化されえないとする(Dreyfus, 1981, Dreyfus & Dreyfus, 1986)。こうした背景の存在を踏まえたうえで、ドレイファスは、記号の組み合わせで表現される明示的表象だけで知的振舞いを実現できる、という GOFAI の前提を批判する。

だが、なぜ背景は明示的に表象されていると前提してはいけないのだろうか。例えば次のように考えることもできるだろう。つまり、我々が日常的活動において背景の明示的な表象を自覚していないからといって、背景が明示的に表象されていないとはかぎらない。なぜなら、我々は単に、背景の表象へのいかなる意識的なアクセスをも持っていないだけなのかもしれないから(Preston, 1993, p. 45)。だとすると、ドレイファスは、いかなる根拠で、GOFAI の前提が間違っていると考えるのだろうか。プリストンによれば、

問題は、いかなる与えられた状況においても、どの背景的事実が関連している（relevant）のかが決定されなければならない、ということである。しかし、このためには、背景が、どの状況でどの事実が関連するのかについての事実を含んでいる必要がある。ところが、これは、他の事実に関する事実という無限後退を引き起こすよう

に見える。[…]したがって、この後退を着地させるには、背景は完全かつ明示的には表象されていない、ということでなければならない。(ibid., p. 46)

こうして、GOFAI の前提が退けられたところで、ドレイファスは、二つのオプションを提示する。

一つの応答は、[…]人間の関心と実践の「知識」は全く表象される必要はない、と言うことだ[…]。

別の可能な説明は、表象の場所を許すだろうが、これらはたいてい形式なしの表象であり、よりイメージに似ており、それによって、私は、私が知っているものではなく、私であるものを切り開くのである。(以上 Dreyfus, 1981, pp. 202-3)

プリストンは、第二のオプションはコネクショニズムによって具体化されるとし、自分が検討したいのは第一のオプションであるとする。本稿もまたこのラインで議論を進めたいので、コネクショニズムには立ち入らず、プリストンが第一のオプションに従った AI の例として挙げる、ブルックスの移動ロボット（mobile robot）の議論に入っていきたい。

2 カンブリア紀の知能

2.1 カンビリア紀の知能の性格づけ

GOFAI が知能を人間に特有の記号操作能力と決めてかかり、トップダウン的に知能にアプローチするのに対して、CI は、むしろ人間と動物の連続性を強調し、ボトムアップ的に知能にアプローチする。こうしたブルックスの知能観は、例えば次のような記述に端的に表れている。

地球上の生物学的進化の時間のかかり方ついて考えてみるのは有益である。単細胞の存在者は、およそ 35 億年前に、原初のスープから生じてきた。光合成を行う植物が現れるまでに 10 億年が経過した。その約 15 億年後、つまり約 5 億 5000 万年前、最初の魚類と脊椎動物が出現し、4 億 5000 万年前に昆虫が出現した。すると、事態は急速に動き始めた。3 億 7000 万年前には虫類が出現し、3 億 3000 万年前に恐竜が、2 億 5 千万年前にはほ乳類が続く。最初の霊長類が現れたのは 1 億 2000 万年前で、類人猿の直接の祖先はわずか 1800 万年前である。人がほぼ現在の姿で出現したのは 250

> 万年前である。人は 1 万年前に耕作を発明し、書くことは 5000 年前よりもこちらであり、「専門的な」知識は、過去数百年の間に発明されたにすぎない。
>
> このことが示唆するのは、問題を解決する振舞いや、言語、専門的知識や適用、理性、これらすべては、存在し反応するための不可欠の要素がひとたび利用可能になれば、きわめて単純なことだ、ということである。この不可欠な要素とは、生命の維持と再生産に十分な程度まで周囲を感覚し、動的な環境において動き回る能力である。知能のこの部分は、進化がその時間を集中してきたところである―そこがより難しい部分なのだ。(Brooks, 1991, p. 140)

