

## デカルト（心の哲学の出発点）

[デカルトの推論の誤り]

心と身体の間にはどのような関係があるのか。この間に対して、心と身体は二つの異なる実体であり、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。このデカルトの伝統的な主張は今でも私たちの日常生活での基本的な常識となっている。憂鬱になれば、身体活動は鈍る。病気になれば、心が暗くなる。このような心身の密接な結びつきを前提にして私たちは毎日を過ごしている。このような伝統的態度とは違って、20世紀に展開された心の諸説の特徴は圧倒的に反デカルト的である。デカルト的な解答は自然主義的な解答ではないこと、20世紀の諸説は自然主義的であること、この二つから自ずと20世紀に展開された諸説の大半は反デカルト的ということになる。まず、デカルトの主張を推論の形で批判的に復習してみよう。

デカルトは自分が心をもつことを疑うことができない。

デカルトは自分が脳を持つことを疑うことができる。

それゆえ、心と脳は別のものである。

（結論を「デカルトは心と脳は別のものであると考える」としたら、この推論は正しいか。）

「心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、心と身体が異なることをライプニッツの不可識別者同一の法則（ $A$ と $B$ が同一とは、 $A$ のもつ性質はみな $B$ のもつ性質であり、その逆も成立する）から導き出す。ここで次のような例を通じて上の議論がある誤りを含んでいることを見てみよう。その誤りはフレーゲの意味論的な指示（reference）と意味（sense）の違いに関係しており、命題的態度（propositional attitude）と志向性という概念を通じて考えられてきたものである。

史門は明けの明星を観察したい。

史門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、史門が観察したいという性質は明けの明星にはあるが、金星にはないことから、明けの明星と金星は異なるものとなる。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。

史門は明けの明星を観察する。

史門は金星を観察する。

上の文では明けの明星と金星は同じものである。これは、史門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。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だけから、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。

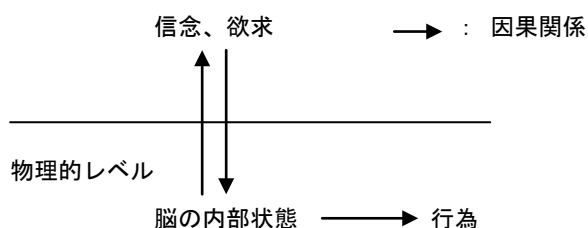
「疑う」や「望む」は命題に対する私たちの心的な態度である。たとえある命題が疑われ、他の命題が疑われなくても、最初の命題が述べているものと二番目の命題が述べているものが異なるということは導かれぬ。「私が脳をもっている」と「私が心をもっている」が異なった命題であって、一方は疑うことができ、他方は疑うことができないとしても、そこから脳と心が異なるということは結論できない。したがって、デカルトの論証は誤っていたということになる。デカルトの論証から、心と脳は異なるという結論は出てこない。しかし、このことから心と脳が同じことも、心身間に相互作用がないことも証明されたのではないことに注意してほしい。

（問）意味と指示の違いをフレーゲはどのように考えたか述べよ。

（問）上の問から、信念と事実の違いはどこにあるか、再考してみよ。

(デカルトの心身二元論の図式)

心的レベル



デカルト以後、心身関係や心の構造について多くの試みがなされてきた。その中には実に優れた考察が数多く含まれている。とはいえ、組織的に心を研究し出したのは20世紀に入ってからである。そこで、歴史には忠実でないが、20世紀の組織的な研究の追及に話を移そう。

(問) 心身二元論の図式で想定されている心的レベルと物理的レベルの相互作用はどのようなものと考えたらよいだろうか。(物理世界での通常の因果的な相互作用、社会の中での相互作用と比較すると、心身の相互作用はどのように異なっているか。)

## 行動主義<sup>1</sup>

### 1 論理的(分析的)行動主義

[心は世界のどこにもない]

哲学的あるいは論理的な行動主義 (*Philosophical behaviorism, or Logical behaviorism*) 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。ライル (Gilbert Ryle, 1900-1976) はデカルトの心身二元論を批判し、それを「機械の中の幽霊ドグマ (*the ghost in the machine dogma*)」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。この主張は誤っているというのが彼の考えである。というのも、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いるのではないからである。これは心理主義への強い反撥である。この反撥は、次の論証に明確に表われている。

心的状態が行動の内的な原因であるとすれば、それを見ることができないため、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。

(最初の前提は何を述べているだろうか。経験主義の主張は何だったか。)

[自然主義的誤謬]

