

哲学入門

00 哲学とはどのような営みなのか

「世界には驚くべき現象はなく、私たちは何の疑問も持たず生活できる」と仮定してみよう。実際、私たちは疑問を持って立ち止まることなく生活の大半を過ごしている。電車の通勤途中で「なぜ今日も太陽が昇ったのか」という疑問を持つ代わりに、「今日の仕事はどのようにするか」といった問題の処理を考えている。今日も太陽が昇ったことは多くの人にとって驚くべき現象ではない。無神論者が「神は死んだ」といっても誰も驚かないが、神父が同じ台詞を言ったなら、誰もが驚くだろう。同じように、「明日は太陽が昇らない」と物理学者が言ったなら、誰も驚くだろう。

世界で起こること、社会で起こること、心のなかで起こること、それらのどれかに驚き、「なぜ」という疑問を持ったとき、人は哲学への一步を踏み出すのである。驚き、疑問を持つと、人はそのことに焦点を定め集中して考え出す。すると、平穏で事もない世界は一変し、謎を秘めたものになる。その謎の解明は平明な世界の一部をデフォルメし、真理、実在、説明といった概念を使うことを強いるようになる。謎の解明はそのための装置を必要とするからである。問題の主題化は世界の一部に集中的にスポットライトを当て、それをドラマチックに仕立ててしまう。本質を見極め、それを使って謎を解くという構図は平穏な世界を知的な戦場に変えるのである。

このような戦場での戦士がかつては哲学者と呼ばれていた。しかし、戦場は細分化され、高度の装備がなければ戦えなくなると、万能型の戦士ではなく専門家が必要となる。知的な戦争ではこの専門家が個別科学者と呼ばれている。だが、哲学者と個別科学者には大きな違いがある。謎に満ちた世界の一部を情熱的に解明する個別科学者に対して、哲学者は謎が解けたなら元の平穏な世界が復元することを目指している。ミクロな世界がどのようなものかわかることは私たちが住むマクロな世界が以前と同じであることに反しないことを示そうとするのが哲学の仕事である。

本文の各項目は哲学が今まで考えてきた謎や問題の一部を描いている。それらが人間の知的な歴史のなかでどのように取り上げられ、驚きがどこまで納得に到達したかを解説的でない仕方ですべて述べている。したがって、半ば強制的に問いが出され、推論する形式で解決が図られている。これには読者諸兄の忍耐を求めたい。専門的な問いの背後にあるより基本的な問い、驚きと謎に直結した問いとその解決をできるだけ普通の言葉を使って表現したが、それが誤解を生むかもしれないという危険を孕んでいることも承知しておいてほしい。

項目一覧

01 正しい推論

- 02 形而上学
 - 03 アプリオリ
 - 04 エッセンス
 - 05 意味と指示の分業
 - 06 時間の向き
 - 07 確率とその解釈
 - 08 決定論
 - 09 非決定論
 - 10 非線型性あるいは決定論と非決定論の重複
 - 11 因果連関
 - 12 ゼノンのパラドックスと因果的説明
 - 13 知識論と懐疑
 - 14 基礎付け主義と信頼可能性
 - 15 経験論と実在論
 - 16 主観性と客観性
 - 17 二つの同一性
 - 18 心身二元論
 - 19 二つの行動主義
 - 20 唯物論あるいは物理主義（と付随性）
 - 21 意識と無意識
 - 22 人間の本性
 - 23 伝統と偏見
 - 24 正常と異常（アリストテレスの Normal State Model とダーウィンの Variation Model）
 - 25 ポストモダン
- 行動主義と機能主義の対立から

01 正しい推論

前提から結論を導き出すことを推論するというが、推理小説での賢い刑事の頭の動きを想像したらそれがどのようなものか察することができる。あるいは中学校で図形に関する証明をした際の経験を思い出してもよいだろう。私たち人間が他の動物と異なるのは頭を使って、このような推理・推論する合理的な思考にあると言われてきた。「人間は理性的な動物である」と言ったのはアリストテレスであり、彼は合理的な推理・推論を哲学する際の主要な方法・装置と考えた。そこで、私たちもこの合理的な思考の中枢を担う推論から考えてみよう。

推論は論理的な規則に従って行われる。そこで、まず論理に関して考えてみよう。昔から

論理には三つの形式が考えられてきた。演繹論理、帰納論理、そしてアブダクションである。一定の前提から結論を規則に従って導き出す演繹論理は合理的な思考のまさに中核となっている。一方、多くのカラスが黒い色をしていることから、「カラスはみな黒い」という文が正しいことを導き出すのが帰納的な推論である。しかし、有限のカラスに関する情報から、カラス全体について何かを結論することには飛躍がある。このような飛躍のない演繹論理では推論の規則の組がどのようなものがはっきりしており、したがって、私たちではなく機械にもその規則の組を使って推論させることができるが、この飛躍を埋める帰納的な論理の規則は知られていない。そのため、現在の私たちは確率や統計の手法を使って帰納的な推論を考えている。アブダクションは与えられた結果を説明するための仮説を作り出すことで、発見の論理とも呼ばれてきた。ある結果を説明するもっともらしい仮説を推量することは帰納的な一般化以上に飛躍があるように見える。アブダクションは現在では最善の説明をするための仮説設定と考えられており、その大きな飛躍の存在のために帰納的な推論以上に内的な仕組みはわかっていない。

唯一よくわかっている演繹論理はアリストテレスにまで遡る。彼は演繹論理を三段論法中心に最初に組織化した人であり、驚くべきことに 19 世紀まで彼のシステムは多くの人に学問上の必須の科目として受け入れられてきた。しかし、極めて単純な推論にしか適用できないことから演繹論理は実際の役に立たないとも考えられてきた。その名残が「形式論理」という語の「当たり前で、実際の役に立たない推論」という意味である。三段論法では二つの名詞だけを含む文、例えば、「すべての A は B である」と「ある A は B である」、そしてそれらの否定形だけが前提と結論に許される。そして、例えば、「すべての A は B でない」と「ある A は C である」という二つの前提から、「ある C は B でない」が正しい結論として導き出される。しかし、関係を表す二つの前提「すべての A は B より大きい」、「すべての B は C より大きい」から、「すべての A は C より大きい」という結論を導き出すことは明らかに正しいが、アリストテレスの三段論法ではその正しさが証明できない。

しかし、このような閉塞的な事態は 19 世紀中葉のブール、そして後半のフレーゲによって一新される。演繹論理の正しさと適用範囲の拡張が同時になされ、現在の論理学が成立することになる。アリストテレスのシステムでは証明できなかった推論も簡単に証明できるようになったのは言うまでもない。

この論理学の革新は論理的な規則の集まりを抽象的なシステムとして記号言語を使って明解に表現した点にある。私たちの日常言語はその豊かな表現能力のためにしばしば論理的な明晰さを犠牲にするし、文の表現は文法の規則に従わなければならない。この二つの点を克服するには日常言語から離れ、明晰さと論理規則が直裁に反映される人工的な記号言語を使った方が得策である。この人工言語は全くの新規のものというより、既に使われていた数学での記法を生かし、それを普遍化したものである。したがって、ある程度は日常言語の表現に対応しており、中学や高校で習った数学の定理はこの記号言語を使って表現するとよりその内容が明確になる。こうして、フレーゲによって再構築された論理のシ

STEMは数学や哲学の研究の装置として使われ、また言語学やコンピュータ科学に応用され、20世紀の科学の一特徴である記号を駆使した対象の把握を可能にしてくれた。特に、20世紀の哲学、数学の特徴はこの論理学の結果によるところが大きい。言語論的転回、集合論による抽象数学はこのような特徴を示す動向、成果である。

記号言語を使うと明晰になる例を一つ挙げてみよう。「どんな人にも好きな人がいる」と「どんな人にも好かれる人がいる」の違いは何かと聞かれたらどのように答えるだろうか。すぐに説明できる人もいるだろうが、そうでない人はどのように考えたらよいだろうか。この二つの文をそれぞれ言い直してみると、

「すべての人 x について、その人 x には好きな人 y が少なくとも一人はいる」

「ある人 x が少なくとも一人いて、その人 x はすべての人 y に好かれる」

となる。この言い直しだけで相当はっきり違いが出てくるが、さらに、それぞれ記号化してみると、

$$x(F(x) \quad y(F(y) \quad G(x, y))$$

$$x(F(x) \quad y(F(y) \quad G(y, x))$$

となり、簡略化すると、 $x \quad yG(x, y)$ 、 $y \quad xG(x, y)$ となる。 $x \quad y$ と $y \quad x$ の記号の並び方の違いが二つの文のもつ意味の違いを反映していることがわかる。これが日常言語の表現では見えにくい論理関係が記号言語を使って明瞭になる一例である。

記号言語を使って曖昧であった哲学や科学の問題を正しく把握することで、正しい解答への距離が多いに近くなる。また、問題だと思っていたものが擬似の問題であることがわかり、問題が解消する。さらには積極的に問題を攻撃する手段として使う。このように論理学の成果を使って様々な成果が上がってきた。それらを理解することも重要であるが、ここではその根幹にある正しい推論の重要性を認識してほしい。私たちの周りには一見正しそうな推論や推量が溢れている。合理的な精神はそのような見かけの姿の背後にあるものを鋭く見極めるためにある。実際に目にする誤謬の大半は純粋に論理的な誤謬ではなく、私たちの知識や経験と論理がもつれ合った誤謬である。論理だけの誤りは意外に見極めがつくが、それが経験的な知識と結びついた場合、私たちの合理的な追求はその内容に気をとられ誤謬を見逃し易い。経験的な内容と絡み合った例として(後で使う)「デカルトの推論」と「総計の誤謬」を考えてみよう。

次の二つの推論についてそれらが正しいかどうか考えてみよう。

デカルトは心をもっている。

ロボットは心をもっていない。

それゆえ、デカルトとロボットは同じではない。

デカルトは彼が心をもつことを疑うことができない。

デカルトは彼が脳をもつことを疑うことができる。

それゆえ、心と脳は同じではない。

上の推論は文句なく正しいが、下の推論は正しいだろうか。それが正しくないことは次の類似の推論が参考になるだろう。明らかに $2+2$ は $6-2$ と同じであり、したがって、結論は誤っている。二つの推論の違いは前者が事実に関する文からなっているのに対し、後者が事実についての心的な態度（疑う）を表す文からなっている点にある。

デカルトは $2+2=4$ を疑うことができない。

デカルトは $6-2=4$ を疑うことができる。

それゆえ、 $2+2$ は $6-2$ と同じではない。

次の例は、シンプソンのパラドックスと呼ばれている誤謬の、印象的な例はカリフォルニア大学で行なわれた入学試験についての男女差別の疑いである。男女同数の受験者に対して、合格者全体を比べると男のほうが女より多かった。これは男女差別ではないかという疑いがかかり裁判沙汰に及んだ。大学当局が学部ごとに調べ直してみると、二つの学部はいずれも男女の合格者数に関して全く公平であることがわかった。（下の表はこれをわかりやすくしたもので、実際の学部や学生数ではない。）

	学部 1	学部 2	総計
応募者	90女；10男	10女；90男	100女；100男
合格率	30%	60%	
合格者	27女；3男	6女；54男	33女；57男

この表は簡単な数に直してあるが、総計を見ていただきたい。確かに男女の応募者数は同数でありながら、合格者数には差がある。しかし、学部ごとの応募者と合格者はどうであろうか。学部1も学部2も共に応募者の男女比にあった合格者を出している。つまり、各学部は男女差別の内容に配慮し合格者を出したのであるが、応募者数の違いのために総計ではあたかも男女差別があったかのような結果になったのである。総計は各部分の性質を正しく反映してくれない。このような誤謬の原因は論理的なものではなく、統計の初歩の認識にある。しかし、それを適切に指摘し、正しい姿を浮き彫りにするシナリオは論理的な構成なしにはできない。論理的な規則は道具であり、その道具は使ってこそ役立つのである。論理的に正しい推論はその適用される状況に正しく適合してこそ、正しいだけでない。

い信頼できる推論として認められるのである。

02 形而上学

形而上学は哲学の分野の一つで、実在や自然の本性を研究してきた。現象の背後にある究極の実在を明らかにしようとしてきた。その第一の研究対象は存在するもの全体であり、世界を構成する基本的なものを明らかにしようとしてきた。このように表現されるのが形而上学の辞書的な説明であるが、これで「わかった」と合点する人はいないだろう。過去形で述べられた形而上学が現在でも同じ研究対象、研究態度を保存しているのかといった疑問がすぐに浮かぶだろう。

私たちは論理学の開祖としてアリストテレスに言及した。そのアリストテレスの名を有名にしている今一つのものが形而上学である。形而上学を英語で書くと Metaphysics である。物理学は英語では Physics である。「メタ」という表現は最近の日本語ではよく登場するが、「~の後に、次に」という意味である。Meta-physics は、したがって、「物理学の後で研究するもの」という意味になる。実際、アリストテレスの形而上学は自然についての個々の知識を習得した後で、自然の基本的な本性について一般的に研究したものであった。このような語源的な説明で形而上学が何かがわかるものではないが、その歴史は自然に関する哲学と深く結びついたものであった。この自然哲学の傾向はニュートンの物理学を通じて物理学の基礎的な概念の追求となって現在にまで続いている。また、アリストテレスの形而上学のシステムがカトリック神学に取り入れられたため、神と自然や人間の関係についての一般的な関係の考察にも多くの研究が費やされてきた。そのような研究の代表例は神の存在証明である。そこで、その存在証明の一例を考えてみよう。

世界の出来事が無限にないことを仮定するなら、どのような出来事にも原因があり、その原因にはまた別の原因があるという具合に、原因を遡及する系列が考えられる。この系列は無限に遡及できないことから、それ自身では他のものによって引き起こされない第一原因がなければならないことになる。この第一原因は世界の中にあるものではなく、世界の中の出来事を引き起こすのであるから、そのような第一原因を性質としてもつものがなければならない。それが神である。

この推論は正しいだろうか。世界の出来事は無限に続いてはいけないのだろうか。現在の私たちには無限概念は驚くべき概念でも恐れるべき概念でもない。したがって、最初の前提は多に疑いの余地がある。さらに、「第一原因である」という性質から、そのような性質をもつものが存在するという推論も受け入れがたい。確かに、中世は無限概念を嫌い、恐れたが、だからといって無限概念を消し去るわけにはいかない。

ところで、「無限」とはどのようなものか。カントールによって明らかにされた無限概念は集合論という 20 世紀数学の基礎理論を生み出すことになった。「無限に分割する」、「無限に大きい」といった表現に正確で、矛盾のない意味を与えることは、自然数や実数という数学的な対象を正しく把握することになり、自然数や実数を基礎とする数学を確立する

ことになるとカントールは考えた。その結果、現在では公理的な集合論ができあがり、数学の基礎理論として使われている。能書きはこのくらいにして、「無限」に触れてみよう。自然数をまず考えてほしい。自然数はいくつあるか。それはどのように証明できるのか。これらの問題はそれほど厄介ではない。自然数が有限で、したがって、その中に最大のものがあつたとしてみよう。それを n とすると、 n に 1 を加えて自然数がつくられることから、 $n+1$ という自然数があることになる。すると、 n は最大であつたにもかかわらず、 $n < n+1$ であり、これは矛盾である。それゆえ、最大の自然数があるという最初の仮定が誤りであり、最大の自然数は存在せず、自然数は有限ではないことになる。(当然ながら、自然数はその大きさに関して単調に並んでいて、循環しない。)つまり、自然数の個数は無限である。では、実数はどうか。自然数は実数の一部であり、その一部が無限なのであるから、当然実数の個数も無限であると考えることができる。では、同じサイズの無限なのか、それとも異なるサイズなのか。カントールは対角線論法と呼ばれる手法によって実数の無限のサイズが自然数のそれより大きいことを示したが、ここでは直観的に次のことで理解してほしい。自然数は並べることができる。小さい方から順に並べると言われれば、0、1、2、...と並べていくことができる。しかし、実数をこのように並べることができるだろうか。0 から始めて、次に大きい実数は何か。私たちに実際はそれが何かを言うことができない。そのような数が存在することは証明できてもそれが何かはわからない。自然数の無限は番号をつけることができるが、実数は連続しており、可付番ではない。これで少なくとも 2 種類の無限があることがわかった。実際、無限の種類は無限にある。また、自然数と実数の間に別の無限があるかどうかはわからない。

私たちは神の存在証明から無限に話を転じたが、自然哲学と結びついた形而上学にはどのような問題があるのだろうか。それら問題は物理学の基礎に結びついたものが多いと述べたが、そのような代表的な問題を列挙すれば次のようなものがある。

実在、自然法則、時間、空間、因果性、時間の向き、決定論

これらのいくつかは後に議論することにするが、因果性や決定論を考える上で基本的な役割を演じる「ならば」という言い回しについて先に触れておこう。因果作用は原因と結果からなるが、それを「 A ならば、 B である」と表現している。その際、原因 A は結果 B の十分条件である必要はない。また、必要条件である必要もない。つまり、原因、結果と論理的な前提、帰結の関係は類似していても基本的に異なったものである。私たちは前提と帰結の論理関係についてはある程度知っているが、原因と結果の因果関係については経験的な確証が必要である。喫煙と癌の因果関係はいくら頭で考えてもわかるものではない。

「 A ならば、 B である」という表現は単純であるが、原因 - 結果と前提 - 帰結の二つの(根本的に異なる)関係を二重に意味している。それを次の例で実感してほしい。

- (1) $x + y = z$ ならば、 $2z = x + y + z$ である。
 (2) 伊作が怒るならば、史門が泣く。

文(1)の「ならば」は論理的な「ならば」であり、前提 $x + y = z$ と帰結 $2z = x + y + z$ の含意関係を主張している。(実際、変数が自然数や実数であれば(1)は正しい文であり、含意関係は成立している。) それに対して、文(2)の「ならば」は因果的な「ならば」で、伊作のある状態と史門のある状態の間に因果的な関係があることを主張している。二つの「ならば」の違いは極めて重要である。例えば、(1)の前提と帰結はそれらがいつ成立するかは考慮されないが、(2)の二つの状態は時間的な制約を受けている。伊作が先に怒り、次に史門が泣くのでなければ、因果関係は成立していない。日常的な表現である「ならば」が論理的、因果的の二つの意味を併せもつことは日本語だけの偶然的な特徴ではない。英語でも「if then」は二義的に使われている。

このような「ならば」の二つの意味は日常生活だけではなく、哲学や科学の研究でも同じように使われている。物理学の理論とそれが扱う物理的な世界を考えてみるとそれが鮮明になる。例えば、力学は数学を使って表現されている。例えば、運動方程式は論理的な「ならば」を使って変形され、解が見つけれられる。一方、そのような運動方程式によって記述される物理世界の変化は因果的な変化であり、その変化は因果的な「ならば」で表現される。数学で使われる論理的な「ならば」が物理世界を表し、解釈するには因果的な「ならば」にすりかえられる。数学が物理学に役立つのは、私たちがこれら二つの「ならば」の共通部分を巧みに使い分け、相互の関係をつけているからである。だが、概念上、二つの「ならば」は全く異なったものである。

03 アプリオリ

トートロジー(tautology)という語を聞いたことがないだろうか。「今日は晴れているか、あるいは晴れていないかである」という文は「今日は晴れている」と「今日は晴れていない」が「あるいは」という接続詞で結ばれており、「 P あるいは P でない」という形をしている。この文は有用な情報を何も伝えてくれないが、誤まってははいない。それどころかいつでも(つまらない意味で)真である。そのような文を聞いても誰も知識や情報を得たとは思わない。このようにその論理的な形だけから真になる文には「 P かつ P でないことはない」や「 P ならば、 P 」がある。このようないつでも真である文は狭い意味でのトートロジーであり、それは論理的にいつでも真であるという性質をもっている。

いつでも真になる文には、例えば「すべての独身者は未婚である」という文がある。この文は狭い意味のトートロジーではない。単なる同語反復ではなく、何がしかの情報を含んでいるように見えるからである。しかし、独身と未婚が同義であることを思い出すなら、同じものを代入しても結果は同じであるという原則にしたがって、「すべての独身者は独身である」というトートロジーが得られる。このような広義のトートロジーは哲学では分

析的な文と呼ばれ、独身という語を未婚と定義することによって、独身 = 未婚が成立し、したがって、その定義だけから「すべての独身者は未婚である」が真になるような文を意味している。そして、分析的でない文は総合的な文と呼ばれている。定義だけからいつでも真になる文が分析的、つまりは広義のトートロジーである。

ところで、トートロジーや分析的な文は実際に実験や観察によって確かめる必要なく真であることから、私たちの経験に頼って真偽を決める必要がない。一方、分析的でない総合的な文、例えば「小淵首相の後継者の後継者は A である」は普通の人には予め確信をもってその真偽を言うことができない。そこで、分析的な文の真偽のように経験の介在を必要とせずその真偽がわかる文の内容をアприオリな知識、総合的な文の真偽のように経験を必要とする文の内容をアポステリオリな知識と区別することになった。これは分析的、総合的が言語レベルでの区別であったのに対し、私たちがどのように知識を獲得するかという認識レベルでの区別になっている。

さらに、トートロジーや分析的な文はいつでも真であり、その内容は必然性をもっているように見えるが、総合的な文の内容は偶然的で、世界の状況に応じて真になったり偽になったりするように見える。そこで、文の形式的な区別から文が指示する事態が必然的、偶然的と分けられることになる。いつでも必ず真である事態が必然的、そうでない事態が偶然的であり、これは存在レベルの分類である。

こうして、分析的 - 総合的、アприオリ - アポステリオリ、必然的 - 偶然的という関連する区別が考えられることになる。分析的 = アприオリ = 必然的、総合的 = アポステリオリ = 偶然的という等式が成立すれば、すべてはすっきりしていて、問題は生じない。しかし、密接な対応関係はあるがそれらが微妙に一致しない点に問題が出てくる。

カントはアприオリな真理が二つの領域に見出されると考えた。その領域とは数学と経験を組織化するカテゴリーとの二つである。そして、アприオリな真理をさらに総合的、分析的の二つのカテゴリーに分ける。伝統的には数学的命題は分析的でアприオリとみなされてきた。しかし、カントは数学とカテゴリーの両方とも総合的でアприオリと分類した。数学の命題が総合的でアприオリなのはそれが時間と空間の直観に依存するからである。また、カテゴリーが総合的でアприオリなのはそれらの否定が矛盾を引き起こさないからである。以下、それぞれの代表例であるユークリッド幾何学と因果的な決定論がカントの言うようにアприオリかどうか考えてみよう。

ある文 H が定義上真で、経験的な証拠なしに正当化できる、つまりアприオリであることをどのように示したらよいのだろうか。どのような観察も文 H を反証できないように見える場合、通常は最初から文 H が真にアприオリとは考えないで、私たちの想像力が欠けていて適切な経験的証拠を見出せないと考えるのではないか。例えば、「時間に向きがある」あるいは「過去から未来への時間的な変化があり、その逆はない」という文についての物理学的な証拠は見出しにくい。しかし、時間の向きについて考えている物理学者は時間の向きがアприオリであるとは思っておらず、単に自分の想像力が欠けているため解決でき

ないと考えているだろう。そのような物理学者は次のような工夫をするのではないか。文 H だけではなく、文 H と他の経験的な文の集まりを使って、文 H だけからは導き出せないような経験的な文が得られ、それが文 H と違う内容を主張していたとすれば、文 H の真偽の判定に参考になり、そこから文 H が経験的な主張でないということが何を意味しているかわかる。この工夫をカントの場合に使ってみよう。

既述のように、カントはユークリッド幾何学と因果的な決定論をアприオリに真だと考え、どのような観察も二つの反証にはならないと信じていた。しかし、今世紀ユークリッド幾何学は相対性理論と結びつくと、誤った予測をすることが発見された。因果決定論も同様に、それが量子力学と結びつくと誤った予測を生み出してしまうことがわかった。上の H はここではユークリッド幾何学、あるいは因果的な決定論である。 H が誤りを生み出す理由は相対性理論と量子力学の理論（これらは経験的な主張である）にある。実際、相対性理論では非ユークリッド幾何学が、量子力学では非決定論が成立しており、 H とは異なる内容を主張している。異なるだけでなく、ユークリッド幾何学と相対性理論、因果的決定論と量子力学は両立しない。相対性理論や量子力学を正しいとする限り、ユークリッド幾何学も決定論もアポステリオリに偽であることになる。このような結果を一部使うことによって、クワインはアприオリあるいは分析的な文が存在することを否定した。

04 エッセンス

科学が哲学にとって興味深い一つの理由は、科学がしばしば哲学の思考カテゴリーを覆すからである。既に見たように、相対性理論と量子力学はそれ以前の哲学者がアприオリな真理とみなしていた原理を否定した。カントは空間がユークリッド的であり、決定論が正しいと主張しただけではなく、それら原理は経験が可能であるために必要なものとも考えていた。これら原理は私たちの経験を組織化するのに使われるので、どのような経験もそれらに矛盾できない。そこから物理理論もこれら原理に矛盾できないと考えた。だが、それら原理はアприオリに真であるどころか、アポステリオリに偽であることがわかった。