このように、ブルックスは、動的な環境で周囲を感知し動き回る能力を知能の基盤をなすものと考えている。さらに、それさえ実現できれば、より高次の知的能力の獲得は容易いとさえ考えている。このことから、ブルックスの「カンブリア紀」の知能という命名の意図が理解できる。カンブリア紀というのは、およそ 5 億 7000 万年前から 5 億 900 万年前までの期間のことで、上の記述でいうと、単細胞の存在者が生じてから 25 億年後、最初の魚類や脊椎動物が現れた時期にあたる。だとすれば、上述の知能の基盤が初めて獲得されたのは、このカンブリア紀であると言えるだろう。CI という名前で意図されているのはまさにそうした知能の基盤なのである。そして、もし AI でそれさえ実現できてしまえば「事態は急速に動き始め」るのだから、あとは比較的容易に人間レベルの知能にまで発展させていけるに違いない。おそらくこのような青写真がブルックスの頭にはあったのだろう。

こうしたブルックスの立場からすると、AI 研究は、いきなりフルスケールの人間知能を実現しようとせずに、より単純だが、実際の世界で自律的に感覚・行為する完全な知的システムを構築し、そこから徐々に能力を増やしていくべきである(Brooks, 1991, p. 139)。だとすれば、CI はある種必然的に、自律的な移動ロボットというかたちをとることになる。

さて、ブルックスは、こうした移動ロボットを作っていく過程で、次のような「予期しなかった結論（C）に達し、よりラディカルな仮説（H）を持つ」に至る。

> （C）我々がきわめて単純なレベルの知能を調べると、世界の明示的表象とモデルが端的に邪魔になっているのを発見する。世界を世界自身のモデルとして用いるほうがよいことが判明する。
>
> （H）表象は、知的システムの非常に大きな部分を構築するときの、誤った抽象化の単位である。(ibid.)

表象が「邪魔」であり「誤った抽象化の単位」であるとはどういうことか。表象を使ってAIを構築しようとする際の問題は、前節で、関連事実の表象の無限後退というかたちで示された。ブルックスは、同じ関連事実の表象にかかわる問題を、また別の角度から指摘する。椅子に関する関連事実を表象する場合を考えよう。椅子には様々な事実が属している。ある椅子は、平らな座る場所を持ち、背中の支えを持っているかもしれない。椅子は、サイズ、強度、形状に幅がある。椅子は、木や金属やプラスティックでできている場合もあれば、何らかの素材で覆われている場合もある。椅子のある部分は柔らかかったりする。このような無数の事実の中から、AIは関連する事実のみを表象しなければならない。だが、この「知能の本質」(ibid., p. 141)である「抽象化」は、実質的に、AI自身ではなく「研究者によってなされ」(ibid.)ている。そのようなAIを知能と呼べるだろうか。

アンダーソンによれば、この関連性問題から引き出されうる一つの規範は、何のために表象しているのかを知らなければ、表象について考えるのは意味がない、ということである(Anderson, 2003, p. 98)。「あらゆる表象者は必然的に選択的であり、よい表象はそれゆえ、特定の（種類の）エージェントによる特定の（種類の）使用に定位されている」(ibid.)。この考えを拡張していくと、「知覚と行為の間のギャップを埋めることになり、行為の観点から知覚の大部分を捨てることになるかもしれない」(ibid., p. 99)。ブルックスはまさにこの方向で自律的な移動ロボットを構想している。

彼が「クリーチャー」と呼ぶこの移動ロボットに要求される性格は次のようなものである。

- クリーチャーは、適切に、時機を得た仕方で、その動的な環境における変化に対処しなければならない。
- クリーチャーは、その環境にかんして頑健であるべきだ。世界の性質のわずかな変化が、クリーチャーの振舞いを完全に破壊することにつながるべきではない。むしろ、期待されるべきは、環境がより変化するにつれた、クリーチャーの能力における段階的な変化だけである。
- クリーチャーは、多数の目的を維持できるべきだし、また、そのうちで自身を見出す周囲の状況に依存することで、どの特定の目的を実際に追求しているのかを変えることができるべきだ。それゆえ、それは周囲に適合することも、偶発的な周囲の状況を利用することもできる。
- クリーチャーは、世界の中で何かをするべきだ。それは何らかの現存する目的を持っているべきだ。(Brooks, 1991, p. 142)