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と呼ばれるものである。「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。このライルの主張は正しいだろうか。

(i) デカルトは心もち、脳をもつが、それら二つは異なる。

(ii) デカルトは心もち、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、心が脳の上に存在するものであると考えるのは誤りである。

これら三つの命題はそれぞれ(i)が二元論、(ii)が同一説、(iii)が行動主義の主張を表している。これらは言語の規則

<sup>1</sup> 行動主義の特徴付けは E. Sober, *Core Questions in Philosophy* に負う。

に反しているだろうか。反してはいない。それどころか、三つの命題はいずれも有意味な主張である。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である。(デネットは次のような例を挙げている。手術中に体の自由を奪う麻酔と、その記憶を手術後喪失させる薬を併用した場合、それは通常の麻酔と同じ効果をもつだろうか。通常の麻酔なら痛みはないが、この場合はどうか。)

論理的な行動主義は哲学的な主張に基づく徹底した行動主義である。そこには哲学的な主張の素朴なまでの徹底が見られる。

## 2 方法的行動主義

[心は研究対象にならない]

デカルト的な二元論によれば心は非物理的 (延長をもっていない) で、第三者には観察できない。それゆえ、心理学は経験科学としては成立不可能である。経験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。このように主張する方法的行動主義 (Methodological behaviorism) 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、つまり、科学方法論の主張である。まずこの点が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否定しない。しかし、信念や欲求が外側からは観察できない内的状態ゆえに、それらについて心理学的に語ることは方法論上否定される。

[スキナーの主張とチョムスキーの批判]

方法的行動主義の否定的なテーゼは、心理学は信念、欲求を使った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 (Burrhus F. Skinner, 1904-1990) に代表されるが、心理主義的な説明に対する彼の反対理由の一つは、信念や欲求は観察できないというものである。それらは隠れている。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作り上げることができてしまうからである。つまり、信念や欲求をアブダクションで勝手に導入してはならないということである。

しかし、スキナーの行動主義はチョムスキー (Noam A. Chomsky, 1928-) によって鋭く批判される。スキナーの『言語行動』(Verbal Behavior) に対するチョムスキーの書評は次のような内容であった。私たちは有限の語彙と文法規則を使って、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取り扱うか知らない。また、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か。無限のもの、新しいものを行動主義は扱うことができない。この批判は方法的行動主義を崩壊させるのに十分過ぎるものであった。

(問) 論理的行動主義と方法的行動主義の違いを心の扱いを中心に説明せよ。

## 同一説

[徹底した唯物論]

心的な性質 (状態、過程) は脳の性質 (状態、過程) と同じものである。つまり、どんな心的性質も物理的性質そのものである。「気体の温度は気体分子の平均運動エネルギーである」と同じ意味で、心的性質は脳の性質である。これが同一説の主張である。自然主義の典型である同一説には微妙に異なる主張が含まれている。代表的な説を以下に挙げてみよう。

1 スマート (John J.C. Smart, 1920-) の主張: しばしば脳過程 (に関する) 唯物論と呼ばれ、感覚は脳過程と同一であると主張する。

2 アームストロング (David M. Armstrong, 1926-) の主張: 中枢状態唯物論と呼ばれ、心的状態は中枢神経系の状態と同一であると主張する。

3 ファイグル (Herbert Feigl, 1902-1988) の主張: 神経生理学的な用語と心的な用語は同じものを指示すると主張する。フレーゲの意味と指示の違いから、二つの用語の意味は異なるが、指示は同一であると考える。

[経験的な同一説]

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。二元論、生氣論と対比するなら、同一説は一元論であり、非生氣論である。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、その同一性の細部は経験的に確認される。その結果として、心は

還元・消去され、脳の振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、認知科学の内容は必ずしも同一説の立場に立って理解する必要はない。むしろ、その主張は同一説と異なるものである。そして、それは次の機能主義に基礎を置くものである。

(問) 同一説が経験的であるとはどういう意味か。

[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だった。「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。

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、 $(x)(y)[(x=y) \rightarrow (F)(Fx \leftrightarrow Fy)]$  (この論理式は高階である)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ため、それらは同一で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。空間的な局所性による批判を見てみよう。

(空間的な局所性を使った推論)

- 1 脳過程は空間的な出来事である。
- 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批判的な推論を展開できる。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。

## サールの批判

[計算機は何も理解しない]