生物学における分類学は伝統的に生命世界の基本的な種類を記述するものとみなされていた。生物種(species)は種類(kinds)と同義語であった。しかし、近年の進化論は種が種類ではなく、個体(individuals)であることを示唆してきた。種が自然種である、種が個体であるという主張にはどのような違いがあるのだろうか。基本的な違いは性質とその性質をもつ対象の間にある。赤は赤い靴や赤い服とは違う。ものの種類とその種類のものは違う。ある種類はそのメンバーがないかもしれないし、別の種類は多くのメンバーをもつかもれない。普遍と個別、タイプとトークンという対についても同じことが言える。例えば金を例に考えてみよう。これは物質の種類である。私の結婚指輪とペン先は同じ種類（普遍、タイプ）に属している。いずれも原子番号 79 の原子からできているからである。ここで重要なのは、結婚指輪もペン先も共通の祖先をもっている、互いに因果的に相互作用したということが何も求められていない点である。二つのものが同じ種類に属しているのはそれ

らの類似性からであって、それらの歴史からではない。

本質主義は自然種についての標準的な哲学的見解であり、各自然種はその種に属するメンバーだけがもつ性質によって定義できるというのがその主張である。すべての金は原子番号 79 であり、金だけが原子番号 79 をもつ。すべての金は質量をもつが、質量をもつことは金だけの性質ではない。自然種はそのメンバーであるのに必要十分な性質によって特徴づけられることになる。本質主義は次のような一般化を金の本質を述べるものとする。

どのような対象 x についても、

x が金である $\iff x$ は原子番号 79 の原子からできている。

この一般化が本質主義の主張を満たしているかどうか別のものと比較してみよう。世界が有限の金の対象を含んでおり、その金の対象の時空的な場所を (L_1, L_2, \dots, L_n) と特定できるとしてみよう。すると金についての以下のような一般化が得られる。

どのような対象 x についても、

x は金である $\iff x$ は L_1 あるいは L_2 あるいは...あるいは L_n に見出される。

いずれの一般化も真であるが、本質主義者は最初のものだけを金の本質を特定するものとするだろう。これには幾つかの理由がある。まず、金の本質は偶然的でない一般化によって与えられなければならない。金は特定の場所に見出される必要はなく、同じ原子番号 79 をもっていなければならない。第二に、本質は説明的でなければならない。原子番号は金のもつ他の多くの性質を説明するのに役立つ。しかし、金の場所はそれ以外のものを説明してくれない。最後に、金の本質主義的定義は金に内在的な性質に言及している。この性質は、金の置かれた場所が他との関係によって決まるのに対して、そのような他との関係を要求しない。

本質主義は化学元素については妥当に見える考えである。では、生物種についてはどうか。つまり、私たちは次のような形で種（ホモ・サピエンス）を考えることができるだろうか。

どのような対象 x についても、

x がホモ・サピエンスのメンバーである $\iff x$ は Homo Sapiens である。

求められているのは、対象がホモ・サピエンスのメンバーである内在的な性質を特定する一般化だけではなく、必然的で説明的な一般化である。人間の本性は金の本性を本質主義者が概念化するのと同じ仕方で概念化されるのだろうか。種に関する見解として進化論が本質主義を拒絶すると簡単に思いがちである。しかし、これには注意が必要である。進化論が本質主義を否定する推論の例を以下に挙げてみよう。

自然種は不変である。

種は進化する。

よって、種は自然種ではない。

最初の前提は何を意味しているのか。自然種のメンバーは変化するかもしれないが、種それぞれ自体は変化しない性質をもっている。結婚指輪は金貨の材料になり変化できるが、金はその性質を変えない。最初の前提がこのような意味だとすると、この推論は正しくないことがわかる。元素の変換は可能である。元素を破壊することで鉛を金に変えることができる。しかし、これは化学元素が不変の性質をもつという考えを否定するものではない。同じように、ある種に属する集団が別の種の集団になるという事実は種についての本質主義を拒絶するのではない。本質主義は種を個体が属する永遠のカテゴリーであるとみなす。進化は祖先とその子孫が時々異なるカテゴリーに属することを意味している。

このような簡単な推論では本質主義を打倒することはできない。進化論が本質主義に反対することを示すにはカテゴリーそのものが変化することを種分化(speciation)を使って詳細に検討しなければならない。

05 意味と指示の分業

意味と指示は密接に結びつき、そのことによって言語と世界の関係が保証されていると考えられてきた。それゆえ、言語の哲学はこのような意味と指示の結びつきをめぐって展開されてきた。

言語哲学はフレーゲに始まる。彼は言語、論理、そしてそれらの世界との関係について多くを述べている。ここでは「意味とは指示の仕方である」という彼の考えと、それに基づくラッセルの記述理論が伝統をつくりだしたと考えてみよう。すると、指示対象は文(の束)によって表現され、そこで用いられる文の内容の違いが意味の違いを引き起こすことになる。指示内容は語と対象の間の対応関係として考えられ、その関係が文の真偽を決定することになる。このような枠組みでは、指示対象は特定の言語的な表現をもつことはなく、基本的な語とその組み合わせによる文の集まりによって記述的に表現される。したがって、固有名詞は文の集まりの省略表現ということになる。

固有名詞は日常表現にはあふれているが、原理的には消去可能なものであり、科学理論は原則としてこのような固有名詞を含まない。これが科学理論や知識の一般性、普遍性を保証することになる。固有名を担う存在は変化し、消滅するものがほとんどであるが、それが省略表現に過ぎないことによって、形式的には理論的主張はすべて一般的な主張ということになる。いま一つの大きな特徴は、指示は記述を通じて決定されるという点である。観測や実験、その他どのような経験的な道具立てで指示を決定しようと構わないが、最終的に指示を決定するのは記述的な文の集まりである。

このような伝統的な見解は、フレーゲやラッセルに顕著に見られるように論理学と形式言語に強く影響を受けて出来上がったものである。形式言語によって日常言語を解釈した結果といっても言い過ぎではない。

意味が指示を決定し、指示は記述的に与えられるという伝統的な見解への反対はクリプキやパトナムによって主張された。固有名詞がその攻撃の突破口となった。固有名詞は果たして記述的な文の集まりに置き換えることができるのか。ここで力を発揮したのが可能世界という様相論理学のモデル構成に用いられた道具立てである。どのような可能世界においても同じ対象を指示する固有名詞が可能世界に応じてその真理値を変える文の集まりと同値ではありえない。これがクリプキの反対の骨子である。このような考えをさらに自然種(natural kind)を表現する名辞にも一般化したのがパトナムであった。有名な双子地球(Twin Earth)の例は、意味が内在的ではなく、外在的に決定されるという主張を含んでいた。こうして、反フレーゲ的な意味と指示に関する立場は、固有名詞や自然種の名辞が記述の束によって置き換えることができないという主張と、意味は内在的でなく外在的に決まるという主張を展開することになった。

ここで注意しておかなければならないのは、この立場においても支持されているのは意味と指示の密接な関係である。二つが意味論の基本的な概念であることには何の変更もない。この点では伝統的な立場と同じである。意味と指示は記述によって結びつき、固有名は文の省略に過ぎないというのがフレーゲ的な伝統的立場なら、意味と指示は因果関係によって結びつき、固有名はその導入の由来によって意味を与えられるというのが反フレーゲ派の立場である。

表象や概念が心に内在するという、心の内在主義を徹底するならば脳状態論が代表例になるだろう。私たちの表象は徹底して外在的で、「何かについての表象」という表現はこの点を見事に表している。その何か、つまり、表象内容は外部世界に存在する何かであって、痛みといった感覚的なものを除いては外在的である。自己意識さえ自己の思考内容をあたかも外在的な対象であるかのように意識することで成立している。表象は直接の接近が可能と言われるが、眼前の机を見るときに直接知覚と比較して特に直接的というわけではない。このような観点から、内在主義はもっと弱められた主張として理解すべきである。私たちがもつ知識と信念・欲求は一つの独立したシステムをもっており、それが外部からの情報を組織的に扱う、処理している。このようなシステムが自律的であることを内在主義の主張と考えるならば、比較的認めやすい主張となるだろう。

では、内在主義と外在主義はいずれが真相を正しく捉えているのか。いずれもがある側面に関してはより正しいと言うことができる。知識を獲得し、修正し、それを疑っている場合には外在主義的でなければ正しい結果に到達することはできない。これが特に顕著なのは経験的な知識の場合である。一方、内在主義が正しいのは、知識を使って何かを実行する場面においてである。ある定まった意味システムを使って現象や対象に係わるには意味システムが不変で、安定したものでなければならない。このような均衡を保った意味シス

テムであるためには内在化され、信頼できる装置としての確に働くものでなければならない。

このように考えてくると、意味の生成や修正には外在主義が、意味の使用には内在主義が真相に近いことを述べていることがわかるだけでなく、二つの考えの違いも明瞭になる。

伝統的であれ、非伝統的であれ、意味と指示が密接に結びついたものであるという主張は奇妙な帰結をもたらすことを示してみよう。

(意味は新しい知識に関して感受性をもっていない)

内在主義を文字通りに受け取って、ある与えられた状況において誰かがその内在的な意味を用いて事態を知ろうとした場合、その人は既知の事態に対しては既存の意味のシステムをうまく適用できるだろうが、既存の意味システムを使うことだけによって未知の事態にどのように対処できるだろうか。指示対象が未知の場合、内在的な意味システムを使うことは問題を解決するには余りに道具不足である。既存の意味システムは新しいパラダイムを供給できない。むしろ、このような新しいパラダイムを供給できない点にこそ内在主義の自律的な意味システムの主張の根幹があるのであり、意味のシフトを未然に防ぐ点にこそ意味システムの意義がある。これが内在的な意味システムが新しい知識に関して感受性をもっていないということである。

(意味論は科学の研究と区別できない)

最初の奇妙さが内在主義から帰結したのと対照的にこれは外在主義から帰結する奇妙さである。パトナムの双子地球の話でも現れるが、 H_2O 、XYZ がそれぞれ何かは意味システムのどこを探ってもわかるものではなく、端的に科学的な追求の結果である。意味は科学研究の結果として得られるのである。指示の決定は外部環境だけではなく、そのような外部環境に関する科学的な探求も含まれている。意味が科学研究の一部である指示の決定に左右されるなら、意味論は科学研究からどのように区別できるのか。意味システムは意味を構成する場合の知識、特に指示対象の決定に関しては意味論以外のものにほとんどすべてを依存している。

これら二つの奇妙な点は意味と指示の分離によって避けることができる。それを次に考えてみよう。既に、内在主義と外在主義の違いで述べたように、二つが異なる局面に適用されるなら何ら不都合を生じないことがわかった。これは二つの考えの分業を意味している。これと同じように意味と指示に関しても、分業が成立するように考えるのが妥当ではないか。では、意味と指示の分業はどのようにして成立するのか。

意味システムは均衡的なシステムであり、その点では物理的、生物的なシステムと同じように、均衡を保つことが存在するための最低条件になっている。その最低限の条件は無矛盾性である。しかし、そのようなシステムは外部の世界と独立に存在しているのではなく、常に相互作用しており、それによって不断に意味内容が更新されるようになっている。こ

れは生物が生存することによく似ている。意味システムが世界に適合し、それによって私たちが適切に振舞うことができるには、システムの保持と変容の両方が必要である。この異なる特徴を満たすために意味と指示の分業が必要であるし、実際私たちはそのような分業を言語を使い始めて以来行ってきたのである。なぜ、このような簡単なことが不明のまま放置されてきたのだろうか。それは意味と指示が余りにうまく分業を行い、あたかも一つの言語的な事柄の二つの側面といった具合に結び合っていたためである。さらに、新しい概念の構成とその解釈や使用は哲学にとっては切っても切り離せない組み合わせであり、それらを分離したのでは解ける問題も解けなくなると考えられてきた。このことが分業をさらに不明にしたのではないだろうか。

指示対象の本性がどのようなものであるかがそれを表現する語の意味を決定するが、そのために意味論に頼る者がいるだろうか。酸素原子の構造を本格的に調べようとするものは国語辞典を調べるといったことはしない。しかし、「酸素原子」という語を正しく遣いたい者は国語辞典に頼るかもしれない。ここには意味に関する分業がうかがえる。というのも、意味を使うには意味をどのように構成したかはいつも必要なわけではないからである。

意味システムは、したがって、意味の辞書とその辞書づくりの二つが異なる内容と原理に従っていることになる。辞書は自らの自律性を保つことがその使命であり、徒に右往左往するわけには行かない。一方、辞書づくりは自律的なシステムをつくりだすのであるから、できるだけ使用が正しく行われることを目的として行われる。その際には多に新しい知識を活用しなければならないし、場合によっては自ら創り出さなければならない。私たちは意味から指示へのシフト、指示から意味へのシフトを意識せずに行っている。ここには議論の対象と議論の装置の違いが指示と意味を介して存在している。

06 時間の向き

哲学では時間について二つの異なる考えがある。時は流れ、現在は世界の客観的な特徴であるとするグループと、時の流れや時制は単に主観的なものに過ぎないと考えるグループである。以下のモデルはプラトンの静的世界からどのようにして動的なアルキメデス的世界がつくり出され、そこには本来的に方向がないことを直観的に示すものである。

トランプカードを一組選び、そこからスペードのカード 13 枚を抜き取る。各カードには 1 から 13 までの数字が書かれており、その中から任意に 1 枚抜き取る。例えば、3 のカードが無作為に抜き取られたとしよう。以後の操作は次の規則に従うものとする。

<規則 A>

- (1) 与えられたカードの数字を見て、その数字に最も近い数字のカードを探す。(ただし、13 に最も近いのは 12 と 1 とする。)
- (2) 最も近い数字のカードが一枚だけなら、それを与えられたカードの右側に並べる。
- (3) 最も近いカードが二枚以上あるなら、無作為に一枚選ぶ。そして、それを与えられた

カードの右側に並べる。

- (4) 右側に並べられたカードに対して、(1)と、(2)あるいは(3)の操作を条件に合う手許のカードがなくなるまで続ける。

さて、スペードの13枚のカードについて、3のカードから出発した場合、どのような数字をもったカードの系列ができあがるだろうか。

3、4、5、6、7、8、9、10、11、12、13、1、2

3、2、1、13、12、11、10、9、8、7、6、5、4

上の二つの系列が答えである。これ以外にはない。トランプ全体についても考えることができる。規則は同じものを使う。今度は(3)の無作為の抽出がカードを選ぶごとに行われなければならない。(規則(2)は実際上使わないことになる。)スペードの3から始めよう。もしカードの種類をあえて無視するなら、できあがる系列の数字の並びは何と上の二種類の系列と同じになるはずである。カードの4種類それぞれについて、3、4、...となるので、上の二つの系列それぞれについて、 4^{12} 通りの系列がつけられることになる。手許には条件に合わないカードが残るが、スペードだけの場合と同じで、並ぶ数字の列は2種類である。

さて、今までは3のカードから出発することだけを考えたが、他の数字の場合はどうなるか。出発の数字が異なるだけで、規則は同じように適用できる。例えば、7、8、9、.....、あるいは、7、6、5、.....、が得られる。しかし、 4^{12} 種類もの系列がつけられる規則というのは規則としては感心できない。そこで、的確な系列が得られるように規則をさらに厳格にできないだろうか。<規則A>を厳格にしたのが次の規則である。この規則はもっぱらトランプ全体に対して適用される。

<規則B>

- (1) 与えられたカードの数字を見て、その数字に最も近い数字のカードを探す。(ただし、13に最も近いのは12と1とする。)
- (2) 最も近い数字のカードが一枚だけなら、そのカードを与えられたカードの右側に並べる。(*カード全体ならこの規則の適用は実際にはない。)
- (3) 最も近いカードが二枚以上あるなら、数字以外の条件を探し、その条件のもとで最も近いカードを選ぶ。ただし、一度選ばれた数字の異なる種類のカードは選ばない。どのような条件に対しても、それらを満たすカードが二枚以上あるなら、無作為に一枚選ぶ。そして、それを与えられたカードの右側に並べる。(*カードの場合、数字以外の有力な条件は種類である。最も近い数字と同じ種類の二つの条件で一枚カードが決定する。)
- (4) 右側に並べられたカードに対して、(1)と、(2)あるいは(3)の操作を条件に合う手許のカードがなくなるまで続ける。

ではこの規則のもとでトランプ全体の系列づくりはどうなるだろうか。今度はトランプ全体であっても、スペード一種類の場合と同じように、

3、4、5、6、7、8、9、10、11、12、13、1、2

3、2、1、13、12、11、10、9、8、7、6、5、4

の二種類の系列が定まる。

<規則 A> は系列上の数字が連続するように求めており、<規則 B> は系列をつくるカードができれば 1 枚選ばれることを求めている。これらをそれぞれ、連続性の規則、線形性の規則と呼ぶことができる。どうしても 1 枚選出できない場合、それは非線形性の規則となる。ここで、二つの系列のいずれかを優先的に選び出すことができるかという問題を考えてみよう。そのような条件はいくらでも考えられる。例えば、最初に選んだ 3 より大きな数字のカードで最も近いものを選べという条件である。(13 が最初の場合にはさらに条件を加えなければならない。) しかし、それら条件は上述の規則の条件に比べると *ad hoc* なものである。少なくとも、連続性や線形性のように数の本来の性質から導き出されるものではない。このことを数学的に特徴づけるよりは次の直観的な理由で理解しよう。

目の前にある線分をみて、それと同じものをノートに写す際、左から線を引くことも、右から線を引くこともでき、しかもできあがりは同じである。いずれの場合も私たちは連続的に線を引く(連続性)。途中でどこに向かうかの選択を迫られることもない(線形性)。しかし、右と左のいずれから線を引くべきかという条件は、付け加えることはできるが、それで得られるものは何もない。なぜなら、引かれた線はどちらから引こうと同じ線になるからである。そのような条件は線分を引き写すには単に余計なのである。

この簡単なモデルのカードのそれぞれをプラトンの世界、構成される系列をアルキメデスの世界と考えよう。各カードは変化のない対象であり、そこから運動や変化をどのように構成するかが上のモデルでは系列の構成によって表現されている。このような構成が数学的に世界を考え、変化がどのように表現されるかの骨組である。

(リングモデルの構成)

以前のトランプモデルから次のようなリングモデルを作ろう。並べられた 13 枚のカードそのままの順序で輪にする。カードの数字は時計の文字盤のように順序を表す名前として使うことができる。(実際、1 から 13 までの数字でつくられる時計の文字盤を想像し、かつ 1 から 13 までの数字がその中のどれかを出発点として並んでいると考えればよい。) そして、系列の数字の向きは時計回りが反時計回りかによって区別される。私たちが以前につくった系列はリングにすることによって反時計回りと時計回りの違いとして表現されることになる。ここで新しい道具立てを考えよう。反時計回りでボールが 1 個ずつ各カードの上を

通過するとしよう。この輪になったカードの上を赤と白のボールが一枚のカード上に一個のボールという形で、追い越しなしに整然と通過するとしよう。さらに、魔法札と呼ばれる何枚かの特別のカードを通過すると、赤は白に、白は赤にとボールの色が変化するとする。以上がこのモデルのすべてである。ここで次のような命題を考えてみよう。

魔法札の数や赤、白のボールの数が最初どのようであれ、このシステムは赤、白同数のボールという均衡状態に達する。

この命題は、システムは最初の赤、白の数が異なるという不均衡な状態から赤、白同数という均衡状態に至る、つまりはシステムのエントロピーが増加するということを主張している。(残念ながら 13 枚のカードでは赤、白の枚数は 6 と 7 の間を振動する。) 例えば、赤が多い時には赤のほうがたくさん魔法札によって白に変わり、白はそれより赤に変わる回数が少ないのであるから、最初の状態がどのようであれ、最終的には赤、白同数に落ち着くだろうことは直観的に見て取れる。それゆえ、私たちの関心はこの命題の証明そのものではなく、この命題の証明に現れる時間の向きの論証にある。

N を系列のトランプカードの数とすると、 $N = 13$ である。魔法札は任意にその数を定めることができる。例えば、魔法札を偶数札とし、偶数の数をもつカードを通過すると赤、白の色の変化が起こるようにできる。(魔法札は 1 から 12 で何枚でもよいが、1 枚も魔法札がない、すべて魔法札という場合は排除しておこう。なぜこの条件は必要か?) ここで時間 t の白ボールの数を $W(t)$ 、赤ボールの数を $R(t)$ とすると、その正規化された差は $d(t) = (W(t) - R(t))/N$ である。時間 t で魔法札上にある白ボール、赤ボールをそれぞれ $W'(t)$ 、 $R'(t)$ とする。すると、次のことが成立している。

$$W(t+1) = W(t) - W'(t) + R'(t),$$

$$R(t+1) = R(t) - R'(t) + W'(t).$$

全体の中での魔法札の割合を $m = (\text{魔法札の集合の個数})/N$ と表すことにしよう。ここで次の仮定を置く。(この仮定は molecular chaos の仮定に対応するものである。)

$$(\text{魔法札上の白(赤)ボールの数}) / (\text{白(赤)ボールの数}) = (\text{魔法札の数}) / N^{(\#)}$$

(この仮定から、 $W'(t)/W(t) = m$ 、 $R'(t)/R(t) = m$ となる。)

($\#$) この仮定は次の関係が成立することから出てくる。(白ボールの数):(赤ボールの数) = (魔法札上の白ボールの数):(魔法札上の赤ボールの数) 心配の向きは実際に証明してみればよい。例えば、白ボールの数を W 、魔法札上の白ボールの数を M 、赤ボールの数を R 、魔法札上の赤ボールの数を M' とすると、上の関係から、 $W:R = M:M'$ 。 $M/W = M'/R = k$ と

すると、 $(M + M')/N = k$ となって仮定が満たされる。

この仮定を使って、上の関係を表すと、

$$W(t + 1) = W(t) - m(W(t) - R(t)),$$

$$R(t + 1) = R(t) - m(R(t) - W(t)).$$

また、赤、白のカードの差は、 $d(t + 1) = (W(t + 1) - R(t + 1))/N = (1 - 2m)d(t)$ となり、したがって、 $d(t) = (1 - 2m)^t d(0)$. それゆえ、このシステムは冪的にすばやくピンク色になる。(実際にピンク色になるわけではない。迅速にカードが移動した場合に私たちに見えるのがピンク色であり、それは力学的に考えていたのでは現れない emergent property である。) そして、 $m > 1/2$ なら振動を伴うことになる。これが私たちの命題の証明である。赤と白のボール数の差は減っていき、最終的に均衡状態に達する。

では、エントロピーはどのようになるか。パラメータ d は、 $d = (2W - N)/N$ ($W = 0, \dots, N$) であるから、 -1 から $+1$ の間を動くことになる。そこでシステムのエントロピーを次のように定義しよう。

$$S(d) = k \log[(\text{the number of microstates consistent with } d)]$$

システムのボールが取れるミクロな状態の総数は 2^N であり、 W の白ボールの数をもつことと整合的なボールの状態の数は N から W の対象をを選び出すことである。そこで N から W を選び出す組み合わせの数を上の式に入れるとエントロピーが計算できることになる。

以上がこのモデルの記述とそこから得られる結果である。エントロピーの増大が証明され、それが私たちの命題の実質的な内容となっている。どのようなリングモデルのシステムであれ、そのエントロピーは増大することが同じように証明でき、そこから第二法則の正しさが帰結することになる。これが標準的な理解である。

このような標準的な理解に対して、可逆性と再帰性のパラドックスを考えてみよう。

(1) 可逆性

例えば、システムの 75% が白ボールで実験が始まったとしてみる。これはそのまま進行すると、上述の結果となる。つまり、最終的に赤、白のボールの数は同数に近づく。この逆の操作を考えてみよう。ボールが時計回りに動き、魔法札に入る際にボールの色が変わる。このような逆向きの変化は上述の計算の仕方と同じように記述できる。すると、 d は減っていき、したがって、エントロピーは増大する。つまりはピンク色になっていく。時計回りの逆向き変化でもシステムのエントロピーは増大することがわかる。(#) ここで問題が生じる。一定の期間、反時計回りでシステムが動いた後で、今度は時計回りに変化し出すとし

てみよう。すべての変化は逆向きであるから、同じ期間逆の変化が持続した後では、システムは最初の白 75%に戻るはずである。リングモデルのシステムが可逆的な操作だけからなっていることは明らかであろう。ところがエントロピーは増大する。これは矛盾である。

(2) 再帰性

再帰性はリングモデルの場合、極めて簡単である。反時計回りに $2N$ 回ボールが移動するとどうなるだろうか。ボールはそれぞれの魔法札を 2 回通過し、そこで色の変化を被るが、2 回の操作の合成は色の変化をもとの色と同じにする操作と同じである。したがって、2 回転後のシステムの状態は最初の状態と同じになる。後は何回移動させてもこの周期で変化が繰り返されるだけである。システムは偶数回の回転後には 75%の白ボールの状態に復元している。したがって、このシステムは規則的な変化を周期 $2N$ で繰り返すという時間発展の形式を持っていることになる。しかし、上述の計算はこれとは大違いで、システムのエントロピーは増大する。これも矛盾である。