次節では、彼がこの構想を実現するためにどのような戦略をとっているのかを見よう。

2.2 自律的移動ロボット「クリーチャー」

クリーチャーの最も重要な特徴は、世界の完全な表象を生み出す中心的システムがない、ということである。そのかわり、レイヤーと呼ばれるいくつかの活動パターンが並行して作動し、それが結果的にシステム全体を形成することになる。中心に据えられた世界の表象なしに、並行して作動する多重のレイヤーによって形成されるシステムは、世界の変化に対して、いくつかのレイヤーを失うことがあっても、システム全体として動けなくなることはまずない（少なくとも、世界の表象に頼るロボットよりはその可能性が低い）。クリーチャーは、中心的なコントロールを持たない、競合するレイヤーの振舞いの集積である。

このクリーチャー構築の手順は、まずきわめて単純だが完全な自律的システムを作り、それを実際の世界でテストし、そこに新たなレイヤーを追加し、また実際の世界でテストする、ということの繰り返しとなる。ブルックスは、「進化はこのアプローチをうまく使ってきている」(ibid., p. 145)と言い、このクリーチャーの構築を生物の進化と異ならないものと考えている。

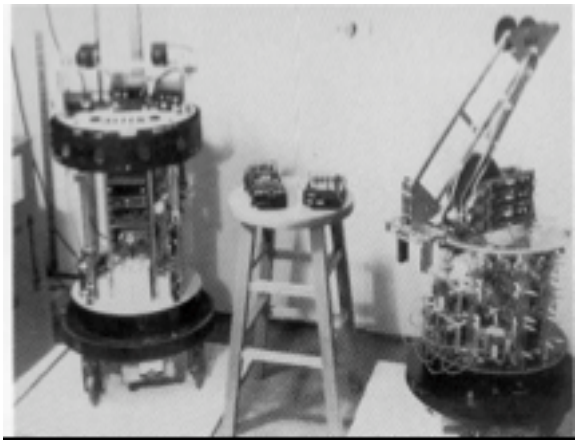

図 1 :MIT 人工知能研究所の 4 つの AI(Brooks, 1991, p. 146, Fig. 1)

ブルックスは、実際の世界（MIT 人工知能研究所の研究所とオフィスのエリア）で動いている図 1 のようなクリーチャーを紹介している(ibid., p. 146)。これらのロボットは全て、服属アーキテクチャ（subsumption architecture）という抽象的なアーキテクチャを実行する。服属アーキテクチャの各レイヤーは、単純な有限状態マシンの、固定されたトポロジーネットワークからなる。有限状態マシンは非同期的に作動し、有限の長さのメッセージを、ワイヤーを通じて送受する。各マシンは内部にタイマーを持ち、メッセージの到着か、指定された時間の到来によって、自身の状態を変化させる。各レイヤーは、抑制（suppression）と禁止（inhibition）と呼ばれるメカニズムによって組み合わされている。新たなレイヤーが追加されると、新たなワイヤーが既存のワイヤーにサイドタップされる。この新たなワイヤーからメッセージが到着すると、既存のワイヤーからのメッセージは、入力側で一定時間「抑制」される。

また他方で、この新たなワイヤーのメッセージは、既存のワイヤーのメッセージを一定時間「禁止」する。クリーチャーは、こうしたメカニズムによって、多重のレイヤーの並行的な作動を可能にしている（詳しくは、ibid., pp. 146-7）。

2.3 カンブリア紀の知能の問題点

以上のような、単純なシステムから始めて徐々に高次の知能を実現させていこうとするブルックスのアプローチには問題がある。まず、クラークは、近年の認知科学における、GOFAI 的な前提から脱却したいくつかの仕事の特徴を、次のように挙げている。

1. 緊密に連結された組織体と環境の相互作用をモデル化し、理解する試み。
2. 適応成功を支えるうえでの、複合的で、局所的に有効な、こつ、ヒューリスティックス、近道（しばしば行為を含む）の役割がますます認められること。
3. 社会的、文化的、技術的に分散し、局所的な環境へと、実質的な問題解決の活動の「負荷を降ろす（off-loading）」、問題解決の様々な形式への注目。(Clark, 2001, p. 126)

