

第1節 人工の感情と人工のクオリア

感情を作る

心とは何だろうか。学生の試験答案から見る限り、それは感情だ。「ロボットは将来心を持てるようになるか」という試験問題を前にして、学生たちの大半は、心を持つとはどういうことだろうかと自らに問いかけ、それは感情を持つことだと答える。もちろん彼らは、感情を正確に分析しているわけではない。しかし、それでも学生たちは、ロボットが心を持つには感情を持たねばならないと考え、感情は人工的に作れないがゆえにロボットは心を持ってない、と答えるのだ。この答えは、脳神経科学者のダマシオの答えと重なる。彼は、意識をもつ人工物を創造することができるのかと問い、それに「イエス」と「ノー」の二つの答えを与えている。「ノー」は、いわば掛け値なしに人間と同じだと言えるような意識を創造することに対して、「イエス」は、意識的行動にそっくりなことを行う形式構造上の意識もどきを作ることに對して(1)。そして、ダマシオはこの「イエス」と「ノー」を分かちものをこう述べている。「要するに、感情がバリアーである。なぜなら人間的な意識の実現には、感情の存在が必要だからだ。『見かけの』情動ならシミュレートできるが、感情のようなものをシリコンに複製することはできない。」[Damasio, 2000, p.\*\*\*]。

しかしなぜ、感情を人工的に作るのは不可能だと思われるのだろうか。一つには、感情を人工的に作るというときに何を作ればいいのかを判然としない、ということがあるだろう。感情の重要な一部が感情表現であることは間違いない。現在作られている多くの「癒し系ロボット」が、触れられたり話しかけられたりすると人間や動物の感情表現の多くを模倣する「感情ロボット」であるのは興味深い。人間とロボットのなめらかなコミュニケーションの基盤が感情表現であるかのようだ。しかし、感情表現と感情そのものとは同じではない。喜びの表現の仕掛けをいくら巧妙に作り出したとしても、喜びそのものを作り出したことにはならないだろう。感情表現は、感情が果たしている機能の一部ではあろうが、喜びの表現が他者に対して喜びの表現として認知され、それとして機能できるのは、そもそも喜びがその機能を当人の内部で果たしている結果である。

では、感情の表現ではなく、何を人工的に実現したら感情そのものを作ったことになるのだろうか。少し考えてみれば分かるように、悲しみの感情だけを実現する機械、あるいは嫉妬の感情だけを実現する機械というのは、悲しみの感情を実現するだけの機械、嫉妬の感情を実現するだけの機械と同様に、何か極めてナンセンスな代物である。つまり、感情は、見分けることや推論することなどと異なり、人間の認知能力の部分的なシミュレーションの対象となるには不向きなのだ。それは、感情がそれらの認知モジュールと同列の一つの認知能力などではなく、それらを駆動し、それらの機能に介入し、それらを束ねる最上位の機能だからである（[柴田, 2001]は、それを通常の認知機能に対する2階の機能として分析した）。

つまるところ、われわれが期待するフルヴァージョンの感情は、さまざまな認知機能を搭載したフルヴァージョンの完成体のロボットくらいにしか実現できないだろう。もっとも、感情が、現在の人間を完成品として目論んだある設計者の

手になる単一の存在なのではなく、何百万年もかけて進化の途上でそのつどそのつど、その場しのぎ的に獲得されてきた多層的で、モザイク的な機能の集合体であるなら、その〈感情もどき〉を部分的にはシミュレーションするくらいはできるだろうが・・・（本書出版のきっかけとなった科学研究費「意識と感情をもつ認知システムについての哲学的研究」は、感情の最下層にあると仮定されたく共感〉のニューラル・ネットワークによる実現に焦点を合わせている）。いずれにせよ、感情に特異なこうした事情を考慮に入れたとしても、われわれが感情の人工的实现に関してできることは、感情の機能を再現することだけであろう。

しかし、この感情機能の再現に関しては、先のダマシオのように「ノー」と答える必要はないと私は考える。ダマシオの論点は、実際にはシリコンというより、身体と脳の相互作用だ。「感情は、肉体が複製されなければ、肉体への脳の作用が複製されなければ、そして、肉体へ脳の作用が及んだあとの脳による肉体の状態探知が複製されなければ、複製することはできないのである。」[Damasio, 2000, p.\*\*\*]。ということは、それらが複製されれば、感情は複製されうることになるだろう。もちろん、ダマシオの言うように、その際のロボットは身体から切り離された単なる脳＝コンピュータではないはずだ。しかし、人間の〈身体＋脳〉が果たす機能を逐一マップした「肌理の細かな」(fine grained)な機能同型体をシリコンや他の物質で構成することができない、という理由はアプリアリには見当たらない。

感情の存在意義がアドレナリンやドーパミンといった神経ホルモンの物質素材そのものではなく、それらが身体や脳との関係で果たす因果的な役割、すなわち機能であるのは、治療に用いられるさまざまな薬や麻薬などの代わりに別の化学物質を用いても同じ効果が得られるという事実からも明らかだろう。要するに、感情の生成と変化を司るとされるさまざまな神経ホルモンがなぜそれらの神経ホルモンでなければならなかったのか、ということの理由は、脳や身体の生化学的なシステムが現にある物質素材から作り上げられたから、ということである。もう少し機能主義的に言えば、脳や身体の物質素材がタンパク質や脂肪や水分といった現にある素材であるがゆえに、それらの複合体に感情の機能を持たせようとしたら、その機能を果たすことのできる物質がアドレナリンやドーパミンという化学物質であった、ということだ。それゆえ、肌理の細かな機能同型体である人工の〈身体＋脳〉が感情の機能を持つには、人間にとっての感情の実現物質である神経ホルモン群ではなく、何か他の素材を、その人工物の素材に合わせて発見しなければならない。それは困難な課題ではあろうが、アプリアリに不可能な課題ではなく、機能に導かれた経験的な探求に他ならない。したがって人工の感情は、感情機能に関する限り原理的には作成可能だろう。

#### クオリアを作る

しかし、問題はここから先であり、それこそが感情を人工的に作ることが不可能と思われる真の理由であり、また感情機能と感情クオリアの関係の解明という本章の主題でもある。冒頭の学生たちの答案に戻ろう。彼らがロボットには感情が持てないがゆえに心は持てないと述べる時、彼らが念頭に置いているのは、明らかに、感情が果たす何らかの機能ではなく、感じられる限りでの感情、つまり感情のクオリアである。悲しいと思われたり、嬉しいと思われたり、悔しいと思われたりするときに感じられるあの感じそのものだ。それこそが〈生きている〉ということの実感を与えてくれるものであるから、学生たちも、それを心の実質と考えるのである。そこで、問題はこうなる。感情の機能が人工的に作れたとして、それによって、感情のクオリアまで作成したことになるのだろうか。ある

いはもっと率直に言えば、感情のクオリアはいかにしたら人工的に作ることができるのか。

感情機能を人工的に作成するという時の困惑は、クオリアを人工的に作成するという時の困惑ほど大きくはないだろう。なぜなら何かを作るといっても、そもそもわれわれには、極微少の組み替え遺伝子から巨大な宇宙船まで、既存の物理的なものをどうにか操作して新たな物理的なものを作る以外にはないからである。心に一番関連の深いわれわれの操作さえも、脳という物理的な装置への物理的もしくは化学的介入である（言葉による他者の心への訴えすら、音声や文字という物理的な手段を必要とする）。つまり、心的なものそのものを何も物理的な支えなしで作るということは、およそ途方にくれる課題なのだ。ましてや、クオリアはその心の持ち主が第一人称的に、しかも彼だけが、内側から経験するものである。したがって、純粹に心的なものとしてのクオリアを人工的に作り出す方法が理解しがたいだけではない。たとえクオリアの可能的経験者を物理的に作り出したとしても、その当の存在者がクオリアを経験しているかどうかは、第三人称の現象とは異なり確かめようがないのである。

さて、このような人工の感情クオリアに関して私が用意できる答えは、現実世界においては、感情機能を人工的に作ることによって、またそれによってのみ、感情のクオリアを人工的に作ることができる、というものである。

以下では、この答えへの一歩として、人工のクオリアではなく、自然のクオリア、つまりわれわれ人間のクオリアの存在論的な枠組みを検討しよう。その際、私は、物理的存在とクオリアに関する以上のような乖離性を正面から受けとめたい。つまり、機能としての感情とクオリアとしての感情の間には、科学や技術によっては埋めることのできない存在論的な亀裂がある、ということを確認すべきだと私は考える。しかし同時に、機能とクオリアの両者には恐らく限定的だが存在論的な繋がりもまた存在し、それが両者の法則的な繋がりとの基盤を与えているはずである。しかし、この見方はもちろん、現在のところ私の単なる提案であり、しかも本章で扱うことができる議論も、機能を担う物理的存在とクオリアとの存在論的關係を通じた、機能とクオリアの關係についてだけである。

#### チャルマーズのゾンビ論法

そこで、物理的な存在とクオリアの乖離に関して、まずチャルマーズのいわゆるゾンビの論理的可能性を出発点にとろう。この余りにも有名になった議論は、チャルマーズが物理主義の誤りを示すために提出したものだが、それを感情機能と感情クオリアの關係に関する一般的ケースとして取り上げる。そのゾンビ論法を本節で確認した後、偶然性のテーゼ（ローカルな真理）としての物理主義(2)と性質二元論的物理主義について概観し(2節)、ゾンビ世界とわれわれの物理主義的世界の違いをチャルマーズの「現象的判断のパラドックス」から論ずる(3節)。そして最後にそこから、クオリアを含めた現象的性質と機能との關係について若干の考察を引き出したいと思う。

まず、チャルマーズの物理主義の定義とその反駁を確認しておこう。チャルマーズが物理主義の誤りを論証するステップは次の通りである。

- (1) われわれの世界には、意識経験が存在する。
- (2) われわれの世界と物理的に同一だが、意識に関する肯定的事実は成立していないような論理的に可能な世界が存在する。
- (3) したがって、意識に関する事実は、われわれの世界の物理的な事実を超

えでたさらなる事実(further facts)である。

- (4) それゆえに、唯物論（／物理主義）は誤りである[Chalmers, 1996, p. 123.]。（「物理主義」の語は引用者による挿入）。

要するに、われわれの世界とまったく物理的に同一であるが、意識経験が生じていないようなゾンビ世界が論理的に可能であるがゆえに、物理主義は誤りだ、というのが趣旨である(3)。したがってここで、物理主義は、あらゆる事実に関してチャルマーズの言う論理的スーパーヴィーニエンスの主張を含意すると見なされていることになる。つまり、意識やクオリアは物理的事実に論理的にスーパーヴィーンしないがゆえに、物理主義は誤りだという論法だ。そこで、チャルマーズの言う論理的スーパーヴィーニエンスとは何かを押さえておこう。

「B 性質が A 性質に論理的にスーパーヴィーンするのは、A 性質に関して同一であるにもかかわらず B 性質に関して異なっているような、二つのいかなる論理的に可能な状況も存在しない場合である。」[ibid., p. 35]

大ざっぱに言えば、ここでの論理的可能性は概念的な想像可能性(conceivability)に対応する。神は、論理的に可能なものが存在する世界を作りえたであろうが、不可能なものが存在する世界を作ることはできない。もし意識が物理的事実に論理的にスーパーヴィーンするなら、2つの可能世界で同じ物理的事実が成立しているのに、片方では意識が存在しもう片方では意識が存在しない、というのは論理的に不可能である。したがって、その場合には神はゾンビ世界を作りえないことになる。この論理的スーパーヴィーニエンスの特徴は、次の自然的スーパーヴィーニエンスと対比することによってもっとはっきりするだろう。

「B 性質が A 性質に自然的にスーパーヴィーンするのは、同じ A 性質をもつものの二つの自然的に可能な状況も同じ B 性質をもつ場合である。」(ibid., p. 36)

自然的に可能な状況とは、いかなる自然法則も侵犯することなしに自然に生じうるような可能性であり、これも大ざっぱに言えば、実際の経験的可能性に対応する。例えば、反重力装置や永久運動機関は論理的には可能であっても、自然的に可能ではない。言いかえると、「もし・・・であったなら」という反事実的条件文の真偽がすべての自然法則の保存の下で値踏みされる場合、その評価の行われる場所が自然的に可能な世界なのである。したがって、意識が物理的事実に論理的にはではなく、自然的にしかスーパーヴィーンしないなら、2つの可能世界で同じ物理的事実が成立しているが、片方では意識が存在しもう片方では意識が存在しない、というのは自然的に不可能であっても、論理的には可能である。つまり、このような可能世界（ゾンビ世界）が、自然的に可能な世界群の外側のどこかに存在することになる。

## 第2節 偶然的真理（ローカルな可能世界群における真理）としての物理主義

### 最弱の物理主義

さて、「物理的事実」ということで、物理的性質のすべての例化と配置（全基本粒子の状態）のみならずすべての物理的法則も含めて考えるならば[ibid., p. 33]、物理主義の主張が、「世界のすべての事実は物理的事実に論理的にスーパーヴィーンする」というものだと解釈されるのは、一見すると当然であるように思われる。というのも物理主義的直観とは、ひとまずは素朴に、「世界には物理的なもの（個体と性質）以外は存在しない」という唯物性の主張か、少なくとも、「物理的なものがすべての事実を決定している」、もしくは「物理的に同一でありながら、それ以外の何かが異なっているようなことはありえない」という決定

性の主張として理解されるようなものだからだ(4)。チャルマーズの言い方を借りれば、もし物理主義が正しければ、神はいかなる世界であれその世界の物理的事実を決定しさえすれば、それでその世界のすべてを創造したことになる。なぜなら、物理主義は、物理的事実が同一でありながら他の事実が異なることはありえない、つまり他のいかなる事実も物理的事実に論理的にスーパーヴィーンするという主張だからである。

そこで、チャルマーズの論点をもう一度まとめるならこうなる。物理主義は、一切の事実が物理的事実に論理的にスーパーヴィーンするという主張にコミットせざるをえない。しかし、われわれと物理的に同一の事実が成立していながらも意識経験を欠いた世界は概念的に可能である(ゾンビ世界の想像可能性)。それゆえ少なくとも意識経験という事実は、物理的事実に論理的にスーパーヴィーンしない。したがって、物理主義は誤りだ。

ところで、チャルマーズの議論においては、明らかに物理主義は、物理的なものが存在するあらゆる可能世界をスコープとする主張として理解されている。したがって物理主義は、物理的なものが存在しないあらゆる可能世界においても空虚に成り立つがゆえに、あらゆる可能世界で成り立つ必然的真理だ、というテーゼとして理解されている。しかし、そうする必要はあるのだろうか。仮に物理主義が正しいとして、それは端的に必然的な真理である必要があるだろうか。あるいは逆に、実体二元論が誤りだとして、それが偽であることは必然的真理なのだろうか。もう少し別の言い方をすると、霊や天使たちが物理的実体と共に「うようよしている」(後述の[Lewis, 1983]の訳者たちの言い回し)ような世界は論理的に不可能なのだろうか。確かに、われわれの世界やわれわれの世界が属する可能世界のある種のグループ(集合)においては、そのようなものは存在していないだろう(最弱の物理主義者たらんとする私も、そこまで譲歩する気はない)。しかし、もし物理主義の唯物性の主張が必然的真理なら、そのような実体二元論の誤りは概念的に不可能なことを主張することにあつたのだということになるが、果たしてそうだろうか。私にはそうは思われない。

また、物理主義が物理的決定性のテーゼを必然的真理として主張するというなら、次のような場合はどうだろうか。実体に関してはそもそも心的(霊的)実体しか存在しない、という世界は唯物性のテーゼは満たさないが論理的に可能だとしよう。これはいわば、心的実体一元論の世界である。そしてその世界に、われわれの世界の物質の一部、例えば1ダースほどの水素原子のみが片隅に加わった別の可能世界は不可能だろうか。さらに、その世界と物理的に区別つかないもう一つ別の可能世界、つまり物理的にはその1ダースの水素原子のみがまったく同じ仕方で存在する世界に、今度はさっきとは別の心的実体どもが「うようよ」存在するといったことは不可能なのだろうか。物理的決定性のテーゼが必然的なら、それは不可能なはずだが、その場合、物理主義は、たった1ダースの水素原子に関する事実が、何百万もの心的実体に関する事実をすべて決定しているのだ、と主張しているのだろうか。

私はここで、物理主義も二元論もともに、可能世界のあるグループにおいては真だが、別のグループにおいては偽であるようなローカルに成り立つ真理、つまり偶然的な真理の主張であると解したい。しかし、その主張は偶然的ではあるが、ある範囲、あるグループの可能世界のすべてにおいて成り立つがゆえに、自然法則にも似た限定的な必然性をもつだろう。

すると、とくに、われわれの現実世界と物理的に同一の可能世界なのだが、物理的以外の事実が現実世界と異なるような可能世界もまた存在するだろう。現実世界とゾンビ世界はそのようなペアである。だが、そのことは、物理的に同一な

可能世界のどれもが、物理的以外の事実においてもまったく同一だ（つまりスーパーヴィーニエンスが成り立っている）、というような無数の可能世界の存在を排除するわけではない。物理主義は、現実世界を含むそれらの可能世界のグループにおいて真である。したがって、問題は、物理主義が真であるような可能世界のグループを、すべての可能世界の中から適切に切り出すことである。

### 性質二元論的物理主義

さて以下では、そのようなグループの切り出し方を、個体と性質の（非）同一性という観点から少し描いてみよう。私は、世界の唯物性に関する主張の点でも、また物理的決定性に関する主張の点でも、物理主義が必然的真理であるわけではない、つまり偶然的真理だという出発点は比較的健全だと考える。デカルト的実体二元論の成り立つ可能世界も、性質二元論的可能世界も、性質一元論的可能世界も、ともに論理空間のどこかに棲み分けている。つまり、それらの主張が互いに論理的に矛盾しないような解釈が可能である。しかし、ここでは実体や性質の組み合わせに関するあらゆる論理的可能性を議論できるわけではないので、これまでの物理主義の主張にとって関連のあるいくつかの可能性だけを取り上げることにしよう。また議論の単純化のために、実体つまり個体と性質に関して、それらのタイプを、一元か二元かの場合のみで考えることにしよう。

まず、実体（個体）と性質を可能世界における〈同一性の必然性〉の観点から眺めてみよう。もしある個体  $a$  が個体  $b$  と同一であるなら、 $a = b$  はあらゆる可能世界において成り立つ、とされるのが可能世界論者の間では一般的である。つまり、その同一性は端的に必然的だ。私もここで、クリプキ流のその同一性の解釈にしたがう。そこで、実体（個体）一元論的物理主義の主張を、「現実世界におけるいかなる心的実体（個体）も何らかの物理的実体と同一である」というものだとしよう。すると仮にこの主張が真であるなら、現実世界の心的個体がどんな物理的個体でもないような実体二元論的可能世界は存在しない。なぜなら、現実世界において何らかの物理的個体と同一なら、その心的個体はあらゆる可能世界でその物理的個体と同一であり、したがって、それがいかなる物理的個体とも同一でないような可能世界は存在しないからである。

しかし同時に、実体一元論的物理主義が現実世界を含むある可能世界群で成り立つにすぎないローカルな真理であるなら、どこかの可能世界群には、現実世界には存在しない心的個体が少なくとも一つ存在し、それはいかなる物理的個体とも同一でない、ということではなければならない。そのような心的個体をルイスの用語を借りて、現実世界にとって「異世界的 alien」と呼ぶことにしよう（Cf. [Lewis, 1983, p. 211]）。つまり、いかなる物理的個体とも同一でないような異世界的な心的個体が存在せず、すべての可能世界に渡って、心的個体がすべて何らかの物理的個体と同一であるなら、実体一元論的物理主義は必然的真理となるだろう。物理的個体と同一でない霊や天使という概念に嫌悪を覚える物理主義者は、この選択肢を積極的に取り、このタイプの物理主義を形而上学的に必然的な真理として掲げるかもしれない。

この事情は、実体二元論の場合も同様である。この主張が現実世界を含むある可能世界群で成り立つローカルな真理でしかないなら、〈（非）同一性の必然性〉により、どこかの可能世界には現実世界に存在しない異世界的な心的個体が存在し、それが何らかの物理的個体と同一である、と言えなければならない。というのも、この場合、実体二元論は偶然的真理であるがゆえに、どこかの可能世界では偽でなければならないからだ。逆に、実体二元論を端的に必然的な真理として主張したい人は、すべての可能世界におけるいかなる心的個体も物理的個体と

同一でない、と主張していることになる。

また仮に、現実世界において、物理的個体と同一である心的個体も、またそうでない心的個体もともに存在するなら、現実世界においては、哲学者がふつう主張するような実体二元論も実体一元論もともに偽であるだろう。ここでは、実体二元論を、少なくとも「現実世界におけるいかなる心的個体も物理的個体と同一でない」と主張する立場だと解している。

さて、私としては、物理的実体一元論くらいは、この現実世界において現に真であり、しかも偶然的に真であると主張したい。すなわち、現実世界を含むある可能世界群においては、いかなる心的個体も何らかの物理的個体と同一であるが、この可能世界群にとって異世界的な心的個体が、それら以外の可能世界において存在し、それらは、いかなる物理的個体とも同一ではない。つまり、お望みならこう言おう。われわれの属する可能世界群の外側には、いわゆる霊や天使などのような霊的実体が住む可能世界が存在する。それを認めることはそれほど危険だろうか？ あるいは不健全だろうか？ それが不健全だとすれば、それは可能世界概念の不健全さとちょうど同じ程度だと私には思われる。つまり私には、例えば霊的存在がいかなる物理的個体とも同一でなく存在するということには、論理的矛盾も概念的矛盾も今のところは見いだせないのだ。

いずれにせよ、実体（個体）に関する一元論 vs. 二元論の争いは現在のところ収束状態だと言っていいだろう。現実世界における個体に関して＜唯物性の主張＞が成り立つであろう、ということはいまやほぼ認められていると思われる。むしろ問題は、性質一元論、もしくは性質二元論的物理主義である。そこで、性質も、個体と同様の＜（非）同一性の必然性＞の論理に服する存在だとしよう。

先ほどと同様に、性質一元論的物理主義が現実世界を含むある可能世界群で成り立つローカルな真理であるなら、どこかの可能世界では、現実世界で例化しない異世界的な心的性質が例化し、それはいかなる物理的性質の例化とも同一であってはならない。逆にもしそのような異世界的な心的性質がいかなる可能世界においても例化せず、すべての可能世界に存在する心的性質が、現実世界で例化する性質だけで尽くされているなら、現実世界で成り立つ性質一元論的物理主義は、実は必然的真理であったということになるだろう。

また、性質二元論的物理主義が現実世界を含むある可能世界群で成り立つローカルな真理でしかないなら、どこかの可能世界には、現実世界で例化しない異世界的な心的性質が例化し、それが何らかの物理的性質の例化と同一である、と言えなければならない。また仮に、現実世界において、物理的性質と同一である心的性質も、またそうでない性質もともに例化しているなら、現実世界においては、哲学者がふつう主張するような性質二元論も性質一元論もともに偽であるだろう。

一元論も二元論もともに真であるとしても偶然的真理であるべきだという直観がどれほど尊重されるべきかは、にわかには見積もることができない。しかし、性質同一性の主張でなければ物理主義ではない、といういわゆる還元的タイプの物理主義解釈もまた、それほど強い説得力をもっているわけではないのだ。物理主義の主張が偶然的真理であるか必然的真理であるかを分かつ分水嶺が、現実世界に存在しない異世界的個体あるいは異世界的性質の存在とその有り様だということ、それらが限定的とはいえ必然性の主張を伴っているからに他ならない。とすれば、仮に一元論と二元論のどちらかが現実世界で真だとしても、それが偶然的真理と成り立っているのかあるいは必然的真理として成り立っているのかは、現実世界に関するいかなる探求によっても判明しないことになるだろう。とくに、この現実世界が性質一元論的世界なのか性質二元論的世界なのかは、もし

心的性質が現実世界を含むある範囲の可能世界群において物理的性質にスーパーヴィーンするなら、いかなる経験的／科学的探求によっても決着がつけられないだろう。二つの性質の（限定的かもしれない）「必然的」連動関係が、同一性によるものなのかスーパーヴィーンエンスによるものなのかは、われわれには見分けることができないからだ。

#### 可能世界群の棲み分けとゾンビ世界

さて、もし物理主義が以上のような意味で偶然的真理であるなら、とくに性質二元論的物理主義の主張は、現実世界とそれを含むある範囲の可能世界群においては、心的性質と物理的性質という二つの異なる性質が存在し、しかも前者は後者にスーパーヴィーンする、というテーゼだと言うことができよう。言いかえると、それは、個体に関する唯物性のテーゼと性質に関する物理的決定性のテーゼの二つを偶然的真理として主張する最弱の物理主義である。では、それを、可能世界という論理空間と、性質に関する可能世界の配置から描いてみよう。まず第一に、あらゆる可能世界におけるすべての物理的および心的性質の存在に関する主張としての性質一元論から始めよう。それが成立すれば、性質一元論は必然的真理だ。すなわち、あらゆる可能世界において、いかなる心的性質も何らかの物理的性質と同一である。そして、この種の性質一元論の否定の主張もまた、すべての可能世界に渡るあらゆる物理的および心的性質の存在に関する主張という意味では、必然的主張であろう。つまり、「ニューヨークが金沢より北に位置するのは必然的だというわけではない」があらゆる可能世界において成り立つのと同じ理屈である。私が主張したい偶然的真理（ローカルな真理）としての性質二元論は、この非性質一元論が成立するための可能世界の配置として描くことができる。つまりあらゆる可能世界から成る論理空間の中では、次のような可能世界群が相互に棲み分けているだろう。誤解のないように急いでここでもう繰り返しておけば、実体多元論や性質多元論などのようなあらゆる論理的可能性が以下で尽くされているわけではない。（図を参照）。

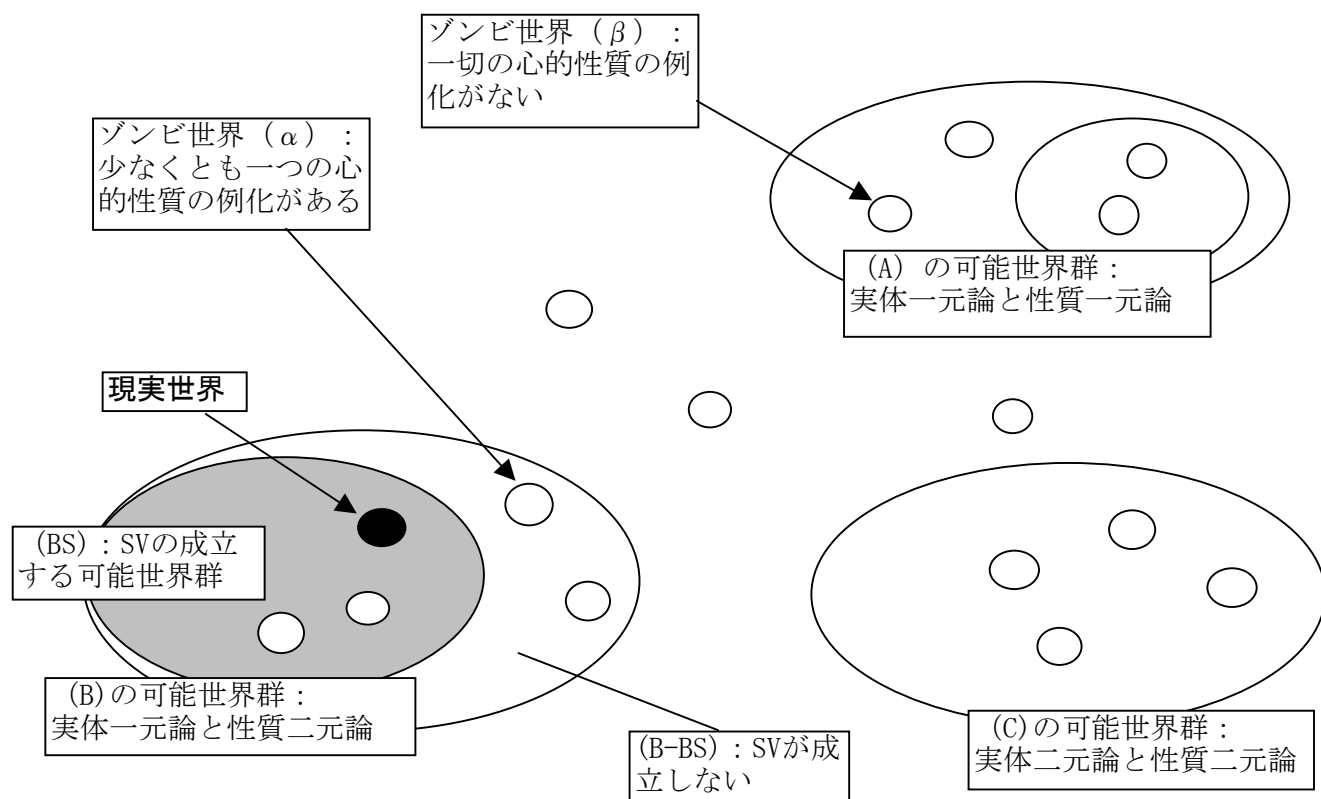
(A)実体一元論と性質一元論が成り立つ可能世界群。その中には、物理的な個体と物理的な性質のみ、および心的個体と心的性質のみが存在する可能世界群があるだろう。

(B)実体一元論と性質二元論が成り立つ可能世界群。それらの世界では、例えば、いかなる心的個体も物理的個体だが、いかなる心的性質も物理的性質と同一ではない。それに対し、すべての個体が心的個体であり、それらがもついかなる物理的性質も心的性質と同一ではないような可能世界がどのようなものになるのかは、読者の想像にお任せしよう。

(C)実体二元論と性質二元論が成り立つ可能世界群。これらの可能世界群においては、例えば霊的実体がいかなる物理的個体とも同一ではないものとして存在し、なおかつ、それらが例化する心的性質はいかなる物理的性質とも同一ではない。またいかなる物理的個体も心的個体と同一であることなく存在し、それらが例化する物理的性質はいかなる心的性質とも同一ではない。



## 性質二元論的物理主義の視点から見た現実世界の位置



私が念頭においているのは、(B)の可能世界群であり、この内部には、さらに、心的性質が物理的性質にスーパーヴィーンする可能世界群 (BS) と、そのスーパーヴィーニエンスが成り立たない可能世界群 (B-BS) を区別することができる。正確に言えば、私が提案している性質二元論的物理主義は、現実世界がこの (BS) の可能世界群に属する、という主張である。この可能世界群においては、先に述べた意味での最弱の物理主義が成り立つ。したがって、チャルマーズの物理主義解釈をとる必要はない。つまり、ゾンビ世界の存在を可能世界群 (B-BS) において認めたとしてもなお、物理主義は現実世界において真である。それどころか、性質二元論的物理主義をこのような意味で偶然的真理として主張することは、性質の依存的連動関係に関しては、どこかの可能世界において心的性質と物理的性質のスーパーヴィーニエンスが破れていることを認めることに他ならないのだから、ゾンビ世界の存在を認めることを含意するだろう (5)。

これは物理主義者にとってそれほど嫌悪すべきことだろうか。そもそも性質一元論を主張する物理主義者の動機は、心的性質に因果的効力を確保し、それによってエピフェノメナリズムを避けることであつた。このエピフォビアから情緒的な動機を消し去ることができれば、性質一元論をくがむしやりに主張しようとする動機はなくなるだろう。そして、性質二元論の健全な形が偶然的真理であるなら、むしろ物理主義者の取り組むべき課題は、エピフェノメナリズムを受け入れた後の倫理的責任に関する議論の方向にあるというべきである。もっとも、

この問題はここではこれ以上深入りすることはできない。

さてもう一度くり返すと、私の主張は、先の意味での性質二元論的物理主義が現実世界を含むある可能世界群(BS)において真だ、ということである。現実世界には、物理的実体と同一でない心的実体は存在しない。しかし物理的性質と異なるものとしての心的性質が例化し、後者は前者にスーパーヴィーンする。先にも述べたように、これは、個体に関する唯物性の主張と、世界を構成する基本性質である物理的性質へのスーパーヴィーンエンスの主張から成っている。私は、この主張をこれ以上うまくどう表現すべきなのかがまだ分からないが、以下のルイスの定式化はその助けになるかもしれない。ルイスは、非物理主義的世界は物理主義的世界に欠けている余分な何かを含んでいるというアイデアと、スーパーヴィーンエンスとを結びつけて、次の物理主義（唯物論）のテーゼ（M5）を提出した。それは、「制限的かつ偶然的であるようなスーパーヴィーンエンステーゼ」である。

M5： われわれの世界にとって異世界的な自然的性質が例化されないような諸世界のあいだでは、いかなる二つの世界も、物理的に異なることなくして、異なることはない。物理的に正確にそっくりであるそのような任意の二つの世界は、複製である[Lewis, 1983, p.212]

このテーゼの下で、ゾンビ世界はどうなるのだろうか。ゾンビ世界は、われわれの現実世界にとって異世界的な自然的性質を例化させていない。それどころか、むしろ、われわれの世界の自然的性質の一部（意識やクオリア）を欠いている。したがって問題の諸世界の仲間に入る前提は満たされているが、しかし、現実世界と物理的に異ならないにもかかわらず、意識やクオリアを欠くという点で現実世界と異なっている。つまり、ゾンビ世界では、M5で主張されているスーパーヴィーンエンスは成立していないのだ。したがって、M5はゾンビ世界では偽であり、ゾンビ世界は、現実世界を含めたわれわれの性質二元論的世界群には属さない。これは、われわれにとって歓迎すべき結果である。

だがもちろん、これで問題が解決したわけではない。われわれが以上のような性質二元論的物理主義を表現するための満足すべきテーゼをいかに工夫しようとも、なぜそのような主張をなそうとするのか、という根本的な動機の説明が残るからである。それは言いかえると、物理主義者を動かしている直観の分析を与える必要がある、ということだ。われわれはなぜ、（性質二元論的）物理主義的に可能な世界からゾンビ世界を排除しようとするのだろうか。その答えは、物理主義はどのようなテーゼとして主張されるべきか、ということの説明を与えることでもある。

### 第3節 現象的判断のパラドックス

#### パラドックスの根

物理主義に関する先の提案は、こうであった。われわれの現実世界が一つの可能世界群(BS)に属し、そのグループの中では、物理的性質への心的性質のスーパーヴィーンエンスが成立するが、その外部の可能世界群(B-B)では、たとえそのグループに属するのと同じタイプの物理的性質が例化されても、対応する心的性質は例化されないことがある。その極端なケースでは、現実世界とまったく同一の物理的世界であるにもかかわらず、心的性質としては現実世界とまったく異なるクオリアが例化されている逆転クオリアの世界や、クオリアや意識が一部にお

いて例化されないようなゾンビ世界が存在する。

しかし、可能世界という論理空間がたまたま前節でのように切り分けられるとしても、さらにそれに加えて、現実世界がたまたま性質二元論的な物理主義的可能世界のグループ(BS)に属しているということに、何かわれわれ自身を納得させる理由があるのだろうか。その理由があるとしても、それは、可能世界全体における可能世界の配置にかかる〈純粹に〉存在論的な理由ではないだろう。というのも、ただそれだけのことなら、現実世界はその論理空間の中のどの位置にあってもよかったからである。そしてまた、恐らくは、その理由は概念的なものを含むとしても、概念分析によって決着のつくようなものではないだろう。なぜなら、もしそのような概念上の理由がはっきりとつけられるならば、かくも長きにわたった唯物論と観念論の争いのようなものは起きなかつただろうからである。しかしながら逆に、その理由は、経験の中だけに見出される科学的な理由だけでもないだろう。というのも、経験的な理論だけでは決着のつかない選択肢が残るからこそ、この種の存在論的な問題は、すぐれて宗教や哲学の問題だったのであるから。

したがって、現実世界が性質二元論的な可能世界のグループ(BS)に属しながらも、逆転クオリアやゾンビの世界とは極めて重要な点で異なる、とする理由は、概念的な理由と経験的な理由の間にあるはずだ。いやむしろ、その両者の要素をすべて含んでいる理由なのだから、こう言うべきかもしれない。その理由は、われわれ人類がこれまでに獲得したすべての〈正しい理由〉(最善の説明)についての〈最善の説明〉となるべきだ、と。さて、そのメタ的な最善の説明の骨格を形成するのは何だろうか? 最弱の物理主義者として私が提案できるものは、究極的には物理学が与える〈因果的ストーリー〉との整合性である(6)。つまり、現実世界を可能世界群(BS)の中に数え入れ、逆転クオリアやゾンビ世界を現実世界から異質なものとして遠ざける理由は、後者の世界では、その因果的ストーリーとの整合性が破れている、ということだ。その破れは論理的矛盾ほどあからさまな異常ではないが、しかし、ふつうの経験的なストーリー内部の軋轢とは比べものにならないほどに極端な異常である。私は、この整合性の破れに関して、逆転クオリアを含んだクオリアのズレについては以前に論じたことがある(Cf. [柴田, 2003, pp. 54-6]および[柴田, 2001, pp. 192-211])。したがって、以下では、ゾンビ世界の異常さに関して述べることで、現実世界がどのような意味で可能世界群(B-B-S)には属さないかを示すことにしよう。

ゾンビ世界の異常さは、チャルマーズにとって、物理主義的な還元主義者や消去主義者から論理的矛盾として攻撃されるのを防いでやるべき対象である。論理的矛盾ではなく、むしろ意識という所与の事実が持つ驚くべき不思議さとして、それを彼は、現象的判断のパラドックスと呼ぶ。

・・・意識はわれわれの意識についての主張や判断を説明する上で関与してこない[Charmers, 1996, p. 177]。

このパラドックスの内容は、もう少し詳しく次のように言われる。

(1)物理的領域は因果関係に関して閉じている。(2)意識に関する判断は物理的なものに論理的にスーパーヴィーンしている。(3)意識は物理的なものに論理的にスーパーヴィーンしていない。(4)われわれはわれわれに意識があることを知っている。前提(1)と(2)から、意識に関する判断は還元によって説明できることになる。これが前提(3)と結びついて、意識はわれわれの判断に関する説明に関与しな

いことを意味し、それは前提(4)と緊張関係に立つ。こうして、われわれはパラドックスを手にする[ibid., p. 183]。

私と分子レベルまで物理的に同一のゾンビたちを考えてみよう。彼はこのパラドックスで言われているように、私とまったく同じ現象的判断をする。とくに、チャルマーズの区別する三層の判断レベル、「それは愉快だ」という一次判断、「私はいま、愉快だという感じがしている」という二次判断、さらに「感じというのは不思議だ」という三次判断に至るまで (ibid., p. 176)、私と私の双子ゾンビは同じ判断を同じ状況で形成する。ここからは、極めて興味深い、しかし反直観的な、つまり先に述べた意味でわれわれの描く通常の因果的ストーリーと整合的でない事柄がいくつか帰結する。それらは、チャルマーズが弁明するように論理的矛盾ではない。しかし、それらは、ゾンビ世界を現実世界から決定的に遠ざけるのに十分なほど理解しがたいのだ。

### ゾンビによる現象的判断

まず、判断形成のメカニズムが因果的な機能によって完全に果たされることを確認しよう。つまり、私と私の双子ゾンビは物理的に区別できない複製であるから、もちろん機能的にも区別できない複製、言うなれば完璧に「肌理の細かい」機能同型体である。私がボードの上に娘の受験番号を見分けたとき、現実世界と物理的に区別できないゾンビ世界の私の双子ゾンビも、物理的に同じボードの上に物理的に同じインクで印刷された受験番号を見分ける。その時、第三人称的に見れば、私の知覚メカニズム、信念形成のメカニズム、判断形成のメカニズムと同一のメカニズムが私の双子ゾンビの中でも働いている。ここまではいい。しかし、問題は、現象的判断の場合、その際の判断の対象となっているのが、私には存在し、私の双子ゾンビには存在しない意識体験、例えば、愉快なあの感じなのだ。ところが、その双子ゾンビは、「どんな感じ？」と問われて、私とまったく同じように、「なんだか、ふわっとした、くすぐったいような」と答える。実際は、彼は何も感じていないにもかかわらず。

したがって、自分の経験するクオリアがどのようなものであるかの判断には、クオリアの存在は不要であるように思われる。これは、クオリアを経験している私の場合もそうだということになるだろう。つまり、微細な感情の起伏といった現象的経験に関する判断であれ、そしてそれが私自身の判断であれ、その判断の対象であるはずの当のクオリアの存在すら必要がないのだ。これはすなわち、一見して、クオリアや意識に関する消去主義を含意する事態であるように見える。しかしながら、ゾンビ世界の想定は、ゾンビはいざしらず、私自身はクオリアを疑いもなく持っているというのが前提だ。したがって、この事態がこの前提に対する背理法を形成するとまでは解釈されないとしても、きわめて不愉快な緊張がここにあるのは確かである。

またさらに、たとえ消去主義が帰結しなくとも、この想定は、現象的経験をもつわれわれ自身の場合ですら、現象的判断が下されるときに内省的報告の対象になっているのは、実はクオリアではなく、例えばニューロンの発火パターンや脳の感情メカニズムの活動だ、というある種の還元主義を招き寄せるだろう。つまりこの立場では、ちょうど、クオリアをもたないロボットが自分の内部メカニズムや内部状態を体内センサーによって把握し、それを報告するのと同じことを、われわれ自身も、われわれのゾンビもやっているのに他ならないということになる。だがチャルマーズが言うように、ロボットのようない認知システムが「なぜ自分がいま悲しいと分かるんだ？」と聞かれて、「しかじかの仕方で内部センサー1

6-57 が活性化されているから、自分が悲しいんだということが分かります」と答えるようでは、自らの感情内容に間接的にしかアクセスできない「役立たずの不自然な」認知システムだ、ということになるだろう。まともなロボットなら、「なぜって、本当に悲しいんだから・・・」と答えるはずだ (Cf. [ibid., pp.185-6])。したがって、認知システムが自分の内部状態を知るとき、それが<経験の主体>と言える程度に高度のものであるならば、そのシステムは、第三人称的に自らの内部メカニズムを知るのではなくて、第一人称的にそれを知っているはずである。そうでなければ、それは一つの自律した認知システムではないだろう。しかし、第一人称的に自らの内部メカニズムを知ること、それこそまさしく意識やクオリアを内容とした現象的経験をもつことに他ならない。

したがって、私の双子ゾンビは私と同じように、「いま自分が悲しいのはウソでも何でもない、悲しいから悲しいんだ」と言ったとしても、彼は、ある種の神経興奮の第一人称的な内容を第一人称的に報告しているのではない。その神経興奮の第三人称的に観察可能で特定化可能なメカニズムが、その発言を引き起こしているだけなのだ。なぜか。彼は心的性質を一切欠いているからである。それを一切欠いているにもかかわらず、彼はそれを持っているわれわれと同じ答えをすることができる。これこそがパラドックスの正体だ。しかし、心的性質を一切欠いた個体は、もはや心的個体ではない。そして、この意味で、心的性質は認知システムが<他ではない自己>として経験をもつための必要条件であるから、心的個体でないものは、われわれが理解する意味での経験の主体でもないだろう。言いかえると、そのような個体は、少なくとも第一人称的な経験の把握、例えば「そのロボットであるとはどのようなことか？」ということに答える内容をもたない限り、経験の主体ではないということだ。したがって、ゾンビの想定は、ある種の還元主義を招き寄せると同時に、私と同一の経験の報告をなしながらも、そもそも経験の主体ではないという極めて受け入れがたい帰結をもつ。それは、認知システムとその機能に結びついている<経験の主体>というわれわれの概念が、もはやゾンビの想定を矛盾したものに見せるほど朽ち果てている、ということである。

言いかえると、ゾンビは、程度の低いロボットのように、現象的判断を迫られて自分の人体や脳を第三人称的に内部スキャンして、その結果を報告するわけではない。私が、自分の経験している感情クオリアを報告するのに自分の身体や脳をこれまで一度も第三人称的に内部スキャンしなくとも、またそのようなことができなくとも、それについての現象的判断を下すことができるように、私の双子ゾンビも、彼の現象的判断においてそのようなことをやっているわけではない。彼の内部でも、私の第一人称的経験に対応する同じ脳プロセスが生じ、現象的判断に対応する同じ認知メカニズムが働いている。しかし私の場合には、問題の内部状態に感情クオリアがスーパーヴィーンし、それが私の現象的判断の内容となる。しかし、ゾンビにはそのような感情クオリアが一切存在しない。それにもかかわらず、なぜ彼は私と寸分変わらない判断、例えば「晴れ晴れとするようなこの浮き立つ感じ」といったことを報告できるのか。私は心的個体でもあり、それが有する心的性質として、この判断内容を経験する。私の双子ゾンビはもはや心的個体ではなく、第一人称的経験を一切もたないにもかかわらず、存在しないはずのその第一人称的経験に対して、なぜ私と同じ現象的判断をなしうるのか。その符合は奇跡ではないか。確かに、ゾンビ世界とはそういう奇跡の世界なのだ、という説明しかもはや残されていないように思われる。しかしそれはただ、ゾンビ世界が論理的に不可能なわけではない、ということを使い換えたにすぎない。

このようなゾンビの想定が、論理的可能性というきわめて<薄い>内容以上の

ものを持たないということを見るために、さらに、われわれ自らが「ゾンビであるとはどのようなことか？」と問うてみよう。われわれは、もしゾンビになったとしても依然とまったく変わらない経験の報告をし、現象的判断を行い、他のゾンビたちもまた、まったく現実世界と同様にそれに反応するだろう。しかし私は、ゾンビであるのだから、実際は何も経験しない。つまり、私は機能的には「悲しみに沈んでいる」という状態にあるが、実際には何も感じていない。というのも、悲しみを感ずるということは、少なくとも現象的经验としては、それに至る微妙なニュアンスの感情の積み重ねを経験することを含むが、それらのクオリアが一切私に欠けているからである。それは、他の感覚様相に関しても同じだ。世界のさまざまな色や音や香りや味を、私はすべて機能的には「見、聞き、嗅ぎ、味わっている」が、しかし現象的经验としては何も感じていない。それどころか、先に述べたように、私はそもそも経験の主体ですらない。チャルマーズが言うように、私と私の双子ゾンビとでは、とてつもない違いがある。私は経験し、彼は経験しない[ibid., p.199]。したがって、「ゾンビであるとはどのようなことか？」と問われたら、ゾンビである私がどう発言しようとも、その真の答えは<無>である。私なら答えうるような第一人称的经验は、彼には何もない。だから、何も答えることができないのだ。ある意味では、<私>はもはや存在しないのである。

あなただけがゾンビであることを免れたゾンビ世界では、現象的经验の内容に関して他人が語ることはすべて偽である。偽である発言がその世界には充ち満ちている。しかし、そのことは、その世界にはあなた以外、実は<誰>一人として本当は存在しない、ということに比べたら大して不気味ではないだろう。だが不気味さはともかく、このようなゾンビの想定は、経験の主体を構成するものとしてのみ理解可能なすべての認知的ストーリーが、世界の实在とかみ合っていないということの意味している。その巨大なスケールでのズレは、論理的矛盾ではなくとも、逆転クオリアの場合以上にわれわれの世界理解にとって異質なものである。

このことを、「水槽の中の脳」のパトナムにならってこう述べてもよい。「私はゾンビである」は、自己挫折的な言明だ[Putnam, 1981, p.7, p.15]、と。なぜなら、この言明が真なら、この言明の本来の発言主体である<私>は存在しないからだ。もちろんわれわれは、真なる発言の発言主体の概念には、第一人称的な経験の主体であるという要素が含まれていると考えている。この自己挫折性は、チャルマーズが執拗に擁護するように、その異様さにもかかわらずなお概念的な矛盾ではないだろう。だが、それはもはや、何も包むことのできない、論理的可能性という薄っぺらな布きれにすぎないのは明白である。したがって、ゾンビ世界が現実世界から十分すぎるほど遠いことを、われわれは確信することができる。

ゾンビ世界の異常さは、結局、物理的もしくは認知的ストーリーには<必然的に>現象的经验、つまりクオリアや意識が対応して存在する、ということのをわれわれに説得的に物語っている。それは言いかえると、少なくとも現実世界は、心的性質と物理的性質に関してスーパーヴィーニエンス関係が成り立っている可能世界群に属している、ということの意味している。つまり、以上のゾンビ世界の不都合さは、性質二元論的な可能世界グループの中で、スーパーヴィーニエンスが成り立つ可能世界群(BS)とそれが成り立たない可能世界群(B-BS)を分かつ理由を形成するものだったのである。

世界がクオリアを授ける

翻ってみれば、チャルマーズの述べた現象的判断のパラドックスは、やや形を

変えた形で、心的因果性の議論においてすでに提出されたものである。そして、この問題に関して、私は、物理主義的エピフェノメナリズムこそが唯一整合的な解答だと考えている(Cf. [柴田、2004], [柴田、2006])。その地点から眺めるなら、現象的判断のパラドックスはパラドックスでも何でも無い。ただ、現実世界の有意味な変容という意味での可能世界群において、問題の二性質間のスーパーヴィーニエンスを認めることを意味するだけである。しかし、そのことには、われわれの主題に関して重要な二つの意味がある。一つは、ある種の粗野な還元主義は成り立たないということであり、もう一つは、とりわけゾンビの議論で重要だった<認知的ストーリー>との整合性を考慮に入れるなら、感情クオリアは感情機能に対して世界が授けるものだけということである。

まず、粗野なタイプの還元主義ということによって私が意味しているのは、認知に関する第三人称的な物理的／機能的ストーリーを描くことで、認知に関するすべてを描いたと考えるような還元主義である。性質に関する存在論的立場をどう取ろうと、認知に関する第三人称的な説明を与えることは、決して第一人称的な経験を持つことと同一ではない。それを同一視することは、単なるカテゴリーミスイクだ。したがって、例えば、「そのロボットであるとはどのようなことか」という経験の第一人称的内容が、そのロボットの認知機能に関する第三人称の説明によっては与えられないということと、それにもかかわらずその第三人称の説明はあらゆる点で科学的に完全だ、ということとは両立可能である。ダマシオにもう一度語ってもらおう。あなたが、サンフランシスコ湾を眺める私の脳を驚異的な新型スキャナーで脳の最深部までスキャンし、私の網膜、外側膝状体核、初期視覚皮質部位、それにさまざまな大脳皮質部位の発火パターンを鮮明に捉え、その上、その物理学的、化学的記述も得たとしよう。すると「・・・たぶんあなたは、私の心の中のイメージ・コンテンツに対する、一連の驚くべき『相関物』を手にするだろう。しかし、あなたに申し上げておこう。そのときあなたは決して、私の、そのイメージの『経験』を手にはしていない、と」[Damasio, 2000, p.\*\* \*]。

このことは、先ほどの議論に登場した<経験の主体>という概念が、決して第三人称的な機能的概念によっても内容を尽くしえないということの意味している。つまりゾンビがいかに<経験の主体>であるかのような錯覚を周囲に生み出そうとも、概念上それは第一人称的な経験を持たないのであるから、言葉の真の意味での<私>にはなりえないわけだ。したがって、先程述べたような粗野な還元主義は退けられるべきである。しかし、だとすると、物理的にしか構成されていないはずのわれわれの一人ひとりが意識やクオリアを持ち、何ごとかを<経験>できるのはいかにして可能なのだろうか。それはますます奇跡のように思われてこないだろうか。

ここで、第二の、そして最後の論点に取りかかろう。それは、意識やクオリアは物理的な存在の上にくでたらめに分配されているのではなく、法則的に分配しているという点に関係する。われわれは、ゾンビの陽炎のような危うさが、世界の物理的ストーリー、とりわけ認知主体の認知的ストーリーとの不整合にあるということを知った。してみると、われわれは、最弱の物理主義の物理的決定テーゼ「同一の物理的性質には同一の心的性質（意識・クオリア）がスーパーヴィーンする」を少しだけ踏み越えて、このスーパーヴィン関係は、機能的性質を通して、またそれを通してのみ成立する、と言っているのではないだろうか。つまり、機能的性質（気づきや感情機能）はさまざまな認知主体と物理的素材において多重実現されるが、あるクオリアが経験の主体に生ずるための必要十分条件はその当の機能的性質の出現であり、それを介して、そのクオリアはそれ

らの物理的素材の物理的性質にスーパーヴィーンするのである。

というのも、例えば、喜びの感情クオリアが喜びの感情機能ではなく、悲しみの感情機能や、恐怖の感情機能などにいわばくでたらめに出現しているということの不整合さが、ゾンビをわれわれの世界から遠ざける根拠だったからである。したがって、ここでその議論をくり返す必要もなく、こう言っていいただろう。喜びのクオリアは、喜びの感情機能が働いている場合に認知主体に経験され、また逆に、喜びのクオリアが経験されている場合には喜びの感情機能が認知主体において働いているのだ、と。しかし、第一人称の経験と第三人称の存在物、つまりクオリアと機能ほどに存在論的に乖離した現象を何がこのように法則的に結びつけるのか。しかし、それは私には、「なぜエネルギーは光速の2乗に比例するのであって、3乗に比例するのではないのか」という問いと同じようなものに聞こえる。私には、われわれの世界がそれを求めているのだ、という以外の答えがあるとは思われない。

なお、本稿は、[柴田，2007]を大幅に加筆修正し、そこで積み残されていた問題に新たな展開を試みたものである。また本稿の一部は、平成19年度～平成22年度科学研究費研究課題「認知ロボティクスの哲学」（代表者・金沢大学・柴田正良）の研究成果を含んでいる。

#### 注

(1) 「ノーは、内なる感覚の視点から生み出される人間の意識に類似したのもをもつ人工物を創造する可能性はほとんどないということ。イエスは、本書で提案した意識の形式的構造をもつ人工物をつくることは可能であるということ。たぶんそのような人工物にはある種の意識があると言えるだろう」[Damasio, 2000, p. \*\*\*]

(2) 「偶然的真理としての物理主義」を日本で明快に掲げた例は、[柏端，2006]くらいしか見当たらない。柏端がそこで依拠している D. ルイス[Lewis, 1983]は、これを最小限の唯物論の定式化の問題として論じている。私自身は、以前に、ゾンビの論理的可能性を認める「ひ弱な物理主義」という形で、われわれの世界に存在する心的性質と物理的性質の二元論として主張したことがある（[柴田，2003]）。そこでは、「ひ弱な物理主義」は、自然法則的に可能な世界の集合のさらに部分集合において成り立つテーゼである。

(3) チャルマーズの議論では、いわゆる機能的性質としての心理学的性質は、現象的性質と区別され、それを構成する物理的性質と同一か、少なくともそれに論理的にスーパーヴィーンするものと見なされている。したがって、チャルマーズのゾンビ世界では、信念や欲求などの心理学的性質は第三人称的な性質としては存在するが、それに対応する現象的性質としての意識やクオリアが欠けているがゆえに、ゾンビ世界の存在が還元主義的な物理主義の反駁となるのである。他方、私は、機能的性質としての心理学的性質すらそれを実現する物理的性質には還元しえないと考えている（Cf. [柴田，2006]）。したがって、私としては少なくとも、物理的性質と現象的性質の性質二元論以上の性質多元論を提案したいところだが、以下では、チャルマーズの枠組みにしたがい、心的性質をもつばら意識やクオリアといった現象的性質に限ったものとして議論を進めることにする。

(4) 物理主義をいかに定義するのもまた、一大事である。さまざまな試みがなされ、そのスタンダードな定義はまだ現れていない。しかし、もし(a)の唯物性の



テーゼが成り立つなら、(b)の決定性のテーゼは不要であろう。決定性のテーゼは、少なくとも世界に存在する性質の一部が物理的でない（例えば、心的と称される性質の一部が物理的性質と同一でない）からこそ、必要とされるテーゼである。例えば、物理主義者キムは、機能的性質としての心的性質はその実現性質としての物理的性質と同一であるがゆえに、スーパーヴィーニエンスの概念は物理主義にとって不要だと考えている（Cf. [Kim, 1998]）。

(5) もっとも、ゾンビ世界がいかなる心的性質の例化も許さないなら（注3で述べたように、機能的性質としての心理学的性質ここでは心的性質とは見なされない）、その世界はむしろ、実体一元論と性質一元論の可能世界群(A)に属するゾンビ世界( $\beta$ )であろう。それに対し、ゾンビ世界( $\alpha$ )は、少なくとも一人は意識やクオリアを経験する個体を含んでいるはずである。ただし、性質一元論的物理主義から見れば、現実世界は(A)の可能世界群の中にあり、それはまさにゾンビ世界( $\beta$ )に他ならない。というのも、性質一元論によれば、現実世界のすべての心的性質は何らかの物理的性質と同一であり、したがって、当の物理的性質の例化が存在する限り、チャルマーズのゾンビ世界には現実世界と同一の心的性質の例化が存在するからである（それゆえ、ゾンビ世界( $\beta$ ) = 現実世界）。

(6) <最善の説明についての最善の説明>について、私は、[柴田, 2003, 注(4)]において少しだけ触れたことがある。

なお、翻訳が存在する場合はできるだけそれに従ったが、一部、勝手に訳文を変えたところもある。ここで、それぞれの訳者に感謝の意を表したい。

## 参考文献

柏端達也 (2006) 「選言化する心と二元論的世界」『思想』第 982 号、岩波書店、16-34 頁

柴田正良 (2001) 『ロボットの心』講談社現代新書

----- (2003) 「ゾンビは論理的可能性ですらないか？ ----- チャルマーズに対する pros and cons -----」『コネクショニズムの哲学的意義の研究』（平成 12 ~ 14 年度科学研究費研究成果報告書：研究課題番号：12410003、代表者：南山大学人文学部教授・服部裕幸）、47-66 頁。

----- (2004) 「The Exclusion Problem とエピフェノメナリズム」『理想』No.672、69-82 頁

----- (2006) 「機能的性質と心的因果 ----- キムの還元主義を越えて -----」『思想』第 982 号、53-76 頁

----- (2007) 「人工のクオリアとゾンビ世界」『意識と感情をもつ認知システムについての哲学的研究』（平成 16 ~ 18 年度科学研究費研究成果報告書：研究課題番号：16320003、代表者：金沢大学文学部教授・柴田正良）、44-57 頁。

Chalmers, D. J. (1996), *The Conscious Mind*, Oxford University Press, Oxford. (林一訳『意識する心』白揚社、2006)

Damasio, A. R. (2000), *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, Vintage; New Ed. (『無意識の脳 自己意識の脳 ----- 身体と情動と感情の神秘』田中三彦訳、講談社)

- Kim, J. (1998), *Mind in a Physical World*, The MIT Press, Cambridge, Massachusetts.
- Lewis, D. (1983), “New Work for a Theory of Universals”, in [Mellor, 1997], 188-227. (「普遍者の理論のための新しい仕事」柏端達也・青山択央・谷川卓訳『現代形而上学論文集』勁草書房、2006)
- Mellor, D. H., and A. Oliver(1997), *Properties*, Oxford Readings in Philosophy, Oxford University Press, Oxford.
- Putnam, H. (1981), *Reason, Truth and History*, Cambridge University Press, Cambridge. (野本和幸・中川大・三上勝生・金子洋之訳『理性・真理・歴史』法政大学出版局、1994)
- Stalnaker, R. C., 1996, “Varieties of Supervenience” in [Stalnaker, 2003].
- (2003), *Ways a World Might be*, Oxford University Press, Oxford.