(#)(1)の可逆性の反論については異議のある者が少なからずいるのではないか。その異議はエントロピーが逆向きの変化に対しても増大するという点についてのものである。上述のように任意の赤、白のボールの数の異なる状態から逆向き変化を考えた場合には最初の状態より確かに増大する。より確定的には、反時計回りの初期条件と同じ条件のもとで時計回りを考えた場合に増大する。しかし、白 75%のある状態から反時計回りでのシステムの経過を追い、ちょうどある時点でそれと全く正反対の経過を辿らせた場合、同じ経過の逆経過であるから最初の状態に戻ることになる。そして、リングモデルでの逆向き変化は 1:1 でもとの変化に対応している。この戻りつつある経過の間、エントロピーはどのように変化するのか。白 75%に戻っていくのであるから、エントロピーは減少していくことになる。そこから更に逆向きの変化を続けるならば、エントロピーは増大に向かい均衡状態に達する。したがって、上の内容を正確に言うならば、任意のエントロピーの低い状態から出発しての逆向き変化はエントロピーが増大するというべきなのである。これは私たちが考えているのが単純なモデルであるからこそ明らかになる点である。

さて、この歴史的な反論に対しての典型的な解答はアンサンブルを使ったものである。

A を魔法札の集合とする。その特徴関数を次のように定義しよう。

$$e_p = \begin{cases} -1 & \text{if } p \in A \\ 1 & \text{otherwise} \end{cases}$$

ここで、カードの数をそのまま文字盤の住所にする代わりに反対回りに順序をつけ直し、その番号を p としておこう。(通常の時計の 12 時のところに 1 のカードがくると思えばよい。)ここでボールの色についての特徴関数(正確にはボールの状態の力学的変数)を次の

ように定義する。

$$n_p(t) = +1 \text{ if the ball in the site } p \text{ at time } t \text{ is white}$$

$$= -1 \text{ if the ball in the site } p \text{ at time } t \text{ is red}$$

さて、システムの時間的な発展を追ってみよう。すると、上の定義を使って次のような発展として記述できる。

$$n_p(t) = e_{p-1} n_{p-1}(t-1). \text{ この式から、 } n_p(t) = e_{p-1} \dots e_{p-t} n_{p-t}(0).$$

二つの特徴関数によって赤、白のボールの個数をシステムの単位変化（各ボールがその占めている座を一つ整然と反時計回りに移動する変化）にしたがって計算していくことができる。ボール個数の収支決算の工夫と考えればよい。すると、システム全体の収支は、

$$d(t) = 1/N \sum_p n_p(t)$$

$$= 1/N \sum_p [\sum_l e_{p+l}] n_p(0) (0 \leq l \leq t-1)$$

となる。ここで求めたいのは $\sum_l e_{p+l} n_p(0)$ の値である。それは $[\sum_l e_{p+l}]$ の期待値を求めることである。もし魔法札の各住所 p がランダムであれば e_p は他の住所変数 e_p とも n_p とも相関していない。相関のない変数についての積の平均は平均の積である。例えば、 $abc = a b c$ である。ここで各 e の平均を考えると、

$$e = (+1)Pr(e = +1) + (-1)Pr(e = -1)$$

$$= (+1) \times (N - |A|) / N + (-1) \times |A| / N \text{ (ここで } |A| \text{ は魔法札の集合 } A \text{ の個数とする)}$$

$$= 1 - 2|A| / N$$

$$= 1 - 2m$$

となる。この値は以前の計算結果と同じである。したがって、

$$[\sum_l e_{p+l}] = (1 - 2m)^t .$$

それゆえ、 $d(t) = 1/N \sum_p (1 - 2m)^t n_p(0) = (1 - 2m)^t d(0)$ が得られる。

ここでの導出の特徴はアンサンブルにある。それは各 e の独立性に現れている。可逆性の反論におけるの逆の過程での力学法則は

$$n_p(t) = e_p n_{p+1}(t-1). \text{ (普通の過程では } n_p(t) = e_{p-1} n_{p-1}(t-1))$$

である。さて、普通の過程で t まで発展したとして、次の $t+1$ で逆の過程になったとしよう。すると、上の式の t を $t+1$ で置き換えると、 $n_p(t+1) = e_p n_{p+1}(t)$ であり、したがって、括弧の中の関係から、

$$n_p(t+1) = e_p e_p n_p(t-1). \text{ここで } (e_p)^2 = 1 \text{ であるから、 } n_p(t+1) = n_p(t-1).$$

この式は t まで普通の過程で発展し、 $t+1$ で逆に発展した場合に $t-1$ のシステムの状態に戻ることを表している。ここで $(e_p)^2 = 1$ は相関関係のあることを表しており、独立していない。しかし、上の導出では次の二つの性質が重要であった。

$$e = 1 - 2m, \quad abc = a \quad b \quad c .$$

この二つの性質は相関がないことを表している。ここで重要なのは次のことである。アンサンブルのなかでは因果関係はどこにもない。それはアンサンブルの各要素が因果的には独立したものであり、それら各要素間の関係を因果性で解釈することはできないことを意味している。アンサンブルに依存した概念は因果的に無力、惰性的である。これがアンサンブルに確率を適用した場合の基本的な特徴である。

再帰性についてはどうか。 $n_p(t)$ の積では異なる e_p を含んでいる。これは $t < N$ について言うことであって、 $t = N$ では同じものが 1 回以上現れる。(Addition and subtraction of the index p is modulo N .) このような繰り返しも含んで計算がされており、それでも一般的な性質としてシステムのエントロピーの増大は証明されるようになっている。アンサンブルではシステムの各状態の輪廻的な変化は完全に消えてしまっている。

以上がアンサンブルを使った証明の概略であり、かつ可逆性と再帰性に対する解答である。アンサンブルを使った証明は確かに可逆性と再帰性の反論を見事にかわしている。可逆性も再帰性もシステムの時間的な発展のもつ性質である。アンサンブル解釈はこの時間発展を集団の平均として表現し、特定の時間発展に関与しない。それは時間発展するシステムの統計集団に関する主張であって、システムの時間発展の性質に関する主張ではない。両者は類似した性質として表現できるが、その存在論的な根拠、つまりは何についての性質かという点で決定的に異なっている。

アンサンブルを考えることによって、システムの個別的な時間発展という観点からは見えない性質がわかるとともに、時間発展という観点でこそ見える性質が見えなくなる。

これが逆にアンサンブル解釈の欠点ともなってくる。それはシステムの因果的な変化とは無関係である。この無関係性がシステムの時間的な発展に関する主張をしようという場合には不利に働くことになる。ところで、時間の向きはシステムの時間発展に密接に関係している。標準的なシステムがあるとするれば、その時間発展の向きが時間の向きであるとい

ってもよいくらいに密接な結びつきをもっている。したがって、アンサンブル解釈は時間の向きに対して、それにコミットしないという理由から有益であるとは言えない。

07 確率とその解釈

確率や統計は 20 世紀の知識の性格を他の世紀から区別するような大きな役割をもっている。最初は天文観測の誤差をどのように処理するかから始まった確率の研究は社会科学や自然科学の中で積極的に活用され、今やそれ無しには日常生活さえまなぬものになっている。確実でない、蓋然的知識が市民権を得て、確実な知識と肩を並べている。「明日の降水確率」はそのような例の一つである。しかし、「明日は 20%の確率で雨が降る」と言われた時、その正確な意味を即答できる人は少ない。誰も一日の 20%の時間、つまり 4.8 時間雨が降るとは思わないだろう。では「明日の降水確率」の確率とは何なのか。

命題や集合を考えよう。論理の規則や集合の演算に関して閉じた命題や集合の集まりに測度 $P()$ を定義できる。この関数は命題や集合を実数に写像する。そして、任意の命題や集合 A 、 B に対して、次の条件を満たすとき $P()$ は確率測度と呼ばれる。

$$0 \leq P(A) \leq 1$$

もし A が真の命題か全体集合なら、 $P(A) = 1$

もし A と B が両立不可能 (排反的) なら、 $P(A \text{ or } B) = P(A) + P(B)$

これらはコルモゴロフの確率の公理系である。確率は通常特定の背景知識をもとにして、確率モデルを組んでその中で確率測度を与える仕方で行われる。例えば、サイコロ振りを考えてみよう。サイコロに関する背景知識によって私たちはそれが 1 から 6 までの数字の目をもつ立方体であることを知っている。さらに、サイコロが公平であるとすると、通常考えられるモデルは「1の目が出る」、「2の目が出る」、...「6の目が出る」の各命題に対して $1/6$ の確率測度を与えるようなものである。さらに、この測度は $P(\text{「1の目が出るか3の目が出るかである」}) = P(\text{「1の目が出る」}) + P(\text{「3の目が出る」})$ を満たしている。

この数学的な確率概念の特徴は確率を尺度、物差しとして考えるところにある。この他にも確率の定義はあるが、上のような形式的な定義と並んで、そのような形式的な定義をもつ確率とはそもそも何かという解釈の問題が議論されてきた。それらのうちから主要なものを考えてみよう。

命題や集合は出来事を意味しており、その出来事の集団における実際の頻度が確率の解釈の一つである。コインを 100 回投げ、そのうち実際に 39 回表が出たとしてみよう。これを H という命題で表せば、 $P(H)$ は 100 回のコイン投げの実際の頻度と解釈できる。この解釈は上の公理をすべて満たしている。実際の頻度を使った解釈は客観的な解釈である。それはある出来事が集団内でどの程度の頻度で実際に生じたかによって確率を解釈している。

主観的に確率を解釈するのは別の解釈である。私たちは与えられた命題が真であることに

どのくらい信頼性を置くべきかについて語るができる。この概念は心理的であるだけでなく、規範的でもある。それは私たちの信念の度合が何であるべきかを述べているからである。そして、この信念の度合はやはり上の公理を満たしている。

三番目の解釈は出来事の確率はその仮説的な相対頻度であるというものである。公平なコインはそれが有限回投げられたとき、正確に同じ回数で裏と表が出なくてもよい。しかし、何度も投げていけば最終的には 0.5 に収束する。 x の確率値は x に等しい実際の頻度を帰結する必要はないが、無限に続く仮説的なコイン投げでの頻度は x の値に収束することを帰結する。実際の頻度も信念の度合の解釈も、いずれも確率を何か別のものを使って解釈するものであるが、この三番目の解釈はそうではない。この解釈は循環している。

その理由を知るために、無限の回数の公平なコイン投げが 0.5 の相対頻度に収束しないとしてみよう。そこで何か特定の系列を思い浮かべてみよう。HTHTHT.....という交互に表裏が出る系列の場合、コイン投げの回数が増えれば、そのような系列の出る確率は 0 に近づいていく。無限の回数の試行では、どのような特定の系列もそれが達成される確率は 0 になる。しかし、どれか特定の系列は実際に起こる。したがって、確率 0 と不可能を同じとみなすことはできない。同じように、確率 1 を必然とみなすこともできない。それゆえ、公平なコインが 50%の相対頻度で表が出ることに必然的に収束するわけではない。もし表の相対頻度がそのコインの表の出る真の確率に収束する必要がないのであれば、どのように二つの概念は関係しているのか。大数の法則がその解答を与えてくれる。

$$P(\text{表が出る} | \text{コインが投げられる}) = 0.5$$

$P(\text{表の頻度} = 0.5 \pm e | \text{コインが } n \text{ 回投げられる})$ が n が無限に近づくと 1 に近づく

(ここで e は任意の小さな数である。)

(($P(A | B)$ は条件付き確率と呼ばれ、 B が生起するときの A の確率である。)

試行の回数が増えると、 $0.5 \pm e$ 内になる確率は高くなっていく。ここで の両側に現われる確率概念に注目してほしい。仮説的な相対頻度解釈は解釈ではない。というのも、 の両側には共に確率概念が使われているからである。

最後の確率解釈は、同じような欠点をもつとはいえ、人気のある傾向性解釈(propensity interpretation)である。傾向性は確率的な性向(probabilistic disposition)である。では、この確率的な性向とはどのようなものか。性向は「...できる」という言い方をもつ言葉で表現されている。例えば、可溶性は性向の一つである。それは次のように定義できる。

X が可溶である X が通常の条件で浸されるなら、X は溶解する

この定義は、ある「もし...ならば、...である」という文が真であれば、その時に可溶である

ことを述べている。これは X が一度も浸されなくとも構わないことを示している。さらに、通常の条件も重要である。また、この定義は決定論的な表現になっている。可溶性物質は浸されるなら溶けなければならないのである。

確率の傾向性解釈は「もし...ならば、...である」という文に類似の説明をする。コインが投げられると、その表の出る確率は 0.5 であるとしてみよう。もしこれが正しいなら、何がこの正しさを生んでいるのか。コインが特別な性向である傾向性をもっているからであるというのがこの解釈の答えである。もしコインが表の出る確率 0.5 であるなら、それは投げられたとき表の出る強さ 50% の傾向性をもっている。それはちょうど砂糖の塊が水に入れられると溶けるというのと同じである。

傾向性解釈は決定論的性向と確率的傾向性との類比を強調する。ある対象が可溶であるかどうかを見つけるには二つの方法がある。もっとも明らかな方法はそれを水に浸し、それが溶けるかどうかを見ることである。二番目の方法は、その対象が可溶性物理的構成になっているかどうか調べることである。つまり、性向はそれに伴う振舞いと物理的な基盤をもっている。そのいずれかを使うことによって対象が当の性向をもっているかどうか見出すことができる。これは確率的な性向についても正しい。コインが公平かどうかを二つのいずれかの方法によって見出すことができる。実際に何回か投げてみる、あるいはコインの物理構造を調べることのいずれかによって公平かどうかわかる。確率的な性向もその振舞いあるいは物理的構造から見出すことができる。ここには明白な類比が見られる。

それでもなお、この傾向性解釈には疑いの余地がある。まず、説明が十分一般的でない点である。傾向性解釈での原因と結果の関係は「もし...ならば、...である」で表されている。しかし、「もし...ならば、...である」という関係はいつも因果関係を表すわけではない。両親の遺伝子型は子孫の遺伝子型の原因であるが、それと逆のことも「もし...ならば、...である」という形式で問題にできる。条件付き確率はいつでも因果関係を表すのではない。

より基本的な問題は「傾向性」という言葉が「確率」という言葉の別の名前にすぎないのではないかという点である。「傾向性」と「確率」のいずれが明白な意味をもっているだろうか。もし「傾向性」が確率概念を使ってしかわからないのであれば、この解釈は一層事態を複雑にするだけである。

私たちは二つの整合的な解釈、実際の相対頻度解釈と信念の度合いの主観的解釈を述べてきた。もし確率概念が科学において自然に関する客観的な事実を述べているのであれば、私たちは困難に直面する。もし実際の頻度解釈が否定されるなら、公平なコインの表の出る客観的な確率が 0.5 というのはどのような意味をもっているのだろうか。

一つの解決の仕方は確率が客観的であることを否定することである。確率の表現は私たちが原因を知らないために行うものであると考えることである。不完全な情報のもとで私たちがもつ信念の度合を表すと考えることである。この場合、確率は生じることを予測するのに十分な情報をもっていないために存在することになる。

量子力学によれば、偶然は自然のシステムの客観的な性質である。たとえ私たちが必要な

情報をすべて知っていても、量子力学的なシステムの未来の振舞いを正確に予測することはできない。物理学では、したがって、主観的な解釈では十分でないことになる。しかし、量子力学で偶然が客観的なものであるとしても、他の理論でもそうであると結論することはできない。

「04 時間の向き」でトランプモデルで系列 3、4、5、6、7、8、9、10、11、12、13、1、2 をつくったことを思い出してみよう。このような系列世界に確率を導入しようとするれば、 2^{13} のトランプ数字の部分集合を考え、そこに $P(\cdot)$ を定義すればよい。この時注意してほしいのは系列の順序はバラバラになり、出来事も膨れ上がり、 $\{2, 5, 10\}$ といった出来事が加わることである。これは確率モデルがシステムの時間変化を追うのではなく、システムの変化の総体に関して考えられていることを示している。さらに、リングモデルでの赤白の変化のように、世界での変化を確率空間でモデル化する場合、その変化はマクロな変化であり、そのような確率空間は力学的な空間に付随している。

08 決定論

すべての出来事はそれ以前の原因の結果であるというのが決定論である。このような形而上学的な主張はニュートンの力学によって、物理的な世界の決定論として精巧に具体化された。ニュートンの決定論は自然についてのビリヤードボール説とも言われており、その洗練された表現はラプラスの魔物（物理学者ラプラスが思考実験で考えた架空の万能者）によって見事に示された。

私たちはコイン投げやサイコロ振りを確率的な出来事の典型例だと考えている。実際、教科書にもある通り、公平なコインは表、裏の出る確率が $1/2$ とみなされ、確率モデルがつくられる。このような確率的な出来事は私たちの生活に馴染んでおり、公平な選択のためにコインやサイコロが使われ、時には賭けの道具にもなっている。しかし、もしニュートン的な決定論が正しいとしたら、ある公平なコイン投げの表が裏のいずれかが出ることは決まっていないのだろうか。このような疑問に答えるために考えられたのがラプラスの魔物である。

ラプラスの魔物はコイン投げについての完全な知識をもっており、投げられるコインの物理的な運命について完全に予測できる。魔物はなぜコイン投げの過程が確率的と理解されるのかについて次のように説明できた。人間はコイン投げについて十分な物理的知識がなく、正確な予測ができないために、その過程が確率的に見えるに過ぎない。魔物はコイン投げでもそれが生じるときのバイアス（非対称性）は決して見逃さない。コインを投げるときの物理的な状態のバイアスが何であるかを的確に掴み、それが結果にどのようなバイアスを生むかを正確に予測できる。コインを投げて裏が表が出たということは、その結果にバイアスがあったということであり、それは原因であるコイン投げのどこかに最初からバイアスが潜んでいたためである。これは理屈の通った話に思える。というのも、これ

は実は物理学の基本原理であって、対称性の原理(Principle of Symmetry)と呼ばれてきたものの一例なのである。その主張は、

結果に現れる非対称性は、原因にそれを引き起こす非対称性がある、

ということである。この原理が成立している限り、魔物は原因のバイアスに注目することによって結果の裏、表というバイアスの予測を物理学的に行なうことができる。

以上のことから、魔物は物理的な状況に関して予測ができ、確率などに頼らなくても、個々のコイン投げを一回毎に正確に予測でき、したがって、すべてのコイン投げの系列について正確な予測を行うことができる。つまり、魔物にとってはコイン投げの過程は全く決定論的である。それゆえ、確率の使用を主張・擁護する者の理解は誤っており、自然の過程に確率的なものは何ら含まれていないことになる。

この説明によれば、確率は私たち人間には不可避免的に必要であるが、それは私たちが十分な知識をもたないために過ぎない。これは確率の主観的な解釈である。私たちが確率概念を使う理由は私たちの無知のためであり、もし十分な知識をもっていれば確率などに頼る必要はないのである。

さらに現存する確率的な科学法則についても、それは現象的な法則であり、時間対称的な物理学の法則とは違って派生的なものに過ぎないと魔物は結論した。対象の時間発展を述べる法則に対して、そのような統計法則は単なる収支決算の報告の仕方に過ぎず、厳密な意味で法則ではない。そもそも確率が古典的無知の反映であるから、それを使つての確率的な法則は法則と呼ぶに値しない。幽霊はどこにも存在しないが、考え出された多くの幽霊について的一般法則はつくろうとすればできる。統計法則はそのような類の規則であるというのが魔物の結論である。ちなみに、収支決算に過ぎないと言われる法則にはエントロピー増大の法則やメンデルの遺伝法則がある。

ラプラスの魔物は、任意の正確さで初期条件を測ることができ、未来の予測のためには瞬時に完璧な計算ができなければならない。これが決定手続きを考えたときの魔物に課せられる条件である。元来、決定論は実在の決定性を主張するものであり、私たちの認識とは何の関係もないものである。その決定論と予測可能性を同一視させる理由は古典力学の第2法則にある。第2法則と、微分方程式系の解が存在して、しかもその一意性を保証する定理とが結びつくことによって、系の初期条件が定まれば正確な予測が可能であることが数学的に証明できる。これによって現在の状態から演繹される未来や過去の状態が存在するということが保証される。さらに、この決定論は上の予測が実際に構成的に計算可能であるという定理によって強化される。ただ単に予測が可能というのではなく、実際に予測を計算できる。こうして古典的な決定論は予測可能性と同一視されることになる。そして、このような決定論 = 予測可能性という認識的な決定論理解が、ラプラスが魔物に対して与えた役割である。

このような魔物の主張は私たちの行為にも当てはまるのだろうか。自分や他人の行為の予測は大抵できないが、それは私たちの無知のためだけなのか。ここで、決定論と運命論(fatalism)の区別が重要である。物理世界が存在し、ある時点の状態がわかっているならば、ラプラスの魔物にとって古典力学が主張する決定論は運命論である。(したがって、この節のタイトルは運命論としてもよかったのかもしれない。)決定論は、もし過去が異なっていたとすれば、現在も異なっていたらという考えを排除しない。決定論はまた、現在私がある仕方ではなく別の仕方を選ぶならば、私は未来に起こることに影響を与えることができるという考えも排除しない。しかし、運命論はこれを否定する。現在あなたが何をしようと過去と未来はそれとは無関係であるというのが運命論の主張である。つまり、決定論と運命論はほとんど正反対のことを主張している。運命論は私たちの信念や欲求が無力であることを主張するが、決定論では信念や欲求は因果的に私たちの行動をコントロールできることが主張されている。

09 非決定論

私たちは決定論、確率といった概念がどのようなものか既に考えてきた。そこには非決定論が見え隠れしていた。言明の証明可能性や真偽、あるいは関数の計算可能性が決定できないということは現在ではよく知られた事実である。このような非決定論的な事実は私たちの無知によるという主観的解釈によって、世界の出来事の生起に関する非決定論とは区別できる。では、私たちとは独立した世界の事象が非決定論的なものであるとはどのようなことなのか。これがここでの問題である。非決定論的な主張をする理論は量子力学である。そこで量子力学がどのような意味で非決定的な理論であるかを考えてみよう。

トリチウム(三重水素)は水素の同位元素であり、その半減期は12.32年であることが知られている。 n 個のトリチウム原子はその半数が12.32年以内に崩壊し、残りの半数はそのままである。 n 個の原子のいずれが崩壊し、いずれが崩壊しないかは原理的にわからない。それでも、それぞれの原子がそれぞれ特定の時刻に崩壊しているはずであり、あわよくばそのような法則を見つけることができると考えたい。これが不可能な望みであることは量子力学が本質的に非決定論的な理論であることを肯定することから得られる。どの原子がいつ崩壊するかわからないので、非決定的な崩壊のモデルは確率を使ったものになる。多数のトリチウム原子の集団を考えると、一つの原子についてその崩壊時刻を知ることはできないが、時刻 t にまだ崩壊していない原子の個数 $N(t)$ を近似的に予測することはできる。これを利用して崩壊の確率モデルをつくることができる。非決定性は統計・確率概念によって表現される。非決定性と確率は深い結びつきをもっているとはいえ、それは論理的な結びつきではない。この結びつきには経験的な確証が必要である。

量子レベルでの非決定性というときハイゼンベルグの不確定性原理(uncertainty principle)を思い浮かべる人が多い。それは「電子のような粒子の位置と運動量を同時に正確に知ることができない」とよく述べられる。この表現のなかの「知ることができない」理由はそのよう

な値が存在しないからである。そのような値が存在しないので、電子の未来の振舞いは予測できないだけでなく、非決定的である。この非決定性(indeterminacy)が量子力学の領域全体を支配している。波動 - 粒子の二重性をもっとも具体的に示してくれるのは光である。光は物理的には極めて興味深い対象である。私たちの周りにあふれて存在しながら、その正体は長い間曖昧なままであった。この曖昧さは光の本性に関するニュートン以来の論争が見事に描き出してくれる。ニュートンは光を粒子の束であると考えたが、ホイヘンスは光が波であると考えた。不思議なことに光はマクロなレベルでも粒子の束であるという性質と波であるという性質を両方もっている。波の代表的な性質は干渉である。この干渉現象はヤングの二重スリットの実験で見事に示された。更にマックスウェルの電磁気学の理論は光が波であることの理論的な支柱となった。こうして今世紀初頭までの状況は光が波であるという考えがはるかに優位に立っていた。しかし、古典的な電磁気学に問題がなかったわけではない。黒体放射がプランクの量子論をつくらせることになったのは1900年12月だった。アインシュタインによる光量子仮説、コンプトン効果がこれに続く。三者に共通するのは、電磁放射における放出、伝播、吸収は量子として、つまりはエネルギーの局所的な束として行われることを示したことだった。だが、これらのいずれも光が波であることを傷つける結果ではない。光はある場合には波の性質を、別の場合には粒子の性質を示す。これが光の波動 - 粒子二重性である。これを更に物質的な粒子にも拡大したのがドブロイである。 $p = h/\lambda$ はドブロイが見出した関係である。運動量は波長の値に反比例している。この関係を念頭において光の二重性、粒子の二重性の関係を考えてみよう。波は空間に局在しておらず、したがって、位置をもっていない。これは粒子を波と見た場合も同じで、正確な運動量をもつ粒子は決定的な位置をもっていない。波長の異なる複数の波を重ねあわせると干渉が起こる。これをうまく利用して、異なる波長の波の重ねあわせからコンパクトな波束をつくりだすことができる。これを十分に局所化することができる。ただ、この波束は多くの異なる波長の波からなっており、波束の運動量は一定の幅をもつことになってしまう。つまり、波束の運動量は非決定的となる。これは次の関係を示している。

運動量の決定性 - 位置の非決定性

位置の決定性 - 運動量の非決定性

(位置と運動量だけではなく、エネルギーと時間の間にもこのような関係が成立する。位置と運動量はその決定性に関して反比例の関係にあるのである。)