1950年にチューリングは人工的な知能に関するチューリング・テスト(Turing Test)を提案した。このテストは二人の人間とテストされる機械(コンピュータ)からなっている。一人がコンピュータの端末のある部屋にいる。この端末を使って部屋にいる人は二つの会話を行う。一つはコンピュータとの、他はもう一人の人の会話である。但し、その人は会話の相手がコンピュータか人間のいずれであるか教えられていないとしよう。

まず会話がなされ、誰と話したのか尋ねられる。話した相手がコンピュータなのか、それとも人間なのか答えなくてはならない。会話毎にこの質問は何度も繰り返される。繰り返しの後には何回の中したかの結果が出るだろう。50%的中率なら単なる当てずっぽうということになるが、適切な会話時間の長さのもとで、コンピュータが人間と間違われることが十分な頻度で起こったならば、そのコンピュータには知能があると言ってもよいだろう。これがチューリング・テストである。

チューリング・テストが知能のテストとして十分である理由はなにか。テストにパスするには世界や人間に関する広範囲の知識をもつことが必要であるばかりではなく、新しい環境に的確に反応できる柔軟性が求められる。この二つの条件を満たすには知能無くしては不可能である。

これに対して異議を唱えたのがサール(John Searle, 1932-)だった。この批判は、「中国語の部屋の論証(Chinese-room argument)」と呼ばれている。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を解明する原理としては誤っていることを意味している

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知らない。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、そしてその順序を述べているといったものだけである。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。その人にはわからないが、彼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、その部屋の人には自分が何をしたのかさえわからない。サールによれば、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、形式的な規則を使つての漢字の操作は、その人が中国語を理解するには十分でないとするば、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。この議論をまとめれば次のようになる。

コンピュータのプログラムは形式的（統語論的）である。  
人間の心は心的内容（意味論）をもっている。  
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。  
したがって、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もない。

言語を定めるには少なくともその言語の統語論と意味論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、その言語の統語論と意味論の両方を知ることである。だが、コンピュータは統語論しか知ることができない。

サールによれば、何かが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ることである。彼は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は提供していない。それは将来の研究対象である。しかし、私たちは思考するものの具体例と思考しないものの具体例を知っている。

[サールの論証への反応]

サールの論証への反論や批判は実に多い。サールの論証は認知科学の哲学的な支柱である機能主義が十分に私たちの心を捉えることができないという主張であるから、機能主義的に心を捉えることができると考える哲学者、認知科学者の関心を強くかった。

哲学的な反応の代表はチャーチランド夫妻（Patricia and Paul Churchland）であろう。サールの3番目の仮定を信じる根拠がないというのが彼らの考えである。それは次の仮定であった。

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。

彼らは部屋の中の人には中国語を理解していないという点ではサールに賛成する。その人、部屋、規則集、何枚かの紙からなるシステムが中国語を理解していると言うのもばかげている。しかし、このことから記号を操作・処理するシステムが意味論をもち得ないと考える根拠はないと彼らは主張する。私たちは人間の認知の働き、そして記号操作システム一般について十分知らない。したがって、ただ一つの思考実験から一般的な結論を得ることはできない。

サールの思考実験ではこの仮定を証明するのではなく、それを仮定しているだけである。この仮定は強いAIの基本的な仮定の否定に過ぎない。それゆえ、彼の論証は論点先取である。この点を明らかにするために彼らは次のような類比的な論証を考案し、電磁力が光を構成できないことを示す。

電気と磁気は力である。  
光の本質的な性質は輝きである。  
力それ自体は輝きを構成しないし、またそれに対して十分でもない。  
それゆえ、電気と磁気は光にとって必要でも十分でもない。

誰かが手で磁石を振動させても光はつくりだすことはできない。私たちに見える光を生み出すには振動が余りにゆっくり過ぎるからである。いずれにしろ、これは電磁現象のただ一つの例でしかなく、そこから一般化することはできない。

（ネーゲルの物理主義に反対する論証：コウモリであるとはどのようなことか（T. Nagel, 'What is it like to be a bat?',

*Philosophical Review*, 83, 435-50, 1974.)。)

ネーゲルが物理主義に反対する論証を示したというのは正確でないかもしれない。彼の主張は、物理主義がどのようにして真であり得るのか何もわかっていないと述べる方が正確である。彼によれば、有機体が意識経験をもつという事実は、その有機体であるような何かがあることを意味している。これは経験の主観的性格と呼ばれる。それは単一の観点に結びついている。つまり、意識経験があるところには常にその経験をもつ個体が存在し、その個体はその経験に一意的な関係にある。