クラークによれば、ブルックスやコネクショニストのようなアプローチは、上の 1 に注意を払い過ぎた結果、伝統的にターゲットにされてきたものを見失ってしまっている(ibid.)。本稿の文脈で言えば、完全に表象なしでシステムを構築しようとするブルックスは、記号や表象を使った知的活動をあまりにも低く見積もってしまっている。クラークの例を引けば、表象は、パリの夢を見たり、合衆国の銃規制の方針について考えたり、来年の休みの計画を立てたり、ローマでの休日の間に、あなたのニューヨークのアパートの窓の数を数えたりするといった、ありふれた活動に欠かせないものであるように思える(ibid., p. 129)。こうした「表象に飢えている（representation-hungry）」(ibid.)活動の存在は、CI を決定的に論駁するものではないが、少なくとも CI がいずれこうした活動をも実現できるのかどうかにかんして、懐疑的になるには十分である。

また、単純なクリーチャーから徐々に人間レベルの知能へと近づいていこうとする CI は、3 を適切に考慮に入れることができない。クラークによれば、高度な認知の多くは、「外的かつ／または人工的な認知補助の領域」(ibid., p. 131)に合わせられ、振り当てられた能力に依存している。彼はそうした領域を「認知テクノロジー（cognitive technology）」と呼ぶが、それは例えばペン、紙、パソコンのようなものである。人間は、他のどの生物よりも、こうした「非生物学的要素[…]を配置し、[…]拡張された認知システムを作り出すことで、

処理プロセスの基本的な生物学的モードを補完する」(ibid., p. 134)。そして、まさにそうした認知システムの拡張を行う能力の有無に、人間と動物の、「認知進化における断絶」(ibid.)がある。クラークは、ブルックスの「クリーチャー」のような、基本的な適応反応の生物学的モードに微調整を加えることでフルスケールの人間の知能を実現しようとするアプローチを、「生物学的認知漸進主義（biological cognitive incrementalism）」と呼ぶが、そうしたアプローチでは、この認知システムの拡張という側面を適切に扱えないのである(ibid., p. 122)。

では、その拡張された認知とはどのようなものなのか。次節では、この認知をも組み込んだ知能モデルの可能性を考えてみよう。

3. 新石器時代の知能に向けて

3.1 拡張された認知

少し長くなるが、クラークが挙げる拡張された認知の例を見よう。

> 学術論文を書くおなじみのプロセスを取り上げよう。最後に、輝かしい完成作品を前にして、敬虔な唯物論者は、知らず知らずのうちに、いい仕事をした脳に満足しているかもしれない。しかしこれはミスリーディングである。それは（例のごとく）考えのほとんどがどのみち我々自身のものではないからだけではない。脳は絶えずメディアやテクノロジーと相互作用しているわけだが、脳がそれらの様々な特殊機能と協働し、またそれらに依存する複雑な仕方に、最終的な作品の構造、形式、流れが、しばしば著しく依存しているからである。[…]我々は、ある古い注に目を通すことから始め、次にオリジナルな資料を参照したのかもしれない。我々が読むとき、我々の脳は、いくつかの断片的な反応をその場で生み出した。それは、その頁ないし余白に、痕跡として適切に保存された。そのサイクルは反復し、オリジナルな計画やスケッチへと立ち返るために中断したり、それらを同じような断片的な仕方でその場で修正したりする。この、批評したり、並べ直したり、能率化したり、関連づけたりする全体的なプロセスは、外部メディアのきわめて特殊な性質によって性格づけられており、この外部メディアが、単純な反応の継起を、何か論証のようなものへと組織化させ、（願わくば）発展させるのである。(ibid., p. 132)

学術論文は、最終的に一つの論証へと組織化されるが、そのためには、少なくとも、断片的な他人の考えや自分の考えを、保存し、並べ、つなげなければならない。これらは脳が

単独でやることではなく、紙、ペン、パソコンなどのテクノロジー環境も、全体としての論文を書くという作業の一部である。ここでは、脳はむしろ、「脳、身体、テクノロジー環境の間で絶えずループする、複雑で反復的なさまざまなプロセスにおいて、媒介要素として振る舞う」(ibid.)のである。