ドブロイの波動という考えはシュレーディンガーの波動力学に引き継がれる。彼の理論によれば、電子のような量子力学的なシステムは波動関数 (q, t) によって完全に記述される。ここで q はそのシステムの位置を、 t は時間を表している。古典力学の運動方程式と同じように、そのシステムの時間的な変化は波動方程式によって特徴づけられる。以前のシステ

ムの状態がわかれば、その後のシステムの状態もわかるという決定論的な関係はこの波動方程式は表している。しかし、古典力学とは異なって、この後の状態は位置と運動量の両方の正確な値をもっていない。以前のシステムの位置と運動量も正確な値をもっていないのであるから、このことは決定論的な方程式の性格に違反しているわけではない。波動方程式は、もし時点 t のシステムの状態（位置と運動量）が正確に与えられれば、それを使ってそれ以後のシステムの状態が正確に計算できる、予測できることを主張しているだけである。決定論の主張では、ある時点 t でのシステムの完全な記述から、それ以後の時点 t' のそのシステムの完全な記述が演繹できる。

シュレーディンガーの波動方程式はこの決定論の主張を正確に満たしている。1925年にハイゼンベルクのマトリックス力学が出ると、シュレーディンガーはその力学が自分の波動力学に論理的に等価であることを示したが、ハイゼンベルクの力学が彼の不確定性原理を含んでいたことから、不確定性原理はシュレーディンガーの力学においても主要な地位を占めるはずである。

ここで注意しておかなければならないのは、ド・ブロイの波動が通常物質的世界の対象であるのに対し、シュレーディンガーの波動は抽象的な数学的空間の対象である点である。では、どのようにこの抽象的な波動を物理的な実在に適用したらよいのか。1926年にボルンが答えを見出した。波動関数の2乗が電子が与えられた領域に観測の結果として見出される確率であるというのがその答えである。この答えから得られる暫定的なまとめを以下にしておこう。

- (1)物質のミクロな描像が粒子と波動の両面を備えている。粒子と波動というモデルは、物理的には互いに排除の関係にあるにも拘わらずである。
- (2)統計的なアンサンブルについてはその確定記述を完璧に行うが、個々の事象の記述を考えた場合、波動関数の振る舞いを決める何か欠落していると感じられる。これが「不確定性関係」である。そこをコペンハーゲン解釈は「波動関数の収縮」で解釈するが、その解釈が測定、観測、認識といった対象自体でない要素で補完されている点は理論として不完全に見える。
- (3)シュレーディンガーの猫のパラドックスが示すように、測定と過程、対象系と測定器の境界、分離の曖昧さがコペンハーゲン解釈にはつきまとう。
- (4)EPR実験やベルの不等式で議論されているように、存在しているものを局所的に自立した実体と考えることができない。そう考えると矛盾がある。
- (5)量子的対象のイメージが明確でない。「イメージ」として人々が要求するのは物質、本質、実体、存在、実在、実存、秩序、情報、などなどいろいろなものがある。

10 非線型性あるいは決定論と非決定論の重複

確率や決定論が何を意味しているか考えてみよう。線型の方程式で表わされる線型の過程

も、非線型の方程式で表される非線型の過程も決定論的な過程である。状態変化の仕方は決まっている。一方、確率過程は非決定論的な過程の代表である。それゆえ、二つの過程は根本的に異なったもののように見える。しかし、本当にそうであろうか。近年の非線型現象への関心は生命現象の研究をきっかけにしていた。生物集団の人口動態は非線型的な過程であり、交配現象は非線型現象としてみる事ができる。マルコフ過程としての交配は確率的な過程であるが、その同じ過程は非線型の方程式系としても表現できる。一方、非線型の方程式系として表現された交配過程は確率過程として表現することもできる。この数学的事実は形式上、ある条件の下で確率過程と非線型の方程式系が数学的に等価であることを示している。

この結果を使って、決定論と非決定論の関係を非線型の方程式と確率過程の関係に限定して浮かび上がらせてみよう。決定論には線型と非線型の 2 種類があり、次のように分類できる。

A. 線型の法則 ($v = f(u, t)$)

- (1) 初期値を決めると t での状態が唯一つ決定される。
- (2) t による変化は比例的である。

B. 非線型の法則

線型の法則の性質 (1) だけをもつもの

一方、非決定論の代表は確率過程であり、サイコロ振り、コイン投げ等がその具体例である。これらは線型の法則の (1) (2) のいずれも満たさない。いずれの場合であれ、システムの状態は次のような系列で表現できる。

$$x_0, x_1, x_2, x_3, x_4, \dots$$

すると、決定論的な系列の場合、

$$x_{n+1} = f(x_n) \quad (\text{このように表現できるシステムを力学系と呼んでいる。})$$

と表現でき、上の A は 1 次関数、B はそれ以外の関数が対応していることになる。一方、コイン投げは n 番目の結果から $n + 1$ 番目の結果を知ることができない。

このような準備の下で、次のようなことが成立する。適当なコイン投げの系列が与えられたとする。すると、その系列に対して、その同じ系列を生み出す決定論的な力学系が存在する。つまり、非決定論的な系列に対して、決定論的な力学系でその系列を生み出すことができる。この驚くべき結果を考えてみよう。

次のような決定論的な式 $x_{n+1} = f(x_n)$ を考えよう。

$$x_{n+1} = 2x_n \quad \text{if } 0 \leq x_n < 1/2$$

$$=2(1-x) \quad \text{if } 1/2 < x_n < 1$$

このような式のもつ特徴は、区間 $[0,1/2]$ を A 、区間 $(1/2,1]$ を B としたとき、次のようなことが成立している点にある。

$$[0,1] = A \cup B \quad f(A), \quad [0,1] = A \cup B \quad f(B)$$

ここで無限回のコイン投げを考え、表を A 、裏を B とすると、 A と B からなる次の系列ができあがる。

$$(*) \quad w_0, w_1, w_2, \dots, w_n, \dots$$

この系列は、例えば、 $AABABBBBA\dots$ といった系列であり、中には規則的なもの、つまりは周期的なものも含まれている。ここで先の決定論的な式 $x_{n+1} = f(x_n)$ の一つの軌道を次のように表現してみよう。

$$(**) \quad x_0, x_1, x_2, \dots, x_n, \dots$$

ここで注意しておいてよいのは、 $x_i \in A$ or $x_i \in B$ が成立していることである。すぐにわかるのは、(**)に対して(*)の一つが対応していることになる。これは決定論的な式の値がコイン投げの系列の一つと対応していることを示している。既述のごとく、これは不思議なことではなく、コイン投げの結果が規則的になる場合がいくらでも考えられるからである。

不思議なのはこの逆が成立することである。つまり、(*)のような任意の系列、 $w_0, w_1, w_2, \dots, w_n, \dots$ に対して、(**)の系列で、 $x_n = w_n$ が成立する。 f は決定論的であるから x_0 が定まれば系列の他の値は決定論的に定まる。したがって、(*)を任意に与えると、 $x_n = w_n$ となるような系列の初期値 x_0 をうまく見つけることができることを意味している。ここでは証明は与えないが、関心のある向きは力学系のテキストを参照してほしい。

さらに次の事実も決定論と非決定論の違いを曖昧にするものである。 n 番目のコイン投げが表なら、 $n+1$ 番目が裏である確率が p であり、 n 番目のコイン投げが裏なら $n+1$ 番目が表である確率は q としてみよう。すると、 A_n, B_n をそれぞれ n 番目のコイン投げの結果が裏、表であることを表すとすると、

$$P(B_{n+1}|A_n) = p$$

$$P(A_{n+1}|B_n) = q$$

と表すことができる。

ここで、 $P(A_0) = 1$ と仮定しよう。また、 $x_n = P(A_n)$, $y_n = P(B_n)$ としよう。すると、

$$\begin{aligned} x_{n+1} &= P(A_{n+1}) \\ &= P(A_{n+1}|A_n)P(A_n) + P(A_{n+1}|B_n)P(B_n) \\ &= (1-p)x_n + q(1-x_n) = (1-p)x_n + qy_n. \quad (x_n + y_n = 1 \text{ より}) \end{aligned}$$

同様にして、 $y_{n+1} = px_n + (1-q)y_n$. ここで p や q が一定の確率値であるとする、その値は不変であるからそれぞれ定数 k 、 l で置き換えると、

$$x_{n+1} = (1-k)x_n + ly_n, \quad y_{n+1} = kx_n + (1-l)y_n$$

という式が得られる。この式の形を見ただけで、非線型の決定論的な式であることがわかる。こうして私たちは非決定論的なマルコフ連鎖の例から非線型の決定論的な式を得ることができた。

11 因果連関

既に形而上学の主題の一つとして因果性があると述べた。原因と結果の関係はどのようなもので、世界は因果的かどうかといった問題が哲学では考えられてきた。ここでは物理的な因果関係よりは私たちに馴染みのある日常的な世界での出来事の間での因果関係を考えてみよう。私たちの行動も当然ながら因果的な変化であり、物理的な性質や過程を利用しなければなにもできない。しかし、石が転がり落ちるのは違って私たちは自分の意志で歩いたり、走ったりしている。信念や欲求の介在が石の運動と私たちの行動を分けている。心が介在して身体的な運動が始まったり終わったりする。すると、信念や欲求が原因となって行動が結果となるような因果系列が考えられる。心が身体に作用しているというわけである。ところで、このような心的因果についての議論は私たちをどこに導くのだろうか。多くの哲学者が参加して心的因果の姿を捉えようとしている。なかなか一致が得られない理由の一つは心的因果が心の内だけでなく、心の外とも結びついているためである。外との結びつきは信念や欲求が原因になって行為が実行されることにある。

私たちの行為は世界の中で実行される。その世界は因果的な連関をもつ出来事や状態の集まりである。行為もまた出来事の一つであるから、世界の因果連関の中に組み込まれていることになる。このような因果の鎖の一部を取り出してみるなら、次のような一連の関係が見えてくるだろう。

環境(物理的な世界)、遺伝子 心(信念+欲求) 行為

心の関与は私たちの常識的な世界では重要な役割を果たすものとみなされている。しかし、

心の働きと物理世界との関係は疑うことができないように見えながら、決して明らかなものではない。この明らかでない関係が引き起こす典型的な問題が自由と決定に関するパズルである。それは次のように表現できる。人間の信念、欲求、そして行為がその人自身のコントロール外のものによって引き起こされるなら、そこに私たちの自由な裁量は入ってこない。環境や遺伝子は人間の心のあり方を決定している。そして、その環境や遺伝子は私たちのコントロール外のものである。そうであるなら、どのようにして行為が自由選択の結果と言えるのか。これが伝統的なパズルである。

自由な行為に見えないような行為は私たちの周りに溢れている。そして、そのような行為は社会にさまざまな問題を生じさせている。例えば、病的な盗癖をもつ人の窃盗は自由になされた行為ではないことから、法的な責任を免れるのだろうか。このような問題の背後にあって、私たちを悩ませている概念が因果性（因果作用）である。

因果的な連関に関して古来議論されてきた考えは既に述べた通り、次の二つである。

決定論：因果的な事実を完全に記述すれば、何が将来生じるか決定できる。

非決定論：現在の完全な記述が与えられても、将来に二つ以上の可能性を残す。

因果的決定論はあらゆる因果的に関連する事実が与えられれば、将来はただ一つだけ決まると主張する。もしすべての物質が決定論的で、心も物質であれば、人間の行為は物理的に決定されていることになる。実際、古典力学は物質の決定論を主張してきた。この世界観は今世紀まで信じられてきたが、量子力学の登場と共に非決定論的な世界観が浸透し始めている。物理学以外の領域では人間の行為の自由選択、意志の自由が古くから認められてきたため、19世紀には社会科学で既にその自由の入った出来事や状態を確率・統計概念を用いて取り扱ってきた。この二つの流れと心の特徴づけは上述のパズルとなって私たちに突きつけられている。心が物理的なものかどうかを未定のままにしておいても、次のような二つの選択肢の間で決断を迫られることになる。決定論が正しいなら、確率は私たちの知識や情報の欠如であり、主観的なものになる。決定論が誤っているなら、確率は世界についての客観的な事実を述べていることになる。では、非決定論は私たちを自由にするのか。それは私たちに自由を保証する理論的な根拠になるのだろうか。確率が世界の客観的な事実であるなら、上述の行為の因果連関は次のように修正しなければならなくなる。

環境、遺伝子、偶然 (chance) 心 (信念 + 欲求) 行動

そこで、あなたの信念と欲求が環境、遺伝子、そして偶然によっているとしてみよう。もし決定論があなたの自由を奪うなら、偶然もやはりあなたの自由を奪うことになる。あなたが自ら自由に決定するのではなく、偶然に左右されるままになるというのは、あなたにとっては厳格に決定され自由の入る余地がない場合以上に厄介である。とにかく、あなた

の自由な決断は入る余地がない。

因果作用は決定論が成立しない世界でも可能である。非決定論的な世界でも因果作用は存在する。原因と結果の間に確率的なつながりがある場合を考えてみればよい。決定論的でない因果連関が想像できるだろう。つまり、因果性と決定論、因果性と非決定論はそれぞれ両立する。したがって、真の問題は決定論と自由が和解できるかどうかではなく、因果性と自由が和解できるかどうかである。量子力学が決定論は誤りであると主張しても、因果性そのものを否定はしていない。このような意味で、因果性こそが自由を考える上での鍵である。

12 ゼノンのパラドックスと因果的説明

ゼノンのパラドックスの一例、アキレスとカメの競争を具体的に考えてみよう。アキレスとカメは $100m$ 競争をする。アキレスは 1 秒に $10m$ 走り、カメは 1 秒に $1m$ 走る。これでは明らかにアキレスが勝つので、ハンディキャップレースにする。カメはスタートラインから $10m$ 先のところから $90m$ 走ることにする。二人は同時にスタートするが、最初は $10m$ の差がある。 1 秒経つと、アキレスはスタートラインから $10m$ のところまで達する。カメは $1m$ 進むので、スタートラインからは $11m$ のところまで達する。したがって、 1 秒後の二人の差は $1m$ である。次に 0.1 秒後を考えよう。アキレスは $1m$ 進むので、 1.1 秒後の進んだ距離は $11m$ となる。一方のカメは $0.1m$ 進むので、その進んだ距離は $11.1m$ となる。まだ、カメのほうがアキレスより先にいる。次に考えるのは 0.01 秒後である。同じように簡単な計算から、アキレスは $11.1m$ 、カメは $11.11m$ となる。まだ、カメの方が先にいる。次ぎは 0.001 秒後である。このように前の $1/10$ の時間間隔でアキレスとカメの進んだ距離を計算していく。さて、こうしてできあがるアキレスとカメの距離の系列において、いつアキレスはカメを追い抜くことができるのか。この問題に対して、常識はアキレスがカメを簡単に追い越すことを示しているが、系列はいつまで経ってもアキレスの進んだ距離はカメの進んだ距離を超えることができない。したがって、これはパラドックスである。

ゼノンのパラドックスを構成的に考える場合と非構成的に考える場合とに分けてみよう。非構成的な解決の仕方は反事実的な状況を仮に設定して試みることに、あるいは帰謬法を用いてその問題を扱うことである。アキレスが既に a 秒走ったとすればどうなるかを尋ねてみることである。あるいは、アキレスが a 秒後にまだカメを追い越していなかったとすればと仮定して議論を進めてみることである。このような対処の仕方は簡単にアキレスがカメを追い越していることを証明する。例えば、 4 秒後にアキレスがカメを追い越していないと仮定してみよう。アキレスは $40m$ 既に走っている。一方、カメは $14m$ に過ぎない。したがって、この仮定は矛盾している。それゆえ、 4 秒後にはアキレスは既にカメを追い越していなければならない。

一方、構成的な考え方はスタートから一步一步議論を実際に変化する状態に合わせて進める方法である。この方法は反事実的な仮定や帰謬法を用いず、アキレスとカメの走りを再

構成していくという意味で数学の構成主義的な議論の進め方と同じである。このような議論の進め方のすべてについてアキレスがカメを追い越せないというのではない。実際、二人の走りの追跡方法を少し緩和し、1秒毎にアキレスとカメの走りをモニターしたとしてみよう。2秒後にはアキレスはカメを追い越してしまっている。問題はどのような仕方で二人の走りを追跡しても同じように追い越しの確認ができるかどうかである。上述のように時間間隔を $1/10$ にしていく仕方ではそれが確認できないところに問題が生じる。なぜ確認できないかは無限個の位置や時間間隔の存在に由来している。無限の級数の和がゼノンには計算できなかったし、それが計算できるのは19世紀のコーシーまで待たなければならなかった。コーシーの考えに基づいてアキレスとカメのパラドックスを考えてみよう。

まず、無限の数列 $\{S_n\}$ が L に収束する、あるいは極限 L をもつとは次のことである。任意の $\epsilon > 0$ について、ある正の整数 N が存在して、どのような $n > N$ についても、 $|S_n - L| < \epsilon$ となることである。例えば、

$$1/10, 1/10^2, 1/10^3, 1/10^4, \dots$$

という数列の極限は 0 である。このように数列の極限という概念が定義されると、それを無限級数の和を定義するのに使うことができる。無限の級数が $s_1 + s_2 + s_3 + \dots + s_n + \dots$ と表されたとすると、その和を定義するために、次のような部分和を考える。

$$S_1 = s_1$$

$$S_2 = s_1 + s_2$$

.....

$$S_n = s_1 + s_2 + \dots + s_n$$

各 S_i は有限で通常に加算ができるので、問題はない。私たちはすでに無限の数列の極限を定義してあるので、数列 $\{S_i\}$ についてそれが極限をもつなら、上の級数は収束することになる。この結果を容易にアキレスとカメの場合に適用し、アキレスがカメに追いつけることが示される。

非標準解析(non-standard analysis)を用いたゼノンのパラドックスの分析は構成的方法をさらに顕微鏡を使って拡大して分析するものである。事態に密着した方法を更に顕微鏡を通して明らかにするのは反対に、反事実的方法や帰謬法は事態から離れて、一定の距離からその事態を見直すものである。しかし、この好対照の方法はそれぞれの仕方でパラドックスの内側を垣間見させてくれる。

この二つの方法をアキレスとカメの因果関係に適用すると何が言えるだろうか。構成的な因果関係と非構成的な因果関係という区別ができそうである。では、構成的、非構成的とはどのようなことか。よく外挿や内挿という操作を耳に思うが、これらが典型的な

非構成的概念である。「構成的」とはシステムの時間発展をつぶさに観察し、それを正確に追うことである。そこから、そのシステムの時間発展を再構成することができる。これに対して外挿や内挿が入った場合は正確な再構成ができない。

非構成的なものは数学基礎論においては従来積極的な意味が与えられずに来た。非構成性の多様な意味をそのような否定的な意味付けから開放して、さまざまに拡大することができる。具体的な手続きの欠如には実にさまざまなものがある。観測者が問題になる仕方の一つは、視点を自由にいつでも好きな場所に設定できるという点にあるが、その視点の設定は構成的には行われない。どのような視点から見るかは通常非構成的である。これを一般化すれば、認識的な概念のほとんどは非構成的である。構成的な概念は存在化できるのが常であるから、そうでない概念は非構成的にならざるを得ない。構成的、非構成的を対照的に比較すると、現象に肉薄する方法と現象から距離を置く方法と表現できるだろう。帰謬法や反事実的条件法はどの程度距離を置くかの指定は含んでいない。目的論は現象からもっとも距離を置いた方法と言うことができる。というのも、空間、時間的にもっとも距離を置くと、視点は世界の外に出て、超越的な目的論が得られるからである。目的、帰謬法、人間原理等は視点の時間的な移動を自由に行うことによって、非構成的に系列を取り扱っている。

科学的な説明に関する公式的見解には、それは推論である、原因の記述である、現象についての確からしい情報提供であるといったものがある。説明は記述からは独立している。それは記録ではないし、眼を凝らして現象を追うことでもない。それは単なる描写ではない。私たちは日頃記述と説明が一緒になって呈示されることに慣れ過ぎている。映像は記述に見えながら、ナレーションは記述せず、説明する。映画の場面は記述に見えながら台詞が場면을説明したり、説明のための要素として使われたりする。純粹の記述も純粹の説明も実際にはない。それらが融合しているのが私たちの世界である。科学理論もその例外ではない。説明においては私たちの記述、言語、知識、そして推論や推測がすべて総動員される。説明は記述に比べれば遥かに恣意的で、かつ工夫を要するものである。記述できても説明できないものはたくさんあるが、説明できても記述できないものも少なくない。具体的な進化の過程はその代表例である。説明にはいくつかの種類があるが、そのようなことは記述にも言えるだろうか。因果的説明と目的論的説明は説明の代表的な種類であるが、因果的記述や目的論的記述といったものはあるだろうか。記述はすべて因果的であり、目的が未来に実現するような場合には目的論的記述は有り得ないだろう。記述は記録であり、記録に法則を描くことはできない。法則を使って記述する、法則に従って記述する。記録から法則を抽出することはあっても記録が法則をそのまま描くことはない。

説明の現象からの距離の置き方に私たちの工夫が入り込む余地が生じる。それが知識を生み出すことになる。わかることは現象から一定の距離を置いてその現象を眺めることである。この鳥瞰的な見方は単なる遠近法でも、記録でもない。そこには知識の演出が働いている。

ここまでの推論の結果をアキレスとカメのパラドックスへ応用してみよう。微分法と集合論以前の人々のためにこれまでの議論を役立ててみよう。ゼノンのパラドックスは彼らには深刻な問題である。もしそれが記述の問題であるとする、構成的にアキレスとカメの競争を追い、再現できなければならない。ゼノンの述べる分割の仕方にしたがって記録を取っていくなら、その記録のどこにもアキレスがカメを追い抜く記録は現われない。微分法も無限も知らない限り、この問題はそのままパラドックスとして残り、ゼノンのパラドックスはその地位を保つ。しかし、説明という観点からは当時の人々にもパラドックスは見かけに過ぎないことを簡単に説明できる。「100m 走り切ったとしてみよう」という反事実的な条件法も、「仮に 100m 走ってアキレスがカメを追い抜いていないとしてみよう」という帰謬法の仮定も彼らには自由に使えるからである。実際、そのような仮定のもとではアキレスはカメを追い抜いており、どこにもパラドックスはない。つまり、記述のレベルではパラドックスであるが、説明のレベルではパラドックスではない。アキレスとカメの競争の結果は説明によって十分与えられる。但し、その競争の追い抜き場面がどのようなかについては、このような説明では何も述べられていない。

ではカントールの集合論や微分法を知った私たちはどのように異なる解決法を与えるのか。私たちが解決できるのは、私たちが異なる記述の装置を獲得したからである。説明の仕方に新機軸が登場したのではない。現在の私たちも以前の人々も同じ説明の方法しか使用していない。変わったのは記述の方法である。より優れた記述の仕方の獲得によって、それを含む説明はそれを含まない説明より優れることになる。

私たちがアキレスとカメのパラドックスに関心を寄せるのは次の点にある。記述を通じて競争の過程を物理的な過程としてどのように想定すべきか。記述はこの場合ほとんど過程そのものである。私たちがゼノンの論法に惹かれるのは、あらかじめ帰謬法や反事実的な条件法を使って常識が正しいことを確認しながらも、それを構成的に再現するという追確認を求めるからである。記述の表現力が 100%であることを求めるのではなく、得られている説明の追確認に十分な記述を求めるのである。

さて、空間の無限分割性に関するカントール以後の理解は物理学的な内容に数学を適用できることをさらに保証してくれることになったし、確かに非標準解析はそのような技術的な分割可能性を明らかにしてくれた。非標準解析によってアキレスとカメの最後のデッドヒートが描き出せることになった。私たちが内挿によって想像していた部分を数学的なモデルで満たすことができるようになったばかりでなく、カメがアキレスを追い抜く場面さえ可能な場合がある（但し、最終的にはアキレスが勝利する）ということも示してくれた。

アキレスとカメの話はどこか現実離れしている。余りに多くの現実離れがこの話には含まれている。アキレスとカメの x 秒後の位置が正確に計算できるというのも現実離れの一つである。カメは直線的に走らないかもしれないし、アキレスはウサギのように途中で休むかもしれない。このような心配は最初の前提でブロックされるようになってはいるが、実際の物理現象は恣意的なブロックを受けつけない。そのような場合、構成的な因果連関とか

非構成的な因果連関とかは意味をもっているのだろうか。私たちが考える世界はアキレスとカメの話のようにはすべてが決まっていはいない。しかし、全く何も決まっていなわけではない。変化のいくつかの側面は決まっている。しかし、それが何でいつも決まり方が同じというようにはなっていない。したがって、構成的、非構成的の区別は通常明確でない。明確でないということはこのような概念上の区別が無意味であることになるのだろうか。これを無意味とみなすことが実は物理世界に構成的、非構成的という区別をすることを阻んできたのである。構成的とはたとえ大まかであっても決まっているシステムの時間発展に合わせた追跡の仕方、記述であり、非構成的とは大まかな物差しの中で視点を移動させての説明なのである。

