A Philosophical Study of Cognitive Systems with Consciousness and Emotions

メタデータ	言語: jpn
	出版者:
	公開日: 2017-10-05
	キーワード (Ja):
	キーワード (En):
	作成者: Shibata, Masayoshi
	メールアドレス:
	所属:
URL	https://doi.org/10.24517/00034737

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 International License.

人工のクオリアとゾンビ世界

柴田正良（金沢大）

序

本科学研究費の研究は、最終的には、意識とクオリアに関するいわゆるハードプロブレムの概念的解決を目指すものである。しかし、今回の研究は、クオリアに関するそうした問題までは射程に収めていない。われわれが最終的にターゲットにしたい問いは、「悲しく感じられるとはまさにそれであるような、そのそれとは何か」といった感情の主観的・現象的な経験の質である。しかし、われわれが行ったことは、客観的な因果過程の中で第三人称的に観察可能な感情の機能を、しかも人工的に作成するということである(1)。もしも現象を理解するためにその現象を作るつもりなら、感情の機能ではなく感情のクオリアをこそ作るべきだったろう。しかし、われわれはそうはしなかった。というよりも、もちろん、そうできなかったのである。なぜなら、われわれの誰にとっても、クオリアを人工的に作り出すことなど現時点ではとうてい出来ないからである。それはおよそ何かバカげたことのように思われる。しかし、なぜクオリアを人工的に作り出すことはバカげたことのように思えるのだろうか。クオリアの人工的な生成にはしかるべき機能の実現で十分だ、と主張することになぜ深い困難を感じるのだろうか。そこにあるのは何らかの概念的不可能性なのだろうか。

本論では、われわれの研究が実際には扱えなかったクオリアの概念的問題の一部を扱う。それは、偶然性のテーゼ（ローカルな真理）としての物理主義(2)と、チャルマーズのいわゆるゾンビの論理的可能性である。以下では、ゾンビの論理的可能性と性質二元論の概略を確認し（1 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（2 節）。そこから、機能的性質と、クオリアを含めた現象的性質の違いについてなにほどかの考察を引き出したいと思う。それは、うまくすれば、クオリアの人工的生成とはいかなることであるかをわれわれにイメージさせてくれるだろう。

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。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わけだ。ところで、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なんだろうか。

「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)

大ざっぱにいえば、この論理的可能性は概念的な想像可能性(conceivability)に対応する。したがって、チャルマーズによれば、「雄の牝狐 male vixen」という概念は矛盾をはらむがゆえに論理的に不可能であるのに対し、「飛ぶ電話機 flying telephone」は整合的な概念であるがゆえに論理的に可能である。神は、後者が存在する世界を作りえたであろうが、前者が存在する世界を作ることはできない。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性質をもつど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、これも大ざっぱにいえば、実際の経験的可能性に対応する。チャルマーズによれば、高さ1マイルの高層ビルを建てることは論理的にも自然的にも可能だが、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。

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以外は存在しない」、「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」と表現されるようなものだからである。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。

チャルマーズの論点をもう一度まとめよう。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、物理主義は誤りである。もっとも、ここで急いでつけ加えておけば、チャルマーズは、物理的事実への意識経験の論理的スーパーヴィーニエンスは認めないが、自然的スーパーヴィーニエンスは認める。

1. 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

さて、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、そうする必要があるのだろうか。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、霊や天使が物理的実体と共に「うようよしている」（後述の Lweis [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であろう（私もそこまで譲歩する気はない）。しかし、チャルマーズの言い方を借りて言えば、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ったのだろうか。私にはそうは思われない。

いささか想像力の乱用だと思われるかもしれないが、次のような場合はどうだろう。実体に関してはそもそも心的（霊的）実体しか存在しない、という世界は不可能だろうか。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世界は不可能だろうか。その世界と物理的に区別のつかないもう一つの世界、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、今度は別の心的実体どもが存在するといったことは不可能なのだろうか。もし不可能だというなら、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はもっているだろう。

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世界も存在するだろう。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。

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、また物理的なものによる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（ここでは、実体や性質のタイプを、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、それはそれらの世界におけるいかなる物理的個体とも同一でない、ということでなければならない。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. Leiws [1983], p.211, 邦訳、p.184）。つまり、いかなる物理的個体とも同一でないよう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（あるいは逆に、この選択肢を積極的に取り、この物理主義を形而上学的に必然的な真理として掲げたいという人もいるかもしれない）。

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、これ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、例えば霊的存在だけが個体として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＜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べきなら、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。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心的性質が例化し、それがその世界で例化する何らかの異世界的な物理的性質と同一である、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は、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、そのどちらがこの現実世界で成り立っているかは、現実世界に関するいかなる経験的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ニ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。

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。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性質の例化を考えよう。つまり、あらゆる可能世界におけるすべての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主張である。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、性質一元論の否定という意味での主張もまた、すべての可能世界に渡るあらゆる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、「すべての自然数が３の倍数であるわけではない」があらゆる可能世界において成り立つのと同様である。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。

(A)実体一元論と性質一元論が成り立つ可能世界群。例えば、それは、物理的な個体と物理的な性質のみが存在する可能世界群である。

(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。

(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。

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(3)。この可能世界群においては、心的個体に関しては唯物性の主張が成り立ち、かつまた同時に、問題のスーパーヴィーニエンスが成立しているがゆえに物理的性質による決定性も成り立つ。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。

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った。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<がむしゃらに>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（仮に真だとして）、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できないが。

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、実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ニ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、スーパーヴィーニ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ニエンステーゼ」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。（Lweis [1983], p.212、邦訳、p.184-5）

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ニ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。それは、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。

2. 現象的判断のパラドックス

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ニエンスが成立するが、その外部の可能世界群（B-BS)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ズレたクオリアが好き勝手に例化されている世界や、クオリアや意識がまったく（もしくは、「まだら」にしか）例化されていない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＜純粋に＞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っ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？　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＜因果的なストーリー＞との整合性である(4)。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、ということである。しかし、その破れは論理的矛盾ほど明白な異常ではないが、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のタイプについては論じたことがある(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-BS)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)と緊張関係にある。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）、私と私の双生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な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である。

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、私と私の双生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生ゾンビも、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生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、私の双生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、その双生ゾンビは、「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持っているというのが前提である。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、例えば神経の興奮であったり、脳の感情メカニズムの活動である、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、ロボットのような認知システムが「なぜそこにチョコレートがあると分かるのか」と聞かれて、「しかじかの仕方でセンサー78-84 が活性化されているから、それを見ていることが