クラークが言うように、こうした拡張された認知が人間だけに許されたものであるなら、生物学的認知漸進主義をとる CI は分が悪いだろう。とはいえ、人間が進化の産物である以上、CI は全く間違っているわけではない。CI が捨てなければならないのは、生物学的な基盤に微調整を加えていけばいずれ人間の知能に到達できるという、その素朴な楽観論なのである。

3.2 ハイデガー流の人間と道具の相互作用ループ

クラークは、「道具や装置の役割の一般的な強調を伴うこの種のストーリーは、明らかに、そのルーツを、ハイデガー[…]と、ジョン・ホーグランド[…]の仕事に持っている」(ibid., p. 135)と述べているが、それ以上立ち入っていない。以下では、クラークの言う「脳、身体、テクノロジー環境の間」のループのようなものが、ハイデガーにおいて見出せるかどうか真面目に考えてみたい。そのための準備として、まずは次のような状況を思い浮かべてみよう。

(a) 中学生のあなたは、自分の部屋の中で一人、鉛筆を使って、ノートに簡単な数学ドリルの問題を解いている。明日提出しなければならない宿題なのだ。すらすら問題が解けることも手伝って、あなたは高揚感に浸りながら、完全に作業に没頭している。

(b) ところが、突然鉛筆の芯が折れてしまった。一瞬で我に返ると、今まで特に意識していなかった鉛筆やドリルの本が、急に押し付けがましく、よそよそしく感じられた。

(c) それから二秒と経たないうちに、あなたは、机の上の左手前方に電動鉛筆削り機が置いてあるのに気づく。あなたは慣れた手つきで鉛筆を穴に差し込むと、実にスムーズに鉛筆は削り終わり、あなたは尖った鉛筆の先で、再びノートに数字を書き始める。

ハイデガーは、この（a）→（b）→（c）のような体験世界の移行を、形式的に理論化しようと試みている(cf. Heidegger, 1987)。ハイデガーは、体験世界に、(a)、(c) のような没入

モードと、(b) のような反省モードの両極を認め、それらの間にいくつかの程度差を認めている(cf. Heidegger, 2001, S.74)。ハイデガーにおける人間と動物の差異について考えるときにしばしば引き合いに出される「人間は世界形成的である（weltbildend）」(Heidegger, 1983, S. 263)というテーゼは、人間は様々な体験世界に没入することができる、というふうに解釈できる。とすると、問題は、何がある世界から別の世界への移行を可能にしているのか、ということである。

ハイデガーにおいて、多くの場合、ある体験世界への没入は、（どんな些細なものだとしても）何らかの道具の使用である。上の例で言えば、体験世界（a）への没入は、鉛筆の使用である。さて、上の例を注意して見ると、(a) から（b）への移行のきっかけは、鉛筆の芯が折れたことである。逆に、もし鉛筆が折れずそのままドリルを解き続けていれば、あなたは（a）に留まっていた。このことから、ある世界から別の世界への移行を可能にするのは、作業の円滑な進行を妨げる要素（たいていは使えない道具）であると考えられる。

なお、道具が壊れたり、また使いたい道具が手元になかったりして、作業が妨げられるとき、世界は、一度（b）のようなよそよそしいものとなる。ハイデガーの言葉では、このとき世界は「前に-立てられた（vor-gestellt）」ものとなる(cf. Heidegger, 2001, S. 72)。ハイデガーにおいて「表象（Vorstellung）」とは、このように前に立てられたもの（Vorgestelltes）であり、没入状態から一歩引いて、反省的に眺められたものである。ちなみに、上で、(b) を体験世界の一つの極と述べたが、実は（b）のような表象された（前に立てられた）世界は、厳密にはもはや体験世界ではない。体験世界が体験世界であるのは、そこにあなたや私が没入しているかぎりにおいてなのである。