13 知識論と懐疑

代表的な知識論のスケッチをしておこう。プラトンは『メノン』で知識と正しい意見は實際上同じものと考えた。そして『テアエテス』では知識は正しい信念であると提案した。初めて基礎付け主義的に知識を考えたのはアリストテレスで、科学的な知識が証明されるにはその前提となる知識が原初的で他のものによって証明できないようなものでなければならぬと主張した。デカルトの認識に対する関心は自分の心的能力を調べることによって、それに基づいて彼が知ることのできるものを発見することにあつた。その結果、精神として彼が存在すること、神が存在すること、外部世界が存在することが確立された（と彼は考えた）。ヒュームは経験的な知識に関して徹底的に疑った。それは帰納法の正当化ができないという彼の推論に見事に表れている。

ヒュームは心と身体を分けるが、デカルトと違って心を実体と見なすことに反対する。それは彼の二つの原理に基づいている。(1) 心の知覚すべては印象と観念に分けられる。(2) すべての観念は印象のコピーである。(知覚は意識を含み、印象は観念より生き生きして、強い。)(2) は次のことを意味している。すべての観念に対して、それを生じさせた印象(の組)があり、観念はそれを生じさせた印象に似ている。したがって、以前の印象からつくられない、印象から独立した観念は存在しない。意識は経験の主体が存在することを示しているという考えに対して、ヒュームは次のように反対する。印象が観念をつくるのであるから、存続する自己の印象をもたなければ、自己の観念ももてない。印象はいつも移り変わっている。存続する自己の印象はどこにもない。それゆえ、存続する自己の観念はない。このヒュームの批判は、時間を通じて存続する、経験の統合としての自己という哲学的な解釈に向けられている。

デカルトとヒュームは経験的な知識に関して際立って異なった結論を得ている。まず、デカルトの推論からみていこう。

- (1) 私は今私の前にあるのが黒板であると信じる。
- (2) 私の現在の信念は明晰にして判明である。

(3) 明晰にして判明な観念は真である。

したがって、私の前には黑板がある。

デカルトの基礎付け主義は、この推論の前提が疑うことができないものである、この推論の結論が私たちの知る命題であること、これら二つからなっている。

デカルトによれば知識は内的に保証可能である。前提はアプリアリに真であるか、内観によって知ることができる。もし主体が結論の真であることを知るなら、そのとき主体は結合前提が真であることを知らなければならないし、また、そのことを感覚経験とは独立に知らなければならない。つまり、知識は内的に保証可能である。

世界の因果連関に関する信念は理性と経験に基づくと考えられるが、ヒュームはこの理性を拒絶する。知識を疑うことは私たちが世界について部分的にしか知識をもてないということである。私たちは今日太陽が昇ったことを知っているが、明日また太陽が昇ることを知らない。それだけではなく、明日以降の太陽についてもつどのような期待に関しても合理的な正当化の仕方をもてない。

常識的には、予測や一般化について合理的に信じるには、それら信念についての多くの証拠があればよいと考えられる。次の推論は演繹的ではないが、完全に信頼できるものと考えられる。

(一般化)

私は多くのエメラルドを見てきた。それらはことごとく緑色であった。したがって、すべてのエメラルドは緑色である。

(予測)

私は多くのエメラルドを見てきた。それらはみな緑色であった。したがって、私が次に見るエメラルドは緑色だろう。

しかし、ヒュームによれば、このような確信は合理的に正当化できない。では、彼はそれをどのように推論したのか。ヒュームによれば、常識的な確信に思えるのは私たちの習慣にすぎない。それは人間の本性であり、私たちが実際に振る舞う仕方にすぎない。そして、私たちはこの心の習慣を合理的に正当化する推論を知らない。

前述の推論では、前提から結論が得られないので、結論を得るためには新たな前提が必要であるとヒュームは考える。自然の一様性原理(PUN)がその前提で、未来は過去に似ているという主張である。帰納法を使った推論はみなこの原理を仮定しなければならない。では、この原理は正しいだろうか。

(ヒュームの推論)

- (1) すべての帰納的推論はその前提として *PUN* を必要とする。
 - (2) 帰納的推論の結論が前提から合理的に正当化されるなら、それらの前提も合理的に正当化されなければならない。
 - (3) だから、帰納的推論の結論が正当化されるなら、*PUN* についての合理的正当化がなければならない。
 - (4) *PUN* が合理的に正当化されるなら、*PUN* についての正しい帰納的推論、あるいはそれについての正しい演繹的推論がなければならない。
 - (5) *PUN* についての正しい帰納的推論はありえない。というのも、そのような推論は循環するからである。
 - (6) また、*PUN* についての正しい演繹的推論もない。というのも、*PUN* はアプリアリに真でもなければ、私たちが行なった観察から演繹的に導出もされないからである。
 - (7) したがって、*PUN* は合理的に正当化できない。
- それゆえ、私たちは予測や一般化の形式をもつ信念について合理的な正当化をすることができない。

(推論の吟味)

- (4) – (6)を詳しく見てみよう。自然の一樣性とは未来は過去に似ているという主張である。これは帰納法に基づいて真であると知ることができるものであろうか。もしそうなら、帰納的な推論は次のようになるだろう。

自然は私の過去の観察において一樣であった。

それゆえ、自然は一般に一樣である。

この推論は帰納的であり、ヒュームによればすべての帰納的な推論は *PUN* を前提としてもつから、この推論は循環することになる。では、*PUN* の演繹的な正当化はできるだろうか。ヒュームの答えはノーである。上の推論は演繹的に妥当ではない。また、それは定義の問題でもない。

ではヒュームの懐疑論は克服できるのか。帰納的な推論には *PUN* が用いられなければならない。*PUN* は正確に何を意味しているのか。次のいずれなのか。

未来はあらゆる点で過去に似ている。

未来はある点で過去に似ている。

PUN の正確な特徴づけは困難である。これはヒュームの推論の再定式化を示唆している。そのために、次の二つの言明を比較してみよう。

(1) すべてのエメラルドは緑色である。

(2) 2050年まですべてのエメラルドは緑色であり、その後は青色である。

どちらの方がより妥当性が高いか。(1)の方が(2)より妥当性が高いと思われるが、それをどのように正当化するかが哲学の問題となる。

14 基礎付け主義と信頼可能性

懐疑論を克服する一つの方法が知識に関する基礎付け主義であり、知識をその土台となる基礎から階層的に積み上げていくことによって確実な知識の全体を構成しようとする。原理から出発し、そこから個々の知識を導き出そうとする試みである。したがって、基礎付け主義の知識についての見方はユークリッドの幾何学の構成に似ている。あるいは、建物の土台とその上に築かれる上部構造で考えてもよい。デカルトもヒュームも懐疑をバネにして哲学的な考察を展開したが、二人の懐疑の違いを次の表で確認しておこう。

ヒュームの問題	3 予測と一般化	太陽は明日も昇るだろう。
	2 現在と過去の観察	太陽は今昇っている。 私が観察した日には太陽がいつも昇っていた。
デカルトの問題	1 疑うことの出来ない信念	私は今太陽が昇るのを見ている。 私は自分が観察した日には太陽がいつも昇っていたことを思い出している。

ヒュームもデカルトも基礎付け主義者である。その理由は、二人とも、もし信念が合理的に正当化される(知られる)なら、それはより低いレベルの事柄への関係に基づいて正当化される(知られる)、と考えたからである。上の表の1、2、3が各レベルを表わしているとなれば、ヒュームの場合、レベル3の信念が正当化されるのであれば、それはレベル2の事柄だけを基礎にして正当化されなければならないことになる。また、デカルトの場合、レベル2の信念が正当化されるのであれば、それはレベル1の事柄だけを基礎にして正当化されなければならない。注意しなければならないのは、彼らは問題になっている信念が実際に正当化されると主張しているのではない。正当化のために何が真でなければならないかを主張しているのである。

デカルトとヒュームの結論はまったく正反対である。デカルトはレベル1に基づいてレベル2の信念は正当化できると考えたが、ヒュームはレベル3の信念はレベル2からは正当化できないと結論した。デカルトが各レベルを結びつけるのに必要だったのは神の存在と

神が欺かないことであり、ヒュームの場合は自然の一様性原理であった。

もし合理的な正当化が演繹的になされるのであれば、いずれの懐疑論が正しいだろうか。そのために、次の主張を考えてみよう。

(主張)低いレベルの言明は高いレベルの言明の正当化には不十分である。

(1)今あなたの前には黒板が見えている。

(2)今あなたの前には黒板がある。

(1)から(2)が導き出せるかどうかは、あなたがどのような想定のもとに(1)を考えるかにある。(1)だけから演繹的に(2)を導き出すことは不可能である。そのような背景とする知識や信念なしには基礎付け主義は懐疑論に陥る。

このデカルトの知識論とは異なり、結合前提は内観やアプリアリな推理によって知られる必要はないというのが知識の信頼可能性理論(reliable theory of knowledge)である。温度計とその温度表示をアナロジーに使うことでこの考えを説明してみよう。

温度計の目盛りの値が外界の温度を表示する 君の信念が心の外の世界を表示(表象)する

温度計の目盛りの値が正確(不正確)である 君の信念が真(偽)である

信頼できる温度計とはどのようなものか。それは偶然に目盛りの値と温度が一致するものではなく、いつも一致するものでなければならない。そのための信頼性の条件は二つ考えられる。(1) それは正しい環境で使用されなければならない。(2) 温度計の内部の構成は正しくなければならない。信頼できる温度計がその計る温度に関係しているように個人が命題に関係しているなら、その人はその命題を知ることができる。これが知識の信頼可能性理論の直観的な内容である。

S は p を知っている

(1) S は p を信じている

(2) p は真である

(3) S がいる状況で、もし S が p を信じるなら、p は真でなければならない。

この特徴づけはデカルトの場合と随分異なっている。参考のためにデカルトの場合は次のようになる。

S は p を知っている

- (1) S は p を信じている
- (2) p という私の信念は明晰にして判明である
- (3) 明晰にして判明な観念は真である
- (4) それゆえ、p である

また、この特徴づけを使うと、デカルトの黒板についての推論は次のようになる。

- (1) S は今 S の前にあるのが黒板であると信じる。
- (2) S は自分のいる環境で、その前に黒板がない限りその前に黒板があると信じないだろう。

それゆえ、S の前には黒板がある。

上の推論は「知る」を含んでいない。では知識と関係ないのか。信頼可能性理論では、S は結論をその前提が真であるゆえに真であることを知るのか。S は前提が真であることを知る必要はない。また、S は前提に対して感覚経験とは独立の推論をする必要もない。これはデカルトと異なる点である。この理論では「S が p を信じている」ことが S と p 以外のものに依存している。それは S の置かれた環境である。同じ S、p であってもその環境が異なれば、S は p を知っていることにもなれば、知らないことにもなる。

ところで、知識はどのような条件を満たしたならば正しいと言えるのだろうか。まず、S が p を知っているとはどのようなことなのか。従来、S は p を信じ、p は真であり、S が p を信じる正当な理由があるという三つの条件が考えられてきた。これは正当化された正しい信念とまとめることができる。しかし、このような特徴づけは E. Gettier (“Is Justified True Belief Knowledge?”, *Analysis*, 1963, Vol.23, pp.121-23)の次のような例から否定される。

(物語) A は事務所に勤めており、誰かが直に転勤させられることを知っていた。信頼できる上司が A に転勤させられるのは B であると告げた。その時、A は B の財布に 1 万円あることを知った。

- (1) B は転勤させられ、財布に 1 万円もっている。
- (2) だから、転勤させられる人は財布に 1 万円もっている。

このように A が推論することは正しいだろう。しかし、実際は転勤させられるのは A で、そのことを A は知らず、その時 A もちょうど財布に 1 万円もっていたとする。その時、A は (2) を信じ、(2) は真である。A はそれを (1) から演繹したのであるから、(2) を信じることは正当である。(1) は偽であるが、A はそれを真であると考えた十分な理由をもっている。したがって、(Gettier によれば) A は正当に (2) を真であると信じるが、(2)

が真であることは知らない。

まだわかりにくいかも知れないので、さらに二つ例をあげてみよう。

(Russell の例) 公園の信頼できる時計：散歩の途中で9時30分であった。君はそれが信頼できる時刻であると信じることができる。しかし、時計は一日前に止まっていたとしてみよう。それを君は知らない。君は信頼できる時刻 9時30分を信じることができるが、君はそれが正しい時刻であることを知らない。

(別の例) 廊下の途中にその後ろと同じ写真を衝立てにつけて置いてある。それを本物と思う悪漢はその写真のものを銃で撃つ。銃で撃つ前の悪漢の信念は正しいが、銃で撃ったのは本物ではない。

上の例に共通するのは何であろうか。いずれの主人公も信じている命題に関するきわめて信頼できる証拠をもっているにもかかわらず、それが誤り得ないという保証がないことである。ここから導かれるのが懐疑論(Skepticism)であり、上のことから次のような推論が成り立つことになる。

もし S が p を知っているなら、 S が p を信じるのが誤りであることが不可能である。
 しかし、 S が p を信じるのが誤りであることは可能である。
 したがって、 S は p を知らない。

これは正しい推論であるが、その結論は信じがたい。従って、この結論を否定したいのであれば、前提のいずれかを疑って見なければならぬ。

15 経験論と実在論

経験論は文字通り経験についての主張である。まず、経験が必要である、次に、それが十分であるという主張である。何についての必要十分なのか。考えられるのは実在について必要十分であるというもので、その強い主張は「現象的なもの以外にはなにも実在しない」というものである。弱い主張は経験を越えた実在を認めるが、それについて私たちは知識を得ることができないというものである。このような存在論的な区別を離れ、知識という観点から見ると、すべての知識は経験に由来し、経験が知識にとって必要十分という信念が経験論ということになる。経験世界とは私たちが感覚的に経験する世界である。ロックは生まれただての人間の心は *tabula rasa* で、経験によってそこに一般原理や個々の知識が書き込まれていくと考えた。

実在論にも多数ある。例えば、真理が何であるかについての主張として、あるいは、どの

ような真理があるかについての主張として実在論が述べられる。例えば、私たちの経験とは独立に実在があり、その実在を知ることが真理であるというのが実在論ということになる。ここでは真理の本性についての見解としての実在論を考えてみよう。それは意味論的な主張で、文の集合の実在論的な解釈では、人間の思考や言語とは独立に文の真偽が決まると主張される。そこから、文は心から独立した実在について述べ、その実在に真偽が依存していることになる。この見解に反対する立場が検証主義である。文の真偽はその検証の仕方を抜きにしては決められない。それゆえ、文の真偽は検証する私たちに依存し、したがって、それに合わせて真理概念を解釈し直さなければならないというのがその主張である。これは上述の経験論のタイプである。すると、出てくる問題は、真理は実在論的に理解されるべきか、それとも経験論的に理解されるべきかとなる。

真理の具体例は科学理論であるから、それをもとに考えてみよう。科学は経験科学であり、私たちの経験する世界についての経験的研究である。科学における経験論は次の条件法によって展開されてきた。もし私たちの知識が経験を超越することができないなら、(1)知ることのできる命題の集合の範囲と、(2)何が真かを推論するのに適切な方法の集合の範囲とを定めることができない。ここから、経験論者は観察命題と理論命題（例えば、一般法則や原理）を区別し、理論命題の真理が観察命題からは推論できないという主張を明らかにしようとした。ところで、(2)の経験論者が許容する推論の範囲は何か。演繹や単純帰納を使った推論は経験主義者も認めるが、アブダクションはどうであろうか。最善説明への推論は経験を説明するための仮説設定であり、これを別の言い方をすれば、観察的な前提から理論的な結論を導き出すことである。

実在論に反対するフレーゼンの構成的経験論では、科学は観察可能な対象についての言明の真理を語るべきで、観察できない対象に関する真理については何も言えないことになっている。観察可能とは当然私たちにとって観察できるという意味である。しかし、アブダクションが許されるなら、観察可能な前提から観察不可能な、理論的結論を得ることになり、フレーゼンの主張に反することになる。さらに、構成的経験論には問題がある。「すべてのカラスは黒色である」という言明は「すべての x について、その x がカラスなら、その x は黒色である」を意味しており、したがって、この世界のすべてのものについての言明である。もし世界が観察できないものを含んでいたら、カラスの場合のような一般化がその観察できない対象についてもなされなければならない。（カラスでさえ、これから生まれるものは観察できない。）

アブダクションを使った実在論を擁護する推論がパトナムによってなされたことがある。ある科学理論が正確な予測をして受け入れられているとする。では、なぜこの理論は成功しているのか。何がこの理論の予測の正確さを説明するのか。もしこの理論が観察できないものを仮定しているなら、その理論の予測の成功は、仮定された対象が文字通り存在し、その性質が理論の主張する通りのものであるという仮説によってもっともうまく説明できる。これはアブダクションであり、観察できる前提からそうでない結論を引き出している。

アブダクションの使用が論点先取か妥当なものなのかは意見の分かれるところである。理論命題は最善の説明のために要請され、それだけからその内容が実在するというのは確かに論点先取のようにみえる。

経験論者は私たちの知識が経験を超えることができないことに重要な意味があることを示そうとするし、実在論者は観察できないものを知る私たちの能力が観察できるものを知る能力と同じように強力であることを示そうとする。経験論者の考えは科学的推論の能力を制限するという欠点をもつのに対し、実在論はその能力を拡大し過ぎるという欠点をもっている。

では、第三の道はあるのだろうか。実際の経験的知識とその獲得を実情に即して考えた場合、観察できないものについての知識はテストできる命題に現れる語彙に制限を加えるだろうか。何も制限は加えないようにみえる。しかし、経験的に識別できないものについてそれを識別することは科学にはできない。識別できない二つの理論のいずれが正しいかは判定できない。これは経験論だけでは知識にとって不十分であり、だからといって実在論をそのまま鵜呑みにもできないという表明である。科学理論は経験論と実在論の間でその特徴づけを待っている。

16 主観性と客観性

主観性はあるクラスの言明の真偽がその言明をつくった本人の心的状態に依存していることを示すのに使われる。認識論での主観性概念は、知識がある人の知覚に制限されることである。感覚質（色や匂い）の主観性とは感覚器官によって経験される性質が物理的なものに属さず、解釈によるものであることを認める場合に使われる。形而上学では、主観性は唯我論や観念論を含み、その典型例はバークリーの「存在するとは知覚されることである」という主張である。

客観性という語は事物の真理がそれを観察する者から独立しているという見解を指している。客観性概念は事物が心から独立に存在することを意味する。認識論での客観主義的な立場は、認識主体とは独立の真理を認め、したがって、真理の対応説を帰結する。しかし、観念論者は客観性という語を思考における存在こそが真の存在であることを表すのに用いる。プラトンのイデアはそのような存在である。近代の哲学では客観的な領域が存在することを容認するが、それに到達することはできないと考える。カントによれば、私たちの知識は現象的な世界に制限され、物それ自体に直接に近づくことはできない。

このような説明は味気無いので、主観的なものと客観的なものの違いを倫理や道徳の言明を例に考えてみよう。昔から倫理や道徳は客観的な事実であるかどうか論じられてきた。主観的なものに過ぎないという立場では倫理や道徳は単なる意見である。倫理の主観主義 (*Ethical subjectivism*) は倫理や道徳には客観的な事実是对応していないと考える。すべての真なる言明は事実に関する言明であるが、倫理や道徳の言明は当為の言明であり、したがって、倫理的言明は真でも偽でもないというのがその理由である。

「殺人はいつも道徳的に正しくない」

「殺人は時には道徳的に許される」

上の二つの文は共に倫理的な言明であり、かつ一方の否定が他方になっている。したがって、いずれか一方が正しいはずであるが、主観主義はこの事を否定する。つまり、主観主義によると上の言明は共に真ではない。

一方、倫理の客観主義(*Ethical realism*)は、誰かの意見とは独立の、客観的な倫理的事実があると主張する。この二つの極端な立場の間に倫理の規約主義(*Conventionalism*)がある。倫理的な事実は存在するが、それは誰かの意見によって真であるというのがその主張である。誰かの意見は場合によって異なる。国の場合もあれば、個人の場合もある。個人の規約主義の代表例がサルトルである。

事実と当為の違いが倫理の主観主義と客観主義を区別することになったが、それはどのような違いなのか。

(ヒュームのテーゼ) 倫理的な言明は事実に関する言明からは正しく演繹できない。

ヒュームのテーゼから主観主義の主張は導出できるだろうか。その導出のためには還元主義的な前提、つまり、もし倫理的な言明が事実言明から演繹できないなら、「倫理的言明は真でも偽でもない」という言明が正しくなければならない。そして、このテーゼは主観主義者にとっては正しい。すると、次の推論が得られる。

もし倫理的言明が事実言明から演繹できないのであれば、倫理的言明は真でも偽でもない。実際、(ヒュームのテーゼから)倫理的言明を事実言明から導き出すことはできない。それゆえ、倫理的言明は真でも偽でもない。

ところで、次の推論は正しいだろうか。

もし生物学的言明が物理学の言明から演繹できないならば、生物学的言明は真でも偽でもない。実際、物理学の言明だけから生物学的言明を導き出すことはできない。それゆえ、生物学的言明は真でも偽でもない。

この推論が正しいと思う人はいない。すると、倫理的言明に関する推論も疑ってみなければならない。

ヒュームのテーゼを今世紀さらに明確にしたのがムーアである。彼は *Principia Ethica* (『倫理学原理』) で自然主義的誤謬 (*naturalistic fallacy*) について考察した。倫理的性質と自然的

性質は異なり、行為についての倫理的性質はその自然的性質ではないという主張に基づいて、「道徳的に正しい」という表現は「快樂を最大にし、苦痛を最小にする」という表現と同じことを意味していないと考えた。そこから、倫理の理論は倫理的性質と自然的性質を同一視してはならないことになり、同一視する理論は自然主義的な誤謬に陥っているという結論を得た。ここで次のような類似の推論を考えてみよう。

「温度」という表現は「平均運動エネルギー」という表現と同じことを意味していない。それゆえ、「温度」が表す性質と「平均運動エネルギー」が表す性質は同じではない。

この推論は誤まっているだろうか。温度は平均運動エネルギーであるという熱力学の基本的な知識を信用するなら、この推論は誤まっている。つまり自然主義的誤謬は誤謬ということになる。

ここで教訓を一つ述べておこう。客観性や主観性は優れて哲学的な概念に見えるが、それらだけをいくら考えていても何も出てこない。高価なコンピュータも適切に使われてこそ効力を発揮するように、客観性や主観性も適切に使われてこそ議論が進展する。

17 二つの同一性

「あるものが以前のものと同じかどうか」と「あるものと別のものが同じかどうか」とは同一性に関する重要な、しかし、異なった問いである。いずれの問いも私たちの日常生活で時々顔を覗かせる。ここではこれら二つの問いの典型について考えてみよう。

(人格の同一性：ある人が以前のその人と同じとは？)

心や身体についての異なる考えが引き起こす問題は多岐にわたっている。その一例が人格の同一性である。椅子や机が同じものか否かには比較的明瞭で、一致した基準があるが、心をもった人格の同一性の基準は曖昧である。10年前の私と現在の私が同一人物とみなされるのはどのようにしてなのか。あるいは、一般的に時間を通じての人格の同一性は何によって決まるのか。このような問いに対する答えは椅子や机の場合に比べると簡単ではない。どこに簡単でない理由があるのか。通常、椅子や机の同一性はそれらの物理的な連続性(*physical continuity*)が基準になっている。しかし、私たちは心をもっているので、さらに心的な同一性も必要になる。そして、これにはものの場合に対応して心理的な連続性(*psychological continuity*)が基準と考えられている。ここまで述べてくると、人格の同一性は心身二元論が前提になっていることに気づくだろう。気づいても気づかなくても、身体と心の両方の同一性が必要という話が不自然に聞こえないのは、私たちがデカルトの伝統の中に生きていることの証である。

では、心理的な連続性とは何か。昨日のことを思い出す人は今日のその人と心理的に連続している、つまりは記憶の連続性が心理的な連続性と考えられている。通常は身体的同一

性と心的同一性の二本立ての基準によって人格の同一性が捉えられることになる。このような答えは一見すると確実に見えるが、実は多くの疑問をすぐに引き起こす。例えば、意識を失い記憶喪失に陥った人でも、覚醒したとき「私は誰か」と自問できるが、そのときの「私」は誰なのか。「私」は頼るべき記憶を失っている。余りに多くの疑問のため、大抵は人格の同一性に到達する遥か前に追求を放棄してしまうことになる。次のような状況でどのように同一性が言えるか。思考実験してみよう。

- ・船の外板：船を造っている板を一枚ずつ取り替えて、すべての板を新しいものにしたとする。修理された船は以前の古い板の船と同じと言えるだろうか。また、古い板を集めて同じ型の船を再現したときの船は以前の船と同一であろうか。
- ・脳移植：ある人の脳を別の人に移植したとき、移植された人は誰と同一なのか。また、脳を半分移植した場合はどうなるか。
- ・複製：ある人の複製をつくったとき、いずれの人が以前の人と同一なのか。

このような思考実験によってわかるのは、私たちの同一性についての基準が半ば物理的な連続性と半ば心理的な連続性の組み合わせの規約に基づいてきたことである。全くの一致のある規約はないが、だからといって規約がないというわけではない。

(同一説：心が脳と同じかどうか？)

心的性質(状態、過程)は脳の性質(状態、過程)と同じというのが同一説の主張である。痛みとC-線維の興奮は同義語ではないが、にもかかわらず同じ性質を指示する。痛い状態にあるという性質とC-線維が興奮している状態にあるという性質は同一である。心と脳の何が同じかについては、感覚は脳過程と同一である、あるいは心的状態は脳と中枢神経系の状態と同一であると僅かな違いがありながら、心的なものとは物理的なものの同一であることが主張される。また、フレーゲの意味と指示の違いに従って、心的用語と神経生理学的用語は意味が異なるが、指示は同一であると主張される場合もある。

既述のライプニッツの法則は、二つのものが同一なら、それらがもつ性質もすべて同じである、つまり、 $(x)(y) [(x = y) \supset (F(x) \supset F(y))]$ であることを主張していたが、この法則から、心的出来事と脳の活動は異なる性質をもつので同一ではないという同一説批判が出てくる。別の同一説批判によると、心的記述と物理的記述は空間的な局所性、客観的観察可能性、志向性という三つの点で異なっている。ライプニッツの法則から一つの記述によって述べられる性質は別の記述によっても述べられるが、一つの記述の性質は他の記述には適用できない。それを見てみよう。

(空間的な局所性を使った推論)

1 脳過程は空間的出来事である。

- 2 感覚は空間的ではない
- 3 同一性の法則は一方について正しいものは他方についても正しいと主張する。
- 4 それゆえ、二つのものは同一ではない。

客観的観察可能性に関しても類似の推論を展開できる。脳過程は観察可能であるが、経験する感覚はそうではない。また、志向性は心的状態に適用されるが、脳の活動には適用されない。

このような批判を見てくると、別タイプの同一説の可能性が考えられる。その代表はローティやファイアーラーベントの主張で、心的なものを含む話は脳の活動だけを含む話に置き換え、消すべきであるというものである。すべての心的現象についての話は幽霊現象についての話と類似している。

ところで、同一であるとする際の判定理由はいつも心的な出来事、心的な記述をもとに行われる。これは遺伝子とある DNA 断片の同一性が言われる際にも生じることである。心的出来事と脳の活動が同一という判定はいつも心的出来事をもとに行われる。この一方向性は真の同一性確認のの障害になっていないか。遺伝子型と表現型の間の複雑な関係は心と脳の間にもそれ以上の複雑な関係が存在することを示しており、したがって、心と脳の状態や出来事の全くの同一という主張は維持が困難で、それに係わらない還元や消去に議論がシフトすることは当然であるのかもしれない。次の推論を見てみよう。

- 1 ある心的性質 M があって、それを一つの場合に実現する物理的性質の組 $\{P_1 \vee P_2 \vee \dots P_n\}$ があるが、すべての場合に実現する特定の物理的性質はない。
- 2 もし心的性質 M がある物理的性質と同一であれば、それは $\{P_1 \vee P_2 \vee \dots P_n\}$ のどれかと同一である。
- 3 しかし、 $\{P_1 \vee P_2 \vee \dots P_n\}$ のどれも 1 より M と同一ではない。
- 4 M と共通の外延をもつ物理的性質は存在しないので、 M と同一な物理的性質は存在しない。それゆえ、同一説の主張は誤っている。どのような物理的性質とも同一でない心的性質が存在する。

科学が経験的であるのと同じ意味で、同一説は経験的な主張である。そして、それは明らかに唯物論の立場に立つ。したがって、すぐに同一説の説明構図の方が二元論より単純であることがわかる。その主張は単純明快で、心と脳は同一で、それは経験的に確認されなければならない。その結果として、心は還元・消去され、脳の振る舞いによって置き換えられることになる。脳に関する研究者が心について何か述べようとするとき、脳に関する研究は心に関する研究でもあると考えるのは、背後に同一説があると理解しやすい。しかし、認知科学や脳科学の内容を必ずしも同一説の立場に立って理解する必要はない。その主張の多くは同一説ではなく機能主義に基礎を置くものである。

18 心身二元論

形而上学での二元論は、世界は互いに全く異なり、それゆえ互いに他で置き換えることのできない二つのものから成立していると主張する。そのような共役不可能な二つのカテゴリーの例は、心身、善悪、普遍と特殊、現象と本質等がある。二元論とは異なる主張が一元論や多元論である。二つ以上のものを認めることはそれだけ説明や理解が容易になるが、一方で新しく出てくる問題は二つ以上のものが互いにどのような関係にあるかである。物理的なもの、生命的なもの、精神的なものとの関係はどのようになっているのか、世界は階層的な構造をもっているのかという問題である。なかでも、心身の関係はそのような代表的な問題である。心身の二元論を初めて明確に主張したのはデカルトであった。そこでデカルトの心身二元論とそれに対する様々な立場を考えてみよう。

イリアスとオデッセイアでの心、デカルト以前と以後の心、これらの間には心に対する考えの大きな進展がある。それは心概念の進化と呼ぶに相応しい。特に強調していいのはデカルト以後の心の扱いである。二元論は世界には二つのタイプのものがあると主張するが、大抵の場合、一つは物理的で、他は非物理的である。なかでも身体は物理的なものの典型であり、心は非物理的なものの典型である。この極めて常識的な二元論は二つの基本型をもっている。実体そのものが異なるという実体二元論と性質が異なるという性質二元論である。一方、一元論では世界に一つのタイプのものしかなく、そのものが物理的あるいは非物理的であるのに応じて唯物論と観念論と呼ばれる。

一元論とか二元論という区別は現在の私たちの常識に合っているだろうか。誰も世界は一種類のものからできていると思っていないし、無限に多くの種類のものからできているとも思っていない。化学元素を文字通り物質の種類と認めるならば、世界は多元論的である。そこで化学元素はみな物質的なものだと考えれば物質一元論となるが、そのようにみなす十分な根拠を私たちはもっているだろうか。十分な根拠がなければ一元論は恣意的な分類の結果に過ぎないことになる。元素をさらに分解して素粒子のレベルまでもっていても、やはり同じように何元論かは今の私たちには断定できない。何元論かという議論はややもすると、世界の構成に関する真摯な議論というより、私たちの世界に関する考え方を整理するためでしかない場合もある。

さて、心と身体はどのような関係にあるのか。この問いに対して、心と身体は世界の二つの異なる実在であり、それらの間には因果的な相互作用がある、というのがデカルトの解答 = 主張である。この心身相互作用の二元論と対照的なのが、今世紀に登場する心身(脳)同一説である。心と脳は同一で、心的な性質は物理的な性質であるというのがこの説の主張である。心的な用語と神経生理学的な用語の違いは、「水」と「 H_2O 」という用語の違いと同じである。用語そのものは異なってもその指示対象は同一であり、したがって心と脳は同一であると主張される。同一説以外の今世紀の解答もすべてデカルトの主張を出発点にしてそれを批判する、あるいは継承するという形にまとめることができる。デカルト的

な解答は自然主義的な解答ではないこと、今世紀の心についての諸説は自然主義的であること、この二つから自ずと今世紀の心についての諸説の大半は反デカルト的ということになる。

現代の心身相互作用論でもその相互作用の仕方に関しては説ごとに異なっている。相互作用論の最も重要な点は相互作用の特徴づけにある。心がデカルトのようにものとは異なる実体であると、自ずとその相互作用はものの世界だけで特徴づけることができなくなってしまう。認めやすい相互作用は自然の中の相互作用であり、それは心がこの自然の中であって、そこで相互作用するということを意味している。

まず、デカルトの主張を批判的に吟味してみよう。

デカルトは彼が心をもつことを疑うことができない。

デカルトは彼が脳をもつことを疑うことができる。

(ライプニッツの不可識別者同一の原理 (= 二つのものが同一とはそれらが同じ性質をもつということ) を使って)

それゆえ、心と脳は別のものである。

精神をもつことを疑うことはできないが、身体をもつことは疑うことができるとデカルトは考える。そこから、精神と身体が異なることをライプニッツの法則から導き出す。ここで次のような例を通じて上の推論が誤りを含んでいることを見てみよう。その誤りは意味論的な指示(reference)と意味(sense)の違いに関係している。それは哲学では命題的態度(propositional attitude)と志向性(intentionality)として考えられてきたものである。

伊作は明けの明星を観察したい 伊作は金星を観察したくない

ここにライプニッツの法則を適用するなら、伊作が観察したいという性質は金星にないことから、ライプニッツの法則から明けの明星と金星は異なるものになる。無論、これは誤っている。明けの明星と金星は同じ対象を指している。

伊作は明けの明星を観察する 伊作は金星を観察する

上の文では明けの明星と金星が同じものである。これは、伊作が明けの明星を観察したい、金星を観察したくないという事実と何ら矛盾しない。これは最初の二つの文が示している事実が、明けの明星と金星が異なる性質をもつことを示していない。同様に、

私は脳をもっている 私は心をもっている

という二つの文について、一方を疑うことができ、他方は疑うことができないということから、脳と心が異なる性質をもち、異なる対象であるということは結論できない。

ここで重要なのは「疑う」や「望む」は命題に対する私たちの態度であるということである。たとえある命題が望まれ、他の命題が望まれなくても、最初の命題が述べているものと二番目の命題が述べているものが異なるということは導かれない、というのが上の例から言えることである。「私が脳をもっている」と「私が心をもっている」が異なった命題であって、一方は疑うことができ他方は疑うことができないとしても、そこから脳と心が異なるということが結論できないということである。したがって、デカルトの論証は誤っていたということになる。

伝統的な心身問題は、デカルトの場合、意識 - 身体問題といってもよい。意識と身体がどのような関係にあるかという問題である。デカルトは心と意識を同じとみなしたので、心的状態は意識の状態のことである。彼の実体二元論は、したがって、意識的なものと延長するものの二元論である。

19 二つの行動主義

行動主義は心的出来事に関する話は観察可能な行動に関する話に翻訳されるべきであるという哲学的主張であるが、二つの行動主義があり、しばしば混同されてきた。

1 論理的行動主義

論理的行動主義は心理的な用語は意味を損ねることなく行動の用語に翻訳することができるという見解である。翻訳は同じ外延をもつという関係に弱めて考えられる場合もある。また、行動の他に神経的、物理的な用語を含めて翻訳を考える場合もある。論理的行動主義の代表は ライルで、彼の考えは *The Concept of Mind*(1949)で展開され、心身問題に関して大きな影響を与えてきた。彼の行動主義は私たちの日常世界での心的な用語の意味に関する主張で、私たちが個人の思考や信念、意図や感情について語るときに何が意味されているかを明らかにしようとする。その点で科学に関する主張である方法的行動主義とは異なっている。ライルはデカルトの二元論を批判し、それを「機械の中の幽霊ドグマ」と表現する。観察できる出来事が身体と不可分に結びついているように、内的な出来事は心と不可分に結びついている。これは誤っているというのが彼の考えである。心的用語は人が行動する仕方に言及しているのであり、内的な心的状態に言及してはいない。この心理主義への反対は次の論証を考えればよい。

もし心的状態が行動の内的な原因であるなら、私たちは他人の心的状態についての知識をもてないだろう。ところが、私たちは他人の心的状態についての知識をもつ。それゆえ、心的状態は行動の内的原因ではない。

他人の心的状態についての知識は類推によるというのが伝統的な考え方である。これは「他

者の心の問題」として議論されてきた。ところで、素粒子がものの運動の内的な原因であることは、私たちが素粒子について知ることができないことを意味してはいない。同様に、心的な状態が行動の内的な原因であるという見解は、その内的な原因が存在するかどうかに関する懐疑主義にすぐに結びつくわけではない。ここで次の区別は重要である。

命題 p は観察可能な対象に関するものか、命題 p は観察によってテスト可能か。

これら二つの問いは基本的に異なった問いであり、同じ解答をもつ必要もない。素粒子や他人の心的状態に関する主張は直接に観察可能な対象についての命題ではない。しかし、それら命題について観察によるテストを実行することができる。

ライルの二番目の反対は自然主義的誤謬(naturalistic fallacy)あるいはカテゴリーミステイクである。「デカルトは心をもっている」、「デカルトは身体をもっている」はそれぞれまっとうな命題であるが、「デカルトは心と身体をもっている」という命題はライルによれば許されない。心と身体は別のカテゴリーに属するからである。これはほとんど二元論の否定と言ってもよい。この主張のどこがおかしいのか。

- (i) デカルトは心をもち、脳をもつが、それら二つは異なる。
- (ii) デカルトは心をもち、脳をもつが、それらは同一である。
- (iii) 人が心をもつのは正しく、人が脳をもつのも正しいが、前者が後者の上に存在するものであると考えるのは誤りである。

これらは言語の規則に反しているだろうか。反してはいない。またライルは傾向性(disposition)を使って心的状態や性質を分析しようとするが、傾向性で状態を置き換えることはできない。痛みの行動がなくても痛みは存在可能だからである。

ライルは自分を心理学的行動主義者とは考えないが、痛みという例外を除いて、すべての心的状態は行動を通じて分析可能であると考ええる。また、行動に対応する心的状態の存在も否定する。そして、私たちは意識的な心的出来事をもっていないと結論する。ライルは内的状態が傾向性を説明するのに因果的に関連していない、したがって、適切ではないと考える。心的出来事というのはある仕方で行動する傾向性のもとでの人の行動について予測することである。特定の信念は特定の行動を記述するに過ぎない。

2 方法的行動主義

デカルト的な二元論では心は非物理的で、第三者には観察できないゆえに心理学は科学としては成立不可能である。科学として成立するには、刺激、条件付け、反応といった観察可能なものにだけ言及する行動の理論として心理学を確立しなければならない。方法的行動主義は心理学固有の領域は行動の研究であると考ええる。観察できない内的状態に訴える

ことは方法論的に不可能である。このような考えはワトソンによって最初に主張された。彼は人間の活動のうちで公的に観察可能な特徴だけを使って心理学を構成し、その結果、心理学を科学として確立しようとした。人間の心的状態を正確に解釈する唯一の方法は人間の行動を観察することによってであると彼は考えた。彼の弟子であるスキナーは心的状態を記述するのに通常使われる多くの用語と行動を結びつける実験を行うことによって行動主義的な理論を構成した。心的状態の存在に関しては心理学的行動主義者は不可知論を採る場合と、フロジストンやカロリックと同じとみなし否定する場合の二つに分かれる。方法的行動主義は1913年から1957年までアメリカの心理学の主流であった。1957年、チヨムスキーはスキナーの *Verbal Behavior* を鋭く批判したが、その内容は次のようであった。私たちは無限に多くの異なる文をつくり、理解することができる。行動主義は過去の反応と条件付けの歴史を強調するが、どのようにこの無限の内容を扱うか知らない。最初に経験する新しい内容をどのように習慣的な刺激 - 反応の図式で理解できるというのが。

方法的行動主義は心理学がどのように研究されるべきかについての主張である。この点がまず論理的行動主義と異なる。論理的行動主義は、信念や欲求が行動を引き起こす心的状態であることを否定するが、方法的行動主義は心理主義的な用語が内的状態を指示することを認める。しかし、それらが内的状態ゆえに信念や欲求について心理学的に語ることは否定する。方法的行動主義の否定的なテーゼは、心理学は信念、欲求による説明を避けるべきであるというものである。この主張はスキナーに代表されるが、心理主義的な説明に対する彼の反対の理由の一つは、信念や欲求は観察できないというものである。二つ目の理由は説明が余りに容易過ぎるというもので、それらを仮定することで観察結果に合う信念や欲求の話をつくり上げることができてしまうということである。ここで以下の区別は重要である。

(心理主義のテーゼ) 個人の行動はその人がもつ信念や欲求によって引き起こされる。

(特定の信念・欲求仮説) 特定の信念や欲求は行動のデータに対してテスト可能である。

心理主義のテーゼはテスト可能ではないとしても、特定の信念や欲求をそれぞれの行動に帰することはできる。行動は有機体の内的な状態を記述しなくても説明できるというのが方法的行動主義の肯定的な主張である。方法的行動主義は、(1) 対象のいた過去の環境、過去に行なった行動、(2) 現在の環境、この二つから対象の現在の状態を説明できると考える。スキナーの条件刺激の理論はどのようにそのような説明が可能かを示してくれる。

ネズミはベルが鳴るとえさを取る。ベルが鳴る。それゆえ、ネズミはえさを取る。

方法的行動主義の難問は、新しい行動をどのように説明するかである。実際、この点がチヨムスキーの批判であった。過去の状態と行動の間の相関関係は現在でもそのまま適用で

きるのか。過去の相関関係によって条件づけられた行動を説明することは行動主義者には完全に可能であるが、信念や欲求によって行動する場合にはそうではない。つまり、方法的行動主義の環境決定論は誤っている。方法的行動主義は私たちが環境を自主的に変えていくことをうまく説明できない。

20 唯物論あるいは物理主義（と付随性）

中年以上の多くの人がかつて唯物論はマルクス、観念論はヘーゲルと習い、記憶しているだろう。そして、ヘーゲルの観念の弁証法を物質世界に適用してマルクスは唯物弁証法をつくり、歴史の必然性を説いたということも記憶に残っているだろう。このような理解の基本にある唯物論、あるいは物理主義とはどのようなものか。それがここでの問題である。

物理主義とはすべての自然現象は物理的な現象であるという主張である。最近では自然主義という言葉も使われる。存在する性質、状態、対象、出来事は自然的なものである。あるものが自然的とは、それが自然科学の基本的な理論によって理解できることである。このような自然理解は決して今世紀特有のものではなく、昔から存在していた。例えば、デモクリトスの原子論は古代の唯物論の典型である。あるいはルクレティオスでもよい。スピノザがアリストテレスの目的因やプラトンの考えを否定し、道徳の自然性や相対性を主張したのも自然主義の例である。ヒュームが実体という概念を否定し、自由、自己、因果性等を心理学的に扱うのも自然主義の典型例である。このような特徴は「自然化(naturalizing)」という用語に端的に現れている。これらをもとに自然主義の定義を考えれば、次のようになるだろう。私たちがもつ最善の科学理論が存在するもののタイプへの最善の指針となる。状態、性質、対象、出来事は科学理論によって実在的だとみなされる場合に限って実在的である。自然主義の主張は科学理論が世界の最善の像を与えるというものである。多くの自然主義者は科学理論が実在する世界についての理論であると考えている。科学は実在についての科学であるという実在論の主張は自然主義と重なっている。

自然主義について哲学的に議論する際に気をつけなければならない点は自然主義の内容が科学の進展と共に変化する点である。科学理論が与える自然像は科学理論が変化すればそれに応じて変化する。自然主義の変化は科学の変化に呼応している。科学の内容や形式、そして方法が変化するに応じて自然主義の内容も変化する。このテーゼのもとに自然主義も歴史的に変化することを受け入れるならば、自然主義に関する無用の論争の大半は避けることができる。自然主義の範囲や限界に関して語ることは科学の変化を抜きにしては意味がない。また、科学的でないものも実は科学の変化と同じように変化している。科学的でないものの内容の変化に応じて自然主義の内容は変化する。したがって、自然主義の意味は時代や理論変化に応じて変化し、定まらない。科学と非科学の境界が曖昧であるように、自然主義と非自然主義の境界も曖昧である。

自然主義への批判も多い。科学嫌いな人が私たちの心までも科学的に理解すべきだと言われたら、その人が嫌悪感をもつのは明らかであろう。このような科学嫌いではないにして

も、心の能力、人間の行為の理由、倫理や道徳を考慮したとき、かなりの人はそれらが自然化できるものではない、あるいは自然化されるべきではないと思う。あるいは、科学が実在する世界の記述や説明を与えてくれるのではなく、歴史や文化に相対的なものであり、私たちの世界観に依存したものであるから、心をそのような世界観のもとに考えることは単なる一つの立場に過ぎないと考える哲学者も多い。ほとんどの自然主義批判は次の自然（物理）主義の主張に対しての批判である。

（説明的な十分さのテーゼ）

心的状態の説明は物理的状態の説明によってなされる。

このテーゼから心的状態は物理的状態であるという結論が導き出される。例えば、痛みは最終的に物理学によって既述・説明される。その意味で痛みは物理的である。例えば、ローティは伝統的な哲学が消え、その役割の一部は科学が引き継ぐと考えるが、経験的な結果が哲学の問題にいつも答えてくれるわけではなく、科学以外のものが場合によっては適切に答えてくれると主張する。また、パトナムは科学が真理の唯一の供給者であることを否定し、科学的な合理性は相対的なものに過ぎないと考える。ネーゲルは科学理論の説明的な十分さを否定する。物理学が十分な説明を与える点に関して、例えば、マックギンは心的なものは物理的なものであると考えるが、私たちには心的なものと物理的なものとの間の関係を理解することができないと主張する。これらの哲学的な批判の他に、心を自然主義的に理解することへの宗教的、倫理的な批判も多く存在する。

このような批判に対して、物理主義や自然主義を柔軟に理解する術はないのだろうか。以下にそのような一例を考えてみよう。生物が環境にどの程度適応しているかを示す適応度という性質や私たちの心の様々な性質は物理的なものではないように見える。しかし、一方で適応度や心的な性質や状態は物理的なもの（生物や環境、脳）がないならば存在しないように思われる。物理的なものを必要とするが、それ自体は物理的とは呼べないようなものは私たちの周りに溢れている。その適例が情報である。とりわけ、心の状態や性質は情報とも関係して、そのような代表例である。では、すべての対象は物理的なものであるという物理主義は情報、そして適応度や心の性質に関して成立しないのであろうか。

ここで付随性という考えを紹介しよう。対象の性質の集合 Q がその対象の別の性質の集合 P を決定し、逆は成立しないとき、 P は Q に付随する。つまり、 P が Q に付随すれば、 P と Q の間には 1 対多の関係がある。例えば、同じ部屋の多くの温度計が室温 20 度を示しているとき、それら温度計は同じ情報をもっている。しかし、ものとしての温度計はみな異なっている。室温 20 度は異なる温度計によって表示されているにもかかわらず、同じ情報を表している。これと同じように、適応度はその生物と環境の物理的な性質に付随し、心のある性質は脳の物理的な性質に付随することになる。これをさらに一般化すれば、物理学以外の科学で扱われるすべての性質は物理的な性質に付随するということになる。この付

随性を使って物理主義、さらには唯物論の主張を明確にしてみよう。

「すべての対象は物理的な対象である」という主張は何を意味しているのか。ある対象が物理的であるとはその対象のある性質が物理的というのではない。その対象が魂や生命力をもっている、その対象は他に質量や温度をもつことができる。また、その対象のすべての性質が物理的な性質というでもない。実際、適応度や心の性質は物理的ではない。このような状況を損なうことなく、物理主義と物理的でない性質を統合的に扱う際に付随性は役に立つ。「ある対象は物理的である」を付随性を使って言い直すと、「ある対象が物理的であるとは、その対象のすべての性質について、それが物理的でなければ、その対象の物理的な性質に付随する」ということになる。

付随する性質を研究するのは科学であり、したがって、科学は物理的でない性質を研究できることになる。このような拡張は付随する性質はどのような性質かについての哲学的議論を巻き起こすことになる。実際、付随する性質は因果的な原因にはなれない。因果的な効力をもたない性質は科学では何の役割ももっていないのだろうか。そのようなことはない。因果的に無力であっても、説明に関しては効力をもっている。情報、適応度、心的性質といった付随的な性質は科学的な説明では物理的な性質と同程度の効力をもっている。科学的な説明は因果的な過程の記述以外のものも含んでいる。

21 意識と無意識

フロイトに神秘的な魅力を感じる人はどこにその魅力を感じるのだろうか。彼は生理学者、医者、心理学者であり、なにより精神分析の産みの親である。彼は心が複雑なエネルギーシステムであり、その構造的な探求こそが心理学の役割であると考えた。彼はそのために無意識、小児の性、抑圧の概念等を使って心の構造を説明しようとした。その斬新な分析は多くの分野に重要な意義をもっており、特に心の科学を創り出そうとした点は特筆されるが、その科学が真の科学かどうかに関しては現在でも多くの批判、論争が渦巻いている。この批判、論争の理由はどこにあるのか。

フロイトの無意識の理論は極めて決定論的である。心的な領域に最初に決定論的な原理を使い、心的な過程や状態を原因にして人間の行動を説明しようとした。原因を使っての結果の説明、つまり、因果的説明である。無意識的な心的状態があるという仮説はこの決定論の直接の帰結である。人間の異常な心理状態を引き起こすものが意識的な心理レベルでは見出すことができないのであるなら、意識的なレベルでない心理的なレベルにある状態が原因になっていなければならない。無意識は何時でも意識から生じるものではなく、精神分析を通じてでなければ意識の前面に登場することのできないものである。このような無意識の心理状態の仮定は、心が意識と同じではありえないことを帰結する。デカルトが心 = 意識と考えたのとは違って、フロイトにとって意識は心の一部分に過ぎない。

フロイトの本能や衝動の説明は無意識の存在の仮定に深く結びついている。彼は本能を二つの種類に分け、エロス（生本能）とタナトス（死本能）とした。したがって、彼が人間

の行為はすべて性的な衝動によっていると考えていたというのは誤りである。タナトスはすべての性的衝動を破壊するものである。とはいえ、彼の本能と衝動の理論は、人間が身体の快樂を獲得し、それを高めるという欲求によって突き動かされているというものであった。

フロイトの説には多くの用語が常識的な意味で使われている。欲求、衝動、無意識といった用語はどのような心的状態を指しているのだろうか。私たちは自分が日常生活でわかったつもりになり、しかも巧みに使っているこれらの心的用語をいざ正確にしてみると言われるとうろたえてしまう。そこで以下には意識のなかでももっとも単純に見える感覚的な意識について、フロイトの理論からみれば全く単純に見える事柄を考えてみよう。既述の物理主義の立場からは感覚的な性質、色や匂いは物理的な性質でないゆえにどのように説明されるのが問題になっていたことを思い出してほしい。