これに対して、物理的な状態は主観的な性格をもっていない。酸素は 8 の原子番号であるがそれは主体が異なっても同じままである。つまり、物理的な状態は客観的である。そうだとすると、意識は物理主義に問題を投げかけることになる。意識は本質的に主観的現象であり、物理世界は客観的現象であるから、私たちはどのようにして意識が物理現象に還元されるか、あるいは関係しているか理解する術をもっていないことになる。純粋に物理的対象、出来事、状態、過程を使ってどのように主観性が説明できるというのか。彼が使った例はコウモリだった。私たちはコウモリが意識経験をもっていることを疑う十分な証拠をもっていない。たとえコウモリに主観的な経験があるとしても、コウモリの意識経験の主観的性格は私たちにはわからない。ソナーを使って空を飛ぶとはどのようなことかは私たちにはわからない。ネーゲルの論証は次のように整理できる。

- (1) 主観的な経験が存在する。
- (2) 物理主義が正しいなら、主観的経験についての物理的な説明がなければならない。
- (3) しかし、すべての主観的経験が本質的に単一の観点に結びついており、したがって、主観的対象の十分な概念はその主観的観点の採用によってしか得ることができない。
- (4) 客観的な物理理論は主観的観点を採らないので、経験の主観的性格を説明することができない。
- (5) 経験の主観的性格がどのようにして有機体の物理的な操作の中に実現されるかは謎である。
- (6) 物理主義は私たちが理解できない立場である。なぜなら、私たちはそれが現在どのように真であるかについて何の考えももっていないからである。

この論証から何が結果として得られるのか。まず、私たちの知識の限界である。他の動物の主観的経験についての事実は私たちには知ることができない。さらに、そのような事実は物理的事実ではない。(したがって、物理的事実と主観的事実は異なることになる。) それゆえ、物理主義的枠組みで扱うことができない事实在存在し、物理主義は誤りということになる。

(問) 主観的でない客観的経験とはどのようなものか。

(マッギンの物理主義反対の論証(C. McGinn, *The Problem of Consciousness*, Blackwell, 1991.))

マッギンとネーゲルの論証には多くの共通点がある。マッギンは物理主義者である。だが、彼は意識が脳にどんな役割をもつか発見できないと思っている。彼の論証を、すべての出来事は物理的な出来事である、あるいは物理的な出来事によって説明できるという主張であるとするならば、物理主義に反対するというよりそれと両立可能である。彼は次のように議論を展開する。

- (1) 意識を説明する脳の物理的性質がある。
- (2) しかしながら、私たちの心の構造(特に、概念形成能力)はその脳の物理的性質を把握する、理解することができない。
- (3) したがって、脳がどのように意識を生み出すか私たちにはわからない。

彼は(1)が正しいか否かより、私たちはそう信じていると述べるだけである。興味あるのは(2)である。脳や意識の研究を続けるなかで、私たちは意識を説明する脳の物理的な性質をどのように確定できるのだろうか。意識を直接に内観によって調べることによってか、あるいは脳を探索することによって見つけるかであろう。しかし、内観は脳や神経機構については何も教えてくれない。したがって、脳を調べるほかなくなる。脳の直接観察によって何がわかるだろうか。脳がもつ、身体他の部分と同じ特徴、異なる特徴はわかるだろうが、現象的意識の主観的性格はどこにも見出せないだろう。脳の直接観察の結果からそのような性格を間接的に導き出すことはできないだろうか。これもできない。というのも、説明したい主観的性格は理論的前提と観察結果のいずれにも含まれていないからである。脳の理論化(自然化)に意識は含まれていない。脳のある性質が確かに意識を説明する。しかし、私たちが知識を得る基本的なメカニズム(内観、観察、そして理論化)はそれがどんな性質かを言い当てることができない。私たちのも

つメカニズムは意識を知るには余りに不十分、不適切である。そのため、意識がどのように生じるかは原理上知ることができない。脳は、それがスマートである以上に複雑なのである。

私たちは本当に意識について理解することができるのだろうか。私たちの脳は進化の産物であり、したがって、認知に関して限界をもっている。認知に関して完全に適応しているわけではない。ネコやネズミが量子力学を理解できないように、人間は心とものとの関係といったものを理解できない。意識は永遠に人間には理解できないものである。これがマッギンの考えである。この悲観論をもとにそこから反自然主義的な立場を取るといふ方向に彼が進まないのはサールの場合と同じである。

意識はどこにあるのか。デカルトによれば意識は空間には存在しない。空間に場所を占めるというのは心の存在の様式ではない。したがって、心的状態を認識する仕方は知覚を通じて知る際に使う空間的概念を使って把握することができない。物理学には観察できない対象が多く存在するが、それと同じように心的状態を考えることもできない。原子が肉眼では観察できないことと同じ意味で意識は観察できないのではない。単に観察できないどころか、それを空間的な対象としてどのように考えればよいかさえ私たちは知らないのである。したがって、技術的に知覚できないのではなく、原理的に知覚できないのである。意識が延長をもたないというデカルトの直観は今でも私たちの心的なものについての日常的な概念の核心部分を占めている。意識を理論的に扱う場合に、この意識の特徴は空間的な扱いを拒絶することになる。

もし意識が空間的でないなら、それはなぜ空間をもつ世界に生まれ出ることができたのか。意識も進化の過程の中で生み出されたものだとしてみよう。すると、それは宇宙の進化、生命の進化と違って、意識でない空間的なものから空間的でない意識への進化であった。これは大変神秘的に見える。これが意識に関する空間問題である。この問題には歴史的に二つの主要な対応の仕方があった。一つは、意識は物質から生まれたという前提を否定することである。意識は物質から独立したものであると考えれば、デカルトの立場、伝統的な二元論になる。他は、意識が非空間的であるという前提を否定することである。意識状態が脳状態でしかないなら、それは古典的な唯物論になる。この二つの立場に対して第三の途が考えられる。それは次のように述べられる。物質的な発現を保ちながら、通常非空間的な意識概念を否定しないことである。脳が行うことを説明しようとする、脳は現在の物理科学で認識された空間的な性質だけでは十分ではないので、脳はそれ以上の性質をもっている。この性質は今の科学では捉えることができない。特に、意識の発現に関してはそうである。したがって、意識は現在の科学にとって端的に謎となる。

心身問題を解決するには少なくとも空間の新しい概念が必要である。空間がどのように構成され、構造化されるかを考えなければならない。これは単に脳生理学の問題ではなく、物理学まで及ぶ問題である。科学の歴史の中での空間概念、日常生活での空間概念は現在の物理学での空間概念とは大きく異なっている。したがって、物理空間は人間の知覚－行動システムに与えられた空間とは大きく異なっているに違いない。意識はこの世界のどこかに局在してはいないが、この世界の外にあるわけでもない。この意識と空間の特異な関係を記述できるような理論を私たちはもっていない。だが、意識と空間は自然主義的仕方で結びついているはずである。ここに空間問題を解く鍵がある。

(問) 意識が空間的でないのはどのような理由からか。

(ジャクソン (Frank Jackson) の物理主義反対の論証：知識論証 (The Knowledge Argument) (F. Jackson, 'What Mary Didn't Know', *Journal of Philosophy*, 83, 1986, 291-5.))

私たちはバラが赤いと思うし、その赤さを経験できると思っている。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予め組み込まれていないことを示している。さらに、私たちとは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な機能主義者である。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論証の代表がジャクソンの知識論証である。

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。色の知覚に関して知らなければならないあらゆる物理的な事実を彼女は学んだ。物理世界の見目の完全さは、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を示している。これが正しければ、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、彼は次の問いを出す。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」がどのようなものか学ぶのではないか。そのような学習によって新しい知識が得られとすれば、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ことになっているからである。この思考実験を論証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。

(問) ジャクソンの論証についてあなたはどうか。メアリーは最初に色を見たとき、何を経験するのだろうか。

(ジャクソンの批判への反応)

ジャクソンの批判は歯切れがいい。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。以下に四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（コツや技能）に過ぎない。
- 3 彼女は新しい事実的な知識を獲得する。しかし、それは既知の知識を新しい様式で獲得するに過ぎない。
- 4 彼女は新しい事実的な知識を獲得する。しかし、それは物理主義と両立する精妙な仕方では個別化したものに過ぎない。

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。最初のを考えよう。チャーチランド、デネットがその代表で、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。私たちは現在の物理主義的知識と未来のその知識がどのくらい隔たっているかを知らない。

2 番目の反論をチャーチランドは次のようにまとめている。

- (1) 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。
- (2) 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。
- (3) 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。

(1) と (2) 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 (2) 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。別の2に対する反論によれば、メアリーは新しい経験によって獲得したものがあるが、それは事実的知識ではなく、コツや技能に過ぎないというものである。

(問) これまでの話しから君は所謂「心と身体の関係」をどのように考えるか。