さて、ハイデガーにおいて、ある世界から別の世界への移行を可能にするのは、使えない道具だけではない。何らかの目立つ道具が別の世界への移行を可能にすることもある。上の例で言えば、電動鉛筆削り機がこれにあたる。こうした、世界への入り口の役割を果たす、多少とも目立つ道具を、ハイデガーは「記号」と呼ぶ(ibid., S. 80)。記号を見つけ、それを使用することで、使用者はまた新たな体験世界へと没入するのである。

ただし、この記号も、その用途が不明の場合（ハイデガーは例として「ハンカチの結び目」を挙げる）には、単なる使えない道具となってしまうことがある(ibid., S. 81)。結局、ここで登場した 3 種類の道具、すなわち、鉛筆、折れた鉛筆、電動鉛筆削り機は、それぞれ、円滑な作業を可能にする、作業を妨げる、新たな作業へと導く、という機能を持っていたが、それもすべて使用者の目的に相対的なものということになるだろう。

以上で見たハイデガー流の人間と道具の相互作用ループを箇条書きでまとめると、次のようになる。

- 人間は、ある世界から別の世界へと移行することができる。
- この移行は、没入→反省→没入→反省→…、というふうに進行する。
- 没入→反省を可能にするのは、使えない道具である。
- 反省段階においては、世界は表象されている。
- 反省→没入を可能にするのは、目立つ道具（記号）である。
- 各段階での道具の機能は使用者の目的に相対的である。

3.2 ハイデガーの寄与

クラークは、拡張された認知の今後の研究において考えるべき問題を二つ挙げている。

1. 我々をかくも知的にするのはデザイナー環境であるとして、いかなる生物学的差異が、我々にそれらを第一に構築／発見／使用させるのか。(ibid., p. 136, 番号引用者)
2. […]そうした分散されたシステムのさまざまな（生物学的、社会的、人工的）部分の、固有の役割、相互作用的な複雑性の説明、そしてそのシステムが知識産出的であると理解されうるような諸条件の説明[…]。(ibid., p. 141, 番号引用者)

前節で見たハイデガー流のループは、この二つの問題に何か寄与するところがあるだろうか。まず二つ目の問題から考えよう。詳しくは繰り返さないが、前節で述べたように、少なくとも人工物（道具）には、ハイデガーは三つの「固有の役割」を認めている。この三つの役割は、世界の移行ループを成り立たせる上で不可欠のものである。特に注目に値するのは、「没入→反省」という移行を可能にする「使えない道具」が、世界の「表象」をも可能にする、ということである。このことは、ハイデガー流の人間と道具の相互作用ループが、以下のようなクラークの基準を満たす可能性を示唆している。

> 私が提案する、人間レベルの知能を理解しモデル化するための最も見込みのあるルートは、複雑な内的表象（その形式は古典的というよりはコネクショニスト的だが）という考えを真剣に受け止めること、ただしこれを、次のことを特に強調することと結びつけることである。つまり、我々が作り出すきわめて特別な環境との相互作用によって、人間の思考と理性が成形され、高められ、最終的に変形させられる、その固有の仕方の強調である。(Clark, 2001, p. 131)

ハイデガー流のループは、道具との三つの相互作用の仕方であり、そこには不可避的に世界を表象せざるをえない段階が含まれている。もちろん、没入段階において表象はどうなっているのか、といった問題はあるが、作業を妨げる要素が世界の表象を帰結するというテーゼは、一考に値する材料と言えるだろう。

ところで、前節で述べた「人間は世界形成的である」というテーゼは、実は次のような三対のテーゼの一つである。

1. 石は世界を欠いている（weltlos）。
2. 動物は世界が乏しい（weltarm）。
3. 人間は世界形成的である。(Heidegger, 1983, S. 263)

この「動物は世界が乏しい」というテーゼが、動物はあまり多くの世界へと移行することができないということとして解釈できるとすれば、人間と動物の差異は、世界の移行サイクルの頻繁さ、ヴァリエーションの多少によって表現できると考えられる。これはすなわち、人間の方が、道具との三つの相互作用の仕方のヴァリエーションが多い、ということである。また、没入段階に限れば人間と動物はそれほど異ならないと見なせるなら、反省段階、すなわち世界を表象する段階において人間がどのような能力を見せるのか、ということも人間と動物の差異を考えるうえで鍵となってくるだろう。とすると、クラークの第一の問題は、(1) 道具とのハイデガー的な三つの相互作用のヴァリエーションを可能にする生物学的差異、(2) 表象を用いた知的能力を可能にする生物学的差異、というふうに的を絞ることができるのではないだろうか。

そう考えると、NI のプロジェクトのもとでは、ハイデガーを介して、GOFAI と CI をも対立しないかたちで扱える可能性が生じてくる。というのも、ハイデガー流のループにおける没入段階は、CI でモデル化できる性質のものであろうし、反省段階における表象を用いた知的能力とは、まさに GOFAI がターゲットとしていたものにほかならないからだ。

おわりに

本稿は、GOFAI、CI、NI という三つの知能モデルを検討してきた。はじめに述べたように、本稿はそれらに人間と動物の連続性と断絶というテーマを通底させてきた。簡単に振り返ると、GOFAI では、デカルト主義的-認知主義的な、人間に固有の記号操作、表象能力が強調され、人間と動物の間にはこの点で深い断絶がある。CI は、生物学的認知漸進主義であり、ブルックスのクリーチャーのような移動ロボットが、いずれ人間レベルの知能

を実現すると考えている。NI は、知的活動を脳と身体とテクノロジー環境の相互作用のループとしてとらえ、これを人間に固有なものとする。最後に見たハイデガー流の人間と道具の相互作用のループは、NI の実現に向けて、一定の寄与の可能性を示しつつ、大づかみなモデルではあるが、NI の枠組みの中で GOFAI と CI を統合する可能性をも示唆していると言えよう。ハイデガーのこの洞察が何らかのかたちで生かされるかどうかは、このプロジェクトのもとで、認知科学、AI、ロボット工学、哲学など、様々な分野が今後ますます協働できるかどうかにかかっているだろう。

文献

Anderson, M. L. (2003). 'Embodied Cognition: A field guide', *Artificial Intelligence, 149*, 91-130.

Brooks, R. A.(1991). 'Intelligence without representation', *Artificial Intelligence, 47*, 139-59. (1990, 柴田正良訳, 「表象なしの知能」, *現代思想 1990 年3 月号*, 85-105.)

—— (1999). *Cambrian Intelligence: The Early History of the New AI*, MIT Press, Cambridge, MA.

Clark, A. (1996). *Being There*, Cambridge MA: MIT Press.

—— (2001). 'Reasons, Robots and the Extended Mind', *Mind & Language, 16, No.2*, 121-45.

Dreyfus, H. L. (1979). *What Computers Can't Do (Revised edition)*, New York: Harper & Row, New York. (1992, 黒崎政男・村若修訳, 『コンピュータには何ができないか―哲学的人工知能批判』, 産業図書.)

—— (1981). 'From Micro-worlds to Knowledge Representation: AI at an Impasse', in John Haugeland (Eds.), *Mind Design*, Cambridge MA: MIT Press.

Dreyfus, H. L. & Dreyfus, S. (1986). *Mind Over Machine*, New York: The Free Press.

Heidegger, M. (1983). *Die Grundbegriffe der Metaphysik: Welt, Endlichkeit, Einsamkeit*, Gesamtausgabe Bd. 29/30, Frankfurt am Main: Vittorio Klostrmann. (1998, 川原栄峰, セヴェリン・ミュラー訳, 『形而上学の根本書概念　世界-有限性-孤独』, 創文社)

—— (1987). *Zur Bestimmung der Philosophie*, Gesamtausgabe Bd. 56/57, Frankfurt am Main: Vittorio Klostrmann. (1993, 北川東子, エルマー・ヴァインマイアー訳, 『哲学の使命について』, 創文社.)

—— (2001). *Sein und Zeit*, Tübingen: Niemeyer Verlag, 18. Aufl. (2003, 原佑・渡邊二郎訳, 『存在と時間 I -III』, 中央公論新社).

Preston, B. (1993). 'Heidegger and Artificial Intelligence', *Philosophy and Phenomenological Research, 53, No.1*, 43-69.

〔京都大学大学院修士課程・哲学〕