ジャクソンは物理主義に反対する推論として次のような架空の話をつくった。私たちが色と呼ぶものは自然のものに対応していない。にもかかわらず、私たちはバラが赤いと思うし、それを経験できている。一方、科学は可視スペクトルが連続で、私たちが経験するような色の違いが自然にそのまま組み込まれていないことを示している。さらに、私たちが違う生物はスペクトルを違ったように区分することも知っている。すると、ある対象の色は何かと問われた場合、その答えに対応する科学的な事実はないことになる。しかし、色は幻覚や夢とは違って「主観的」ではない。人が異なれば異なるといったものでもない。では、色は一体何なのかということになる。あるスペクトルの範囲が生存と繁殖のために視覚システムによって識別される必要があったので、色を見るという能力が自然選択によって進化してきた。ある色と結びついた内部のシグナルが経験の質的な内容である。これがデネットの答えである。感覚質はシステムの中で機能的な役割を演じる内部状態に還元されている。デネットだけでなく、このように考えるのが正統的機能主義者である。しかし、感覚質は内部状態に還元できないと考える者もいる。機能的でない、感覚質の存在を擁護する推論の代表がジャクソンのメアリーに関する推論である。

ジャクソンは次のような思考実験を考案した。メアリーは完全に白黒の世界で生まれ、育てられ、有能な神経科学者に成長した。彼女の専門は視覚の神経生理学であった。色の知覚に関するあらゆる物理的な事実を彼女は学んだ。物理世界の見た目の完全さが示しているのは、物理的な出来事を引き起こすが、それ自身は物理法則に従わない、独立の心の領域など存在しないことである。これが正しいければ、世界に関する物理的な知識は世界に関する知識のすべてであることになる。物理主義の主張によれば、色の知覚に関するあらゆる知識はメアリーのもっている知識に含まれていることになる。そこで、ジャクソンは次の問いを出す。メアリーが白黒の世界から開放されて普通のカラーの世界に身をさらした時、彼女は「色」がどのようなものが学ぶのではないか。もし、そのような学習によって新しい知識が得られとしたり、それは物理主義的な知識ではない。というのも、彼女は既に色に関する物理的知識をすべてもっているはずだからである。この思考実験を推論の形

に整理すると次のようになるだろう。

- (1) メアリーはモノクロの世界で物理的な世界についてのすべての知識を習得した。したがって、色の経験はないが、色についての物理的な知識はもっている。
- (2) 物理主義は、世界についての知識は物理的な知識であると主張する。
- (3) しかし、メアリーは色の感覚質を知らない。

したがって、物理主義では知ることができない知識が存在する。

この反物理主義の主張に対してどのように物理主義者は返答するのだろうか。以下に二つの返答を挙げてみよう。

- 1 カラーの世界を見ても、メアリーは新しい知識を獲得しない。
- 2 彼女が獲得するのは新しい知識ではなく、別のタイプの知識（こつや技能）に過ぎない。

これらはいずれもジャクソンの基本的な姿勢、つまり認識論的な問題として物理主義の欠点を暴くという姿勢をそのまま認め、反物理主義的な結論は避けようとするものである。まず最初のを考えよう。未来の神経科学は今以上に進歩し、色の質や主観的経験を明らかにしてくれるので、メアリーは初めて色を経験しても驚くことはないというものである。私たちは現在の物理主義的知識とその未来がどのくらい隔たっているかを知らない。これに対するジャクソンの返答は、この推論は物理主義が誤りであるという推論ではなく、物理主義が誤りである可能性が高いということを示唆しているだけだというものである。

2 への反論をチャーチランドは次のようにまとめている。

- a メアリーは部屋を出る前に他人について知るために必要なすべての物理的なことを知る。
- b メアリーは部屋を出る前に他人について知るために必要なすべてのことを知っていない。
- c それゆえ、物理主義的でない他人についての知識が存在する。

a と b の「知る」を考えてみよう。物理的なものの知り方は文によって表現でき、真か偽かの判断がつくような知識である。それに対して、「知る」の異なる解釈がある。それは命題で表現できないような知識で真でも偽でもないような知識である。このような知識が b に含まれていれば、この推論全体は正しくないことになる。

22 人間の本性

アリストテレスによる人間の定義は理性的動物というものだった。人間とは何かをめぐっては多くの分野で多くの人によって考えられてきた。ルネッサンスに登場するヒューマニズムは人間は何かの再考であった。また、デカルトの延長と思考の総合としての人間観も

その後の社会に大きな影響を与えてきた。しかし、デカルトにとって人間と他の生物との関係は歴然としていた。人間以外の生物は単なる機械に過ぎなかった。このような人間中心主義を打ち砕くことになったのがダーウィンの進化論の登場である。人間も他の生物と同じように進化して来たという主張が科学的に正しい仮説かどうか以上に、宗教的、道徳的な問題を引き起こすことになった。そして、それは現在も継続している。

では、進化論による科学的な人間の理解はどのようなものなのか。人間の生物学的な研究は人間についての知識を格段に増やしてくれた。それでも私たちは自然科学の対象であると同時に、そのような自然科学を研究・応用する主人公であることを強く意識している。だから、人間の生理や形態に関する知識は尊重するが、知識、歴史や文化を学び、それを使って生活する場面については生物学は関係ないと思ってしまう。そして、このような人間の社会的行為に関する研究は従来社会科学や人文学の問題とされてきた。身体としての人間は生物学的な対象だが、精神としての人間はそうではないというデカルト的な二元論が現在でも生き残っている。

しかし、このような状況は変わり出している。精神も自然科学の対象になり出している。その一例が認知科学である。ここでは人間の本性に関係深く、かつ倫理の基本概念の一つと見なされてきた利他性に関して、その科学的な追究を考えてみよう。人間は利他的な愛をもち、他人にそれを与える点で他の動物とは全く異なると考えられてきた。このような人間の本性は生物学的な根拠から説明できるのか、それとも人間特有の文化や知識からしか説明できないのか。この問いに対する解答は倫理や道徳の本性を左右する。

動物は利己的であるように見える。したがって、この一般的な本性である利己性を使って、利他的にみえる行動を説明できないかというのが最初のアイデアであった。ハミルトンらの数学的モデルの研究から、従来の適応度（生存と生殖に関する有利さ）概念を拡張して、その個体だけでなく近親者も含めた包括適応度によって、利他的な行動を説明するモデルがつけられた。だが、この結果は利他的にみえる行動が実は利己的な遺伝子によってコントロールされているという主張を確認するものでしかない。

では、利他的な本性はみかけのもの、利己性の変形に過ぎないのだろうか。ここで既に説明した総和の誤謬を思い出してみよう。その際の表を応用して、利他的な行動のほうが利己的な行動より適応度が高い場合があり得ることを示してみよう。そして、このことが利他的な行動が利己的な行動より適応度が低いという一般的にはもっともと思われる仮定の下で、グループとグループの集まりが存在するならば、利他的な行動のほうが利己的な行動よりは適応度が高くなる場合があり、したがって、利他性が集団内に保持され、選択的に有利であることが可能であることになる。これが利他性の存在についての基本的な考え方である。

グループ 1	グループ 2	総計
1S; W = 4	99S; W = 2	100S; W = 2.02

99A; $w = 3$ 1A; $w = 1$ 100A; $w = 2.98$

(この表の S は利己主義者、 A は利他主義者である。 W 、 w はそれぞれ適応度を表している。グループ 1 には利己主義者が 1、利他主義者が 99 いる。グループ 2 は利己主義者が 99 で、利他主義者が 1 であり、その中間のグループも簡単に想像できる。 $W = 2.02$ は $(1 \times 4 + 99 \times 2) / 100$ である。入学試験の表と異なることを表しているように見えるが、上の W 、 w を 1 以下の重みづけにして考えるなら、最初の例とそれほど異なる印象を与えない例となる。1 以下の重みが合格率に対応すると考えればよい。)

このような結果を別の仕方でもとめてみると、次のように言うこともできる。下の推論が与えられた場合、それは正しいだろうか。

どのような部分集団においても、利己主義は利他主義よりも適応度が高い。

適応度の高くない性質はその頻度が低下する。

それゆえ、利他主義はその頻度が低下する。

明らかにこの推論は誤っている。総計の誤謬を犯しているのである。というのも、上の表の総計の数値 ($W = 2.02 < w = 2.98$) がこの推論の結論の反例になっているからである。どのような部分集団においても頻度が低下する利他主義者は、全体ではその頻度が高くなることもあり得るのである。

上の説明は単なる例に過ぎず、もっと議論を慎重に進めなければならない。しかし、その核心はグループという集団概念を導入することによって、利他主義的な性質が集団の中に十分存続できるモデルをつくることのできる点にある。ここから階層的な選択のレベルを考え、群選択を認める考えが出てくる。そのもとでは、利己主義の変形ではない仕方で利他主義の存在を示すことができる。

倫理や道徳という人間特有のものに上のような生物学的な説明を与えることに違和感や危険を抱く人は少なくないだろう。人間の本性が生物学的に説明できると思う人も決して多数派ではなからう。しかし、このような説明の仕方が唯物論や物理主義の真髄にある点を忘れてはならない。人間を科学的に理解する一つが上のような理解の仕方なのである。

23 伝統と偏見

理論的な知識を扱う物理学、道具的な知識を扱う生物学という分け方は現在でも通用しているように思われる。また、心は自然科学が扱うことのない聖域という考えも珍しくはない。心は科学ではわからないと考えている人は少なくない。このような生命や心に関する私たちの常識的な態度は正しいのだろうか。そもそもこれらの常識的な態度はどのようにつくられてきたのか。

中世以来自然の研究は自然哲学 (Natural Philosophy) と自然誌あるいは博物学 (Natural

History) に二分されてきた。それが次第に物理学と生物学に変貌する。自然哲学の代表はニュートン力学である。自然誌の代表はさまざまな博物誌、図鑑である。力学と図鑑は二つの分野の違いを象徴的に表している。自然哲学の目標は自然の基本的からくり(自然法則)を明らかにし、それを使って自然現象を説明することにある。一方の博物誌は、野外でつぶさに観察、記載、分類した対象をまとめたものである。二つの違いは説明と記録の違いである。この違いが物理的なものと生命的なものという区分の背後にある。このような二分化のもとで、生命現象は自然誌の研究対象として扱われた。生命現象は、したがって、説明されるのではなく記録される対象であった。物理学と生物学が異なるという場合、このような伝統的な区別が現在でも暗に尾を引いているようである。あるいは、生命現象を物理現象に還元するという場合、説明が可能な物理現象を基礎にして、説明できない生命現象を物理的に説明するという同じような伝統が生き残っている。また、生氣論の多くも物理的なものと生命的なものとの違いを自然哲学と自然誌の違いから導き出している。いずれにしろ、私たちの多くは物理的・生物的という二つの形容詞で自然現象を区別することに未だにそれほどの不自然さを感じていない。その意味で中世以来の伝統は現在も生き残っている。

心の概念に関する有力な伝統の一つはデカルトの伝統である。心身二元論が彼の主張であるが、生命の場合と同じようにその主張は過去の歴史ではなく、私たちの常識として現在でも生き残っている。デカルトは心が思考という性質を、ものが延長という性質をもち、二つの性質は根本的に異なると考えた。これが心身の二元論である。また、心とものは私たち人間の場合、脳において相互作用していると考えた。そして、自己意識としての心(他人の心はわからないが、自分の心はわかる)が心を考える際の視点であると見なした。心の振舞いを単なる生命現象と同じと考える者は少数派である。心が病めば身体に影響を与え、身体の不調は心を暗くすると考えられている。誰もこれら主張を見当違いのものと受け取らない。誰も自分のことは自分が一番よくわかっている。この常識が正にデカルトの伝統であり、それは現在も生き続けている。

このような伝統は生命や心を考える際にどのような影響を与えたか。生物学は物理学から分離され、そして、心からも分離されるという被害を蒙ることになった。その結果、生物学は物理学と心の研究の両方から締め出されてしまう。そのような生物学の採った途は細胞学という経験的な研究であった。19世紀後半の先端科学は顕微鏡の駆使によって支えられていた。その顕微鏡とともに発展したのが細胞学、そしてそれを基礎にした発生学であった。これは物理学そのものではないが、その研究の仕方は実験や観察の重視から物理学的であった。そこでは物理的な力ではない生命力と生命現象の実際の研究が目的機械論(teleomechanism, T. Lenoir, *The Strategy of Life*, The Univ. of Chicago P., 1982)という危うい関係で結ばれていた。

この伝統 = 偏見に反対するという意味で革命的なのがダーウィンの進化論である。進化論は二つの伝統 = 偏見から生物学研究を解放する役割を果たした。生命現象は力学的な現象

と同じ立場、態度で扱うことができ、心は行動や生態の分析から解明される、というのがダーウィンの主張であった。この主張がダーウィンの考えをニュートン革命に匹敵するものになっている。ダーウィンの『種の起源』は生物進化の事実とその進化の分岐的な仕組みを述べたものである。進化の分岐的な仕組みは自然選択によって説明される。この説明方式に伝統 = 偏見への反対が凝結されている。それをまとめるなら、次のような項目になるだろう。

- (a) ニュートン力学をモデルにする（進化現象を力とその作用の結果として説明する）
- (b) 生物集団を統計的に扱う（生物集団の変化を確率・統計的に説明する）
- (c) 人間の行動や生態を説明する（行動や倫理や道德の規範を進化論的に説明する）

ニュートン力学をモデルに進化を説明するとは二分化への反対である。生物集団に力が働き、その結果として進化が生じるという考え方は力学的な考えと基本的に同じである。この力がダーウィンの場合は自然選択であった。生物個体の自発的な内的変化ではなく、外からの力による変化が進化であるという考えは生氣論とは全く異なった発想、見方である。(b)の確率・統計的な説明はダーウィン以後の総合説において主に使われる説明である。総合説の創始者の一人であるフィッシャーはメンデル遺伝学と自然選択説の総合の際に、それが統計力学に倣ったものであることを明白に述べている。(R. Fisher, *The Genetical Theory of Natural Selection*, Dover, 2nd ed., 1957) 確率・統計的な説明は現在ではあらゆる分野で頻繁に使われている。説明レベルでの二分化への反対は総合説の特徴である。一方、(c)はデカルトの伝統への反対と考えることができる。心と身体あるいは物体の分離に反対して、生物の行動を進化論の中で考えるという視点もダーウィンに起原をもつ。『人間の由来』はそのような最初の著作であり、行動の進化という観点が採られ、心が係わる行動を自然科学的に扱う出発点となっている。そして、行動の性質としての倫理や道德も生物学的な適応として研究されることになる。これは紛れもなく倫理や道德の自然主義的な扱いである。

今世紀後半の生命科学の進展は目覚ましい。その発端はワトソンとクリックによる DNA の二重らせん構造の仮説にある。それをきっかけに遺伝学は分子レベルの遺伝学へと急展開する。分子レベルでの生物学は生命現象を化学、そして物理学へ還元する動きを助長する。しかし、分子レベルには生命現象の法則は見当たらない。このことは生命現象には独自のものはなく、物理的な法則だけで生命現象も説明できることを示しているように見えた。しかし、遺伝情報は物理的なものに完全には還元できない。情報概念そのものが物理的でなく、その情報概念なしに遺伝現象を説明できないことは、生命現象が物理現象に還元できないことの証拠になっている。では、この還元不可能性は生氣論の復活につながるのか。物理現象には還元できないが、それでも自然現象であるような生命現象を研究できるとするものが進化論の基本枠組である。

では、生命現象の基本枠組としての進化論の説明の特徴はどこにあるのか。総合説として

1930年代から整理されてきた進化論は自然選択説と集団遺伝学の総合であり、数学的な構造をもった理論として形式化された。これは自然科学の理論としての条件を満たし、仮説演繹体系として予測まではできないにしても記録のための単なるマニュアルではないことを証明している。さらに、遺伝学の観察・実験は飛躍的に向上し、分子レベルの知見を蓄積してきた。これら二つの点は、特に科学哲学において第三の偏見を生み出すことになる。それはある程度の経験的事実と仮説演繹体系の整備が進化論を物理学の理論と同じような理論であると見なすという偏見である。物理学理論と基本的に同じであるなら、物理学の手法と同じ手法を用いるべきであるということになる。これは背後に科学理論の模範が存在し、科学理論は皆それに従うべきであるという偏見から出てくるものである。

しかし、進化論の説明は起源と適応による説明に特徴がある。それらはいずれも物理学にはない説明形式である。起源と適応を使った説明の具体例を挙げれば、そのような説明が意外に普及していることがわかるだろう。

ゲーム理論等を使った適応による説明

人工生命のシミュレーション

生命倫理や環境倫理

心理と行動 - 社会生物学、進化心理学

これらはいずれも仮説演繹体系としての進化論の枠組だけではその説明が展開できないものである。従来、哲学で機能的な説明、目的論的な説明と呼ばれてきた説明が上述の具体例の正体である。進化論は物理学での因果的な説明とこの機能的な説明を組み合わせる進化現象を説明している。モデルによる研究だけでなく、歴史的な出来事の再構成、進化のシミュレーション、シナリオを含み、人為選択の計画までも入ってきている。しかし、それゆえに進化論は多くの未決の問題を抱え込むことになる。それらはいずれも従来の科学の常識から逸脱する要素を含んでいるからである。選択の本質、浮動、決定論と非決定論といった理論上の問題、生命現象、集団現象、規範の適応的な説明といった応用上の問題は物理学を基本にした哲学的な分析の範囲を超えている。

私たちは生命現象をどのように扱ってきたかを見てきた。二つの伝統 = 偏見は進化論によって原理上は解決できる。しかし、それはあくまで原理上のことであって、実際にどのように具体化するかは現在まだ十分に解決されているとは言えない。進化論が物理学にはない説明方式を考案し、そのことによって物理学とは異なる自然科学の可能性を示したと同じように、心の科学も進化論にはない新しい説明方式をもっているのかもしれない。その可能性は強い。しかし、私たちはそれが何かをまだ確定できない。第三の偏見の克服では足りず、さらに第四の偏見を克服しなければならないのかもしれない。なかでも考えなければならないのは数学や論理学の知識、そして意識である。これらの心の所産は自然主義的に特徴づけることができるのか、それとも独自の存在なのか。心の振舞いの一部は確か

に進化論の枠組の中で扱うことができる。問題は、数学や論理学の知識、さらには言語的な知識がどのようなものかということである。主観的意識は生物的であると考えられることができるが、数学的知識が生物的というのは考えにくい。

進化論は確かに二つの偏見を取り除いた。そして、第三の偏見も克服しようとしている。だが、このことはすぐに心的現象が生命現象に還元できることを保証するものではない。どのようにこの間隙を埋めていけばよいのか。この間隙は例えば、近年の進化心理学や人工生命研究によって埋められ出している。

24 正常と異常（アリストテレスの Normal State Model とダーウィンの Variation Model）

人間を含む生命現象を科学的に扱う以前のことを考えてみよう。例として性と生殖についての神話や伝説を考えてみよう。ギリシャ人は最初の間が半分男、半分女で、それが神によって二つの性に分割されたと信じていた。両性をもった人間は大変強く、それがゼウスを脅かしたので、ゼウスは人間を二つに分割したことになる。旧約聖書ではイブはアダムの肋骨から造られたことになっているが、これはヘブライ語の文献からの誤訳で、最初のアダムは両性を含むもので、片側が男に、他方が女にあっている。原アダムがアダムとイブに分離したのである。これが一夫一婦制の教会による説明のもととなる。結婚は分離した二つの部分を再び結合するものと考えられた。

一方、性や生殖に関する科学的な研究も昔からある。アリストテレスは生殖を他の動物との比較で研究した最初の人であり、彼の考えによれば、精子が胎児を生み出す種をもち、月経の血がそれを成長させる土壌となる。アリストテレスによれば、胎児を形成し、成長させるのは靈魂であり、男はその形相因、機動因であり、女はその質料因である。アリストテレスの考えはトマス・アキナスに引き継がれ、カソリック神学の基盤となる。人間の誕生はまず生物的、次に動物的、最後に人の順で生じる。この考えから初期の教会では嬰兒殺しより、生物的段階の墮胎に対する倫理的な反対は少なかった。また、精子は生命の本質であるので、それを浪費することは重大な罪であった。イスラム教では女の精子も生殖において血や肉を造る素になると考えられ、男の精子だけでは生命はできないので、男の精子の浪費は問題ではないとされた。（こうして、マスターベーションについてキリスト教とイスラム教の二つの異なる態度が出てくることになった。）

科学的な追究と神話や常識との混合が以上のことからある程度はわかったと思われる。このような例は性や生殖以外にもいくらかでも枚挙できる。

性選択はダーウィンによって考えられた性と生殖の重要な仕組みである。この研究は動物行動学という分野に発展することになるが、やはり科学的追究と常識が衝突することになる。性選択とは、形式的には次のように定義される。生殖に限って同種・同性の他の個体に対してある個体もつ有利性が性選択である。

オスとメスの究極的な違いはそれぞれがつくる配偶子の大きさの違いである。小さいものをたくさんつくるか、大きな物を僅かだけつくるか。この起源は別にしても、メスは生殖

と生育にオスより大きな投資をする。この投資のためにメスはオスより生殖には慎重で、それがオス選びに結びつく。選ばれる対象であるオスは上の定義に従って同性の間での競争に打ち勝たなければならない。つまり、オスの間により強い性選択が働いている。オスの間だけの選択は簡単で、オス同士の競争に打ち勝つ限り、形質は選択される。これにメスが入ると複雑になる。オスが配偶子しか提供せず、すべてのメスが交配するような種では、メスの適応度は配偶者の選択によっては影響を受けない。では、オスの誇張された形質に対するメスの好みはどのように進化し、それがオスの形質の進化にどのような影響を及ぼすのか。このような問題が議論され、現在ではメスの選択(choice)がオスの競争より性選択を考える上で重要とみなされている。人間の場合、男優位の社会と性選択における男の役割は伝統的な思想や習慣と生物学的な研究の間で揺れ動いている。

性やジェンダーの議論に限らず、人間の心理や行動を語る場合、「正常」、「異常」という言葉が多用される。「異常性愛」という言葉は正常な性愛があって、そこからのズレとして考えられていないだろうか。「正常」や「異常」には事の「善し悪し」に似た、価値判断が含まれていないだろうか。ここには規範にかなっていれば正常、規範にもとれば異常という判断が含まれているようにみえる。含まれていれば、「正常」や「異常」を含む考察や研究は価値判断を含むものになってしまう。類似の例は「健康」と「病気」である。このような理解の仕方について考えてみよう。

既に述べたように、性や生殖に関するアリストテレスの影響の強さには根深いものがある。そのような一つが彼の正常状態モデルである。どのようなものにもそれ本来の存在の仕方と場所があり、その本来の姿を正しく把握することが本質の理解につながるというのがこのモデルの趣旨である。このモデルの対象は物理的なものだけではなく生物も含んでいる。人間の正常な姿が人間の本質を具体化したものであり、その本質からズレたものが正常でないものである。それら異常なものはたとえ出現しても選択され、支配的になることはない。このモデルは天体の構造や生命現象を大変うまく捉えている。規範、模範になる姿があって、それに外れるものはたとえ存在しても、あくまで例外に過ぎない。現在でも理科室にある動植物の標本はその種を確認するための模範となる標本である。

アリストテレスの正常モデルと根本的に異なるのがダーウィンの変異モデルである。彼は生物集団の中には常に変異が存在し、それが個体差として選択の篩にかけられ、生存と生殖に関して有利なものがその集団の中で多数を占めるようになっていくという、いわゆる自然選択理論によって生物の進化を説明した。この説明の出発点は変異の存在である。この変異、個体差には正常も異常もない。集団の個体に模範となる正常なものも異常なものもない。あるのは個体間の差だけである。そして、この差が進化の原動力になっている。したがって、正常、異常とはある時点の集団の多数派、少数派に過ぎなく、本質的なものではない。

このように見えてくるとアリストテレスとダーウィンの違いは歴然としている。ところで、私たちが生命現象や社会現象を考える際、いずれのモデルで考えているだろうか。多分、

生命現象に関してその原理的な部分ではダーウィン風に、形態や私たち自身の身体的特徴や行動に関してはアリストテレス風に考えているだろう。社会現象や文化現象に関しては近年意外にダーウィン風で、社会や文化の多元論が議論されている。異常な行動は大抵の場合悪い、してはならない行動とさえ考えられている。これだけ述べただけでも、そのような分析が価値判断を含むかどうか、価値判断からは中立かといったステレオタイプの問題ではないことが明らかだろう。

もしアリストテレスのモデルが(かつてそう考えられていたように)正しい科学的なモデルであれば、「正常」、「異常」は優れて科学的な概念であり、それら概念を正しく使った判断は正しい科学的な判断である。一方、ダーウィンのモデルが正しい科学的なモデルであれば、「正常」、「異常」は科学的に誤まった概念であり、それら概念を使った判断は科学的に誤まった判断ということになる。この表現のどこにも価値判断など入っていない。問題は「正常」、「異常」を最初から価値判断が入っていると思いつくことである。確かに、より複雑な人間の行動に関しては歴史、文化、社会の科学的でない基準や約定が関与しており、そこから価値判断を含んだ「正常」や「異常」が生まれ、伝統をつくってきた。しかし、性や生殖の場合のように、それら基準や約定の根拠も科学的な知見に依存している。その科学的な知見が正しいかどうかを判定するのはいずれのモデルを選ぶかという問題であり、それは価値判断とは独立した事柄である。

25 ポストモダン

哲学の歴史にはギリシャ以来合理性を信頼し、そのもとで自然や人間を追求する立場とこの合理性に異議を申し立てる立場が主導権を争ってきた。主流派は合理性信奉の立場であり、反合理主義者は反主流であった。ここにはヨーロッパ哲学の大きな二つの流れとその特徴が見て取れる。一つは合理的追求を根幹におくものであり、私たちの心の創造的な点より、その心によって捉えられる対象とその追求に重点を置く。他はその逆で心の創造的な能力を信じる。簡単な例として心の認知科学的な捉え方を考えてみよう。それは情報処理システムとして心をつまようとする。そこでは私たちの心は外部から入ってくる情報を処理する装置とみなされている。一方、芸術活動は心の積極的な創造によって何かを表現したり、創り出したりするもので、そこでは心の表出(英語の *expression* と *impression* の違いを思い出そう)が重要であるとみなされている。実存主義は反主流の思想であり、その実存主義が退潮をを見せる中で台頭してきた構造主義は主流派に属する。実存主義は戦禍のヨーロッパの特異な状況の中で生まれた思想で、大学の研究室での研究より、与えられた状況をどのように生き抜くかに重点が置かれていた。このような哲学は今まで述べてきた合理的な思索の展開というより、ロマン主義的な哲学に近いものである。構造主義は当然ながら前者の立場で世界を考える。

20 世紀中葉には既に人間の存在に関する構造的な理論が多くあった。言語研究ではソシユールの構造言語学が意味は個々の語の分析にではなく言語全体の構造の中に見出される

と示唆していたし、マルクス経済学は人間存在の真の姿は経済構造の分析によって理解されると述べていた。精神分析学では無意識によって精神構造を記述しようとしていた。1960年代、構造主義の運動はマルクス、フロイト、ソシュールの考えを統合し、実存主義的な各人は自らをつくりだすものという主張に反対した。構造主義者にとっての人間は自らによってはコントロールできない社会的、心理的、言語的な構造によって形成されており、人間はそれら研究手法を駆使して明らかにされるものであった。

このような考えは人間研究に端を発するのではなく、基礎的研究に基づいていた。それが端的に表われているのがブルバキという個人名をもつフランスを中心とする数学者のグループによる数学の再構築と分類である。彼らはヒルベルトの形式主義と集合論を使って 20世紀数学の基礎付け主義的な再構築を行なった。数学は *Mathematics* という複数形で表現されるが、彼らはそれをしっかりした基盤の上に構造の違いから分類された各数学理論を段階的に積み上げる仕方で壮大な理論の建物を築き、一枚岩の単数形の数学家屋 *Mathematic* をつくろうとした。数学の各理論はそれぞれ異なる数学的な内容を扱うが、それら内容はその構造に従って分類される。数学が扱う基本構造は、代数構造、位相構造、そして順序構造である。より一般的な理論から特殊な理論へと階層的な仕方で配列され相互の論理的な関係が明らかにされていく。

この構造主義の後に登場するのがポスト構造主義であり、それと脱構築がポストモダンを生み出していく。

ポスト構造主義と脱構築はポストモダンの理論的表現と見ることができる。啓蒙時代以降の近代は世界を合理的、経験的、客観的な用語で描こうとしてきた。私たちが出した問題に対する解答を得るのに、明らかにされる真理が一つあると仮定されていた。ポストモダンの考えは、このような確信は偽物で理性が約束する確実なものではなく、理性自体特殊な歴史的形態に過ぎないというものである。ポストモダンの主体は真理、倫理、美、客観性に関して評価する合理的な基準をもっていない。

フーコーは構造主義者と呼ばれる。しかし、言語と社会は規則に支配されるシステムによって形成されるが、人間の条件を説明する一定の構造があることも、状況を客観的に研究することもできないと主張する点で他の構造主義者とは異なっていた。フーコーは知識を明らかにする言語行為を分析する。言説に関する行為をその真理によって分析するより、その歴史や発生によって分析しようとした。それが知識の考古学である。真理を主張することの歴史を示すことが彼の分析の目的であった。後期の研究ではニーチェの系譜学的方法やマルクスのイデオロギーの分析を借り、知識の展開が（政治的な）権力とどのように結びついてきたかを示そうとした。フーコーはマルクスと違い、言説や社会を分析する客観的な視点の存在を否定する。『性の歴史』で、彼は医学や精神分析学の勃興が性の言説を深淵で、本能的、神秘的なものにつくりあげ、この言説が主要な説明として採用され、日常生活に浸透することになったと論じた。その結果、個人の性についての経験がそれを説明しようとする言説によって形成、支配されることになった。知識追求は存在する対象を単に明ら

かにするだけではなく、対象自体をつくりだしてしまうと彼は論じた。

デリダにとって言語やテキストは世界の自然な反映ではない。ハイデッガーにしたがって、デリダは言語が私たちをつくり、テキストが実在として理解したものを仕上げると考える。デリダはヨーロッパの思想の歴史を二項対立と見る。これらの対立は階層的に定義され、対立の二番目の項は最初の項の崩壊であり、両者は同等の対立項ではないと考える。善悪、心身、男女はそのような例である。デリダは、テキストは言語に内在する階層を意識して再解釈されなければならないと考えた。だが、そのような解釈が最終の真理に到達することはなく、様々な解釈はテキストのもつ差異によるものであると考えた。

脱構築の結果の一つは、テキスト分析の確実性はないというものである。多くの競合する解釈があり、そのいずれが妥当性をもつかを示す、それ自体解釈を免れない方法はない。否定できない真理に基づく哲学的理解ではなく、脱構築論者は解釈の多様性に合理性を沈めるのである。

ローティはサルトル、ハイデッガーから大きな影響を受け、ヨーロッパ哲学の近代以後の関心をアングロアメリカの哲学での言語ゲームを使って明らかにしようとした。ローティはデカルトが精神についての哲学をつくり、それが世界を理解する確固とした基礎を与えようとするものだったと述べる。カントは私たちが普遍的で、無時間的なカテゴリーを通じて世界を解釈すると考えた。鏡の役割をする非自然的な心とそれに映される自然世界が明瞭に区別された。もし客観的な哲学的観点がないのであれば、哲学は客観的知識の基礎を与えるのではなく、使い古された語彙や態度から私たちを開放するのを助けるに過ぎない。

ここまで読んできて、この節の内容はそれまでの節と異なるタッチで書かれていることに気づかれたのではないか。この節の幾つかの思想の要約の中に、これまで述べてきたものが幾つも成分として含まれているはずである。哲学的な推論を積み重ねた結果としての思想は、それを解説するのも、解説を理解するのもそれ程困難ではない。困難なのはそのような積み重ねを実践することである。積み重ねの実践に必要な材料と手法の要点だけを前節まで述べてきたつもりである。したがって、後は実践だけである。

行動主義と機能主義の対立から

心の理論は行動主義から機能主義へときっぱり転換したと言われてきた。そのためか行動主義は心のない理論とみなされ、投げ捨てられたままになっているように思われる。では、経験主義の一面を見事に表現した行動主義は経験主義的に見えない機能主義に対してどこが劣っているのか。この問いを考える際の要点は行動主義と機能主義のそれぞれの哲学にあるというより、Black Box型説明と Turing Machine 型の説明の違いにあるように思われる。これは思想的な違いではなく説明形式の違いである。

行動主義については既に説明した。機能主義はタイプ - トークンの区別を使うことによって説

明できる。個々の物理的な対象はトークンであり、性質はタイプである。概念とそれを表す具体的な対象がタイプとトークンである。ある物理的な対象は多くのタイプのトークンである。私のパソコン(トークン)はさまざまな性質(タイプ)をもっている。タイプはいくつかのトークンをもつ、あるいは何もトークンをもたない場合もある。人間という性質は多くの個人がそのトークンとなるが、ユニコーンはトークンをもっていない。さて、同一説は心的なトークン、心的なタイプ両方についての理論であるが、機能主義は同一説が心理的なタイプについて主張することを否定する。つまり、心的なトークンは物理的なトークンであるが、心的な性質は物理的な性質ではないと主張する。機能主義は心的性質が物理的なものに付随する(supervene)、したがって心的性質は複数の物理的なトークンによって実現される(multiply realizable)と考えるのがその理由である。心的な性質は物理的な性質ではない。しかし、心的なトークンは物理的なトークンである。これが機能主義の主張である。同一説も機能主義も唯物論である。同一説は心が非物質的なものからつくられていることを否定するが、機能主義はそれを否定しない。その点で機能主義は二元論と同一説の中間にある。

行動主義と機能主義の比較のために次のような図式を考えてみよう。

x y

ここで x 、 y はそれぞれ入力と出力である。中央の \square が Black Box と Turing Machine の場合でどのような違いが生じるのか。 \square が単なる Black Box から計算可能な関数の組に置き換わることは何を意味しているのか。従来、Black Box 型の説明と内部構造を用いた機能主義の説明は全く異なる心の理解の仕方に基づくものとして強調され、それが行動主義と機能主義を峻別すると考えられてきた。この一見説得力があるように見える比較は、しかし、本当に正しいのだろうか。まず、Black Box 型の説明、Turing Machine 型の説明がそれぞれ行動主義、機能主義において如何に重要な役割を演じていることが確認しておこう。

Skinner の攻撃は心理主義(mentalism)にあった。人間の行動を説明する際に内部の心理的現象に訴えて説明するのが心理主義である。そして、Turing Machine 型の説明は心理主義的説明の典型である。心理主義の誤りを彼は以下のように挙げている。

- 1 心理現象は直接に観察できない。したがって、心理現象は推論される。心的状態が推測であるなら、それを説明として使うことはできない。それはフィクションである。
- 2 心理主義的説明は真の説明とならない。それはその場限りの説明でしかない。
- 3 心理主義的説明は行動の事実をより曖昧な言葉で述べ直したものに過ぎない。
- 4 媒介項論証 (middle link argument) によって内部状態は回避できる。内的な心的状態は行動を説明するために仮定されるが、その状態は外部の刺激によって決定される。したがって、それは媒介的なものに過ぎず、安全に無視することができる。 $A \rightarrow B$ かつ $B \rightarrow C$ の B の役割が心的状態であるから、 $A \rightarrow C$ の結論には B は登場しない。
- 5 心理主義的説明は擬人的である。

このような Skinner の考えによれば、Turing Machine 型の説明は推測に基づく、擬機械的な、使わなくとも済む、無用の説明ということになる。

一方、Turing Machine をもとにそこから機能主義的な心の特徴づけを標榜する側にとって、彼らが腐心している機能的過程がすべてBlack Boxという主張は最初から議論の対象にさえならないと感じられる。Black Box型の説明は知的探求を中止した態度であるとともに、心の働きが外部環境に依存したものであるという根本的な誤りを犯しているように映る。Turing machine とコンピュータ科学、論理学との関係は特に密接であるが、機能主義の心の概念にも重要な役割を果たしている。Turing Machine の状態は入力、出力、内部状態によって完全に定められる。つまり、ある状態である入力を受け取ると何を出力し、どのような状態に移行するかだけで Turing machine の一つの状態が定まる。それが何からできているか、何を意味しているかといったことは不要である。機能主義は人間の心についてこの定め方を適用する。心的状態は基本的に Turing Machine の状態と同じであるとみなす。神経生理学的な脳の細部は不要である。こうして、心的状態は知覚（入力）、行動（出力）、他の機能的状態によって定義できる。Turing machine の状態は入力と出力だけでは定まらない。同じように、人間の心についてその定義には入力と出力（＝刺激、反応）だけでなく、心的状態が不可欠であるというのが機能主義の立場である。

機能主義的な心的状態の扱いに対する反論は行動主義からというより、その内部から出された。その代表がSearleの反論である。Chinese-room argument では Turing Test が前提されている。このテストに合格するなら、合格したコンピュータはテストされた事柄に関して人間と区別がつかないというのがこのテストの趣旨である。さて、サールの論証の意図は人間の思考過程に形式的に同型である計算アルゴリズムを実行して Turing Test をパスしても、思考を生み出すには十分では有り得ないことを示すことにあった。もしこの論証が正しければ、これは様々な機能主義に対して同じように言えるので、機能主義は人間の心の解明の原理としては誤っていることを意味している。さらに、サールが強いAI(strong AI)と呼ぶ考えは根本的に誤っていることになる。というのも、強いAIは「思考＝形式的規則に従う記号操作」を仮定する機能主義であり、思考するコンピュータをこの仮定に基づいて作り出すことができ、かつ人間の認知過程をそのコンピュータを使ったシミュレーションによって研究できると主張するからである。

サールは次のような思考実験を考える。漢字を一切知らない人に漢字を操作する全く形式的な規則の組を与えたとしてみよう。その人は記号を識別できても、それが何を意味しているのかわかっていない。規則は漢字が何を意味しているかは述べていない。規則は単にある形の記号が部屋に入ってきたら、紙に別の形の記号を書かなければならない、また、あるグループの記号は別のグループの記号と一緒にいる、そしてその順序はしかじかである、といったことを述べているだけである。その人は部屋に座り、規則を次々に適用して、紙の上に規則によって得られた漢字の別の集合を書き下す。そして、その紙を部屋の外で待っている別の人に渡す。その人には何のことが一切わからないが、彼女が適用した規則は文法的に正しい中国語の文を作り出す。しかし、彼女は自分が何をしたのかさえわかっていない。Searleによれば、その人はコンピュータが行っているのと正確に同じことを行っている。したがって、形式的な規則を使っての漢字の操作は、その人が中国語を理解するには十分でないならば、コンピュータにとっても中国語を理解するには十分ではない。そして、実際彼女には中国語が理解できない。この議論をまとめれば次のよう

になる。

公理 1：コンピュータのプログラムは形式的（統語論的）である。

公理 2：人間の心は心的内容（意味論）をもっている。

公理 3：統語論は意味論を構成しないし、またそれに対して十分でもない。

結論 1：プログラムは心を構成しないし、またそれに対して十分ではない。

言語を定めるには少なくともその言語の統語論(syntax)と意味論(semantic)を定めなければならない。言語の統語論は語彙とその他の補助的な記号を挙げ、カテゴリーに分類し、さらには文法を定めなければならない。ここには意味は一切入っていない。意味が関与する余地はどこにもない。これに対して、意味論は記号の解釈である。ある記号列が何を指示し、何を意味しているか定めることが意味論の役割である。真理、指示、意味はいずれも意味論的な概念である。言語を知るとは、したがって、少なくともその言語の統語論と意味論の両方を知る必要がある。

サールが論じていないのは何だろうか。まず、彼はコンピュータが思考できないとは言っていない。彼は二元論者ではない。人間は単に考える（生物的な）機械であると言い切っている。第二に、彼は考える有機体は生物学的な材料からできていなければならないとも言っていない。生物学的でない材料から考えるコンピュータをつくることは可能である。最後に、彼は思考が記号操作を含まないとも言っていない。

では、サールの積極的な主張は何か。人間の脳で実行される思考に形式的に同型の計算を備える他に「考える」ものをつくりだすには何が要求されているのか。サールによれば、何かを考えているかどうかはその計算アルゴリズム（ソフトウェア）だけではなく、そのアルゴリズムを実行するものの本性（ハードウェア）にも依存している。機能主義が問題なのは、彼によれば、思考に含まれる計算過程の物理的な実現から余りに抽象的なレベルにあり、どのようなものが思考を生み出す因果的な能力をもっているかの基準を提供していない。それは将来の研究対象である。現在の私たちが置かれているのは、思考の幾つかの具体例と思考しないものの幾つかの具体例を知っているに過ぎないという状況である。

しかし、その Searle が拠り所とする進化の結果としての心は生物学的なものであり、それは進化生物学であろうと例外ではない。それは行動主義的な構図と基本的には同じ構図である。そこで援用される進化生物学は純粋に外在主義的であり、それは行動主義と多くの共通点をもっている。外在主義的な理解は機能主義さえ外在化し、行動主義に引き戻すのである。行動の進化は Black Box 型の説明の一タイプなのである。

この結論の理由を述べるために、私たちはまず自然科学の模範である物理学を引き合いに出して二つの説明型を考えてみよう。

「プリンキピア」の自然観の具体化は時間空間的記述にある。この自然観のもとでは、宇宙はもはや原因ではなく、結果に過ぎない。運動方程式は法則であるが、目的ではない。時間空間記述は感覚的なレベルの特殊性から開放されて一般の座標系（ゲージ）に拡張される。このニュート

ンの成果に逆行するようであるが、システム全体に対して成立している運動方程式を装置として内在化すると Turing Machine 型のモデルとなる。原因や目的から開放された方程式に再びその役を担わせようというのは、物理学から見れば逆行以外のものではない。普遍的で、大域的という性質は方程式から取り去られるが、個別的で、局所的な方程式という計算装置が内蔵されるシステムが手に入る。法則の内在化は力学の観点からは逆行であるが、本来法則は局所的であっても構わない。但し、原因、目的といった概念が内在化によって必然的に再登場することは否定できない。これはニュートン的な世界観のもとでは、もし求められるならば目的や原因は外部、それも宇宙の外に置かれることになるのに対して、それらを内在化することになる。したがって、原因や目的をシステムの内部に局所化するのが Turing Machine 型である。しかし、ここには忘れてはならない巧みな工夫が隠されている。Turing Machine 型は物理システムの客観的な記述のための装置ではなく、選択的に情報を処理し、行動に結びつける装置に変身している点である。この変身は情報概念によって掴えることができる。自然の物理変化のすべてではなく、そのシステムの生存に必要な変化だけを選択的に取り出し、それだけを処理する。大域的な力学世界ではその世界の一部の運動、つまり、行動は法則の適用で記述・説明されるが、その法則は内在化してはいない。したがって、Turing Machine 型ではない。では、Black Box型か。すべてが Black Box型ではないが、それを許容する。それゆえ、力学的説明は Turing Machine 型とは両立しないが、Black Box型とは両立する。

次の例は進化生物学である。自然選択は選択がいつも外部から働くという点で Black Box型のモデルである。この自然選択の働き方の模範はニュートンの力学にあったことは周知の事実であるが、重力の法則に対応するような基本法則がなく、それゆえ、世界全体が Black Box型になる。環境に応じて働く力が変わるため、世界全体は収支決算でしか分からないということになる。原因と結果を見て、その時の選択の働き方を推測しなければならない。これはまさに行動主義の主張する Black Boxである。入力と出力の収支決算が媒介関数として捉えられれば、行動主義の主張そのものとなる。当然ながら、進化のレベルで扱われる対象は個々の物理的な対象ではなく、情報である。

直接にわからない現象や対象に対して私たちが採ってきた態度には大きく二つある。それら現象や対象を直接に扱う代わりに、それらを別のものに還元する、あるいは運動や機能から浮き彫りにするという態度がその一つである。これは多くの自然科学が採用してきた方法である。他の方法は、そのわからない現象や対象について内部にメカニズムや原理を仮定するものである。生氣論はその代表であり、Turing Machine 型説明も内部に計算のメカニズムを想定する点では同じ例である。このような分類から、行動主義が科学の常識的な方法を採用し、機能主義が科学では非常識的な方法を採用していることがわかる。

このような一見奇妙な結論はどのように考えたらよいのだろうか。 の意味を外在主義の徹底したものと考えれば、それは外側から の内容を制約するものを表すことになる。一方、Turing Machine のメカニズムを内在主義の徹底と考えるなら、どのような外部環境からの入力に対しても一定の対処をする自給自足型のメカニズムと考えることができる。行動主義の外在的性格はそ

れを進化へと広げることによって、行動の進化までも含まれることになる。

物理学の説明は外在主義的である。それは既述の力学の説明を見れば明らかである。外部の力による対象の運動の説明は対象内部のメカニズムを使っているわけではない。これは極度に広い内容であり、実際広い内容しか含まれていない。力学の対象の説明に広い内容と狭い内容の区別は存在しない。

自然選択のメカニズムと外在主義的な説明の関係を考えてみよう。生氣論を現在主張する者はまずいない。では、生氣論の意義はどこにあったのか。それは内在主義的に生命を捉える点にあった。生命原理は各個体に内在化されている。それが個体を動かす。しかし、生命原理が正しいか否かを決めるには外在化が必要だった。外在化によって、生命原理は物理的な原理と同列に置かれて、その真偽が確かめられることになる。その結果、私たちはどこにも生命原理なるものを見出すことはできなかった。外在化するとは物理的な対象の世界で自然選択や浮動を考えることである。この外在主義的な説明こそダーウィンの進化論を自然科学的な独立した理論として成立させたのである。だが、それだけでは進化論は一つの仮説に過ぎなかった。内在的な遺伝の仕組みがないことには説明は不完全であった。遺伝の仕組みは内在的である。遺伝の仕組みとその仕組みの由来は別であるという認識は内在主義と外在主義の調停が単なる混合というような簡単なものではないことを物語っている。

既に説明した内在主義と外在主義を上述のことと合わせて考えてみよう。まず、デカルト的な二元論を思い出してみよう。そこでは心と身体は全く異なる二つの実体で、私のことは私が直接に内観によって理解できた。この考えは未だに私たちに巣くっている。その一つが私たち自身の考えや信念から出発して世界の事柄を考えようという姿勢である。私の思考や信念の内容には私だけが特権的に近づくことができるので、この私の内的経験を基礎にして外的な世界の像をつくる。これがデカルト的な二元論の目標になる。その出発点を言い換えれば次のように表現できる。私たちの思考や経験は外界から独立している。全く心のうちにある純粋に内的な事柄である思考の表象内容が何にも増してすべての出発点である。

心と脳の二つを認めたとすると、心的状態が外界から独立している、あるいは遮断されているということは、私たちの思考がすべて「頭の中」にあって、脳の状態以外のものに依存していないということになる。心が脳にどのように依存しているかは既にいくつか述べてきた。なかでも付随性と複数実現性は広く使用されている概念であった。それを使うなら、「思考は全く頭の中にある」とは、ある人の思考の内容はその人の脳状態に付随するということになる。これが内在主義と呼ばれている考えである。これに対して、思考内容は頭の外にある、あるいは外部世界に関連しているというのが外在主義である。外在主義は脳の状態に付随せず、外部の状態に付随する心的出来事があることを主張する。

多くの哲学者が外在主義に関する思考実験を試みてきた。パトナムは桶の中の脳という身体から分離された脳が人工的に生かされて、活動している例を考え、そのような脳の思考は外部世界のものごとについてのものではなく、脳自身のメカニズムについてのものであることを論じている。彼の考えでは、志向性のもつ「...について」という特徴（例えば、牛についての表象や思考）

は、思考や表象が外部世界の適切な対象への正しい因果関係を必要とする。彼の主張は、外的な因果関係が志向性には必要である、ということになる。桶のなかの脳は世界からは遮断されている。したがって、そのような脳にとって外部世界についての思考はないことになる。

では、自由と決定は についての Black Box 型説明と Turing Machine 型説明からどのように特徴づけることができるか。Turing Machine 型の説明は自律的なシステムに適合するが、そのような自律的なシステムはそれ自身の行動を自ら決定しなければならない。その決定手続きは内蔵された仕組みによって実行される。その場合、説明を始める起点は Turing Machine の最初の状態である。これは極めて自然な意味で結果として生じる変化の原因と言ってよいだろう。行動の起点としての原因は何か別の原因によって引き起こされているかもしれないが、その Turing Machine にとっては強制的に与えられたものではない。また、その原因に始まる一連の変化は Turing Machine の規則に従ってコントロールされている。いわば Turing Machine の責任の下での決定論的な変化である。この狭い枠組では Turing Machine は「自由に自ら行動する」ことになる。

一方、Black Box 型の説明ではどうなるか。ここでは Turing Machine 内部だけではない広い枠組が説明の対象となる。 を一部とする環境が全体としてどのように変化するかが問題となる。むしろ、その変化はマクロな世界の変化であるから決定論的な変化である。だが、ここでの決定論的な変化は Turing Machine 型の場合と違って、起点は Black Box の中にはない。それは Black Box の遥か外にある。結果も Black Box のすぐ横にあるのではなく、延々と未来につながっている。Turing Machine 型が内部の局所的決定論であるとするれば、Black Box 型は内部と外部の区別がない広域的な決定論である。つまり、より運命論的なのが Black Box 型ということになる。確かに、Black Box 型の説明は Black Box そのものに特権を与えないという意味で、Turing Machine 型の説明よりは決定論的である。

完全な内在主義のもっとも強い形式は独我論である。他者や周りの環境を認めても、「私のことは私しかわからない」という立場を保持するならば、それは完全な内在主義と呼んでいいだろう。この立場を更に弱め、「概念は頭の中にある」という主張として内在主義を理解するなら、それは穏当な内在主義である。外在主義も同じように強い主張から弱いものまで広い幅をもっている。最後に、ここまで議論してきた事柄の趣旨をまとめておこう。

行動主義批判と機能主義擁護は整合性をもっていない。特に、機能主義の生物学的解釈はより広域化した行動主義であり、行動主義批判をそこに含むことができない。

機能主義の因果的解釈が強まれば強まるほど、因果の範囲の広域化によって行動主義に接近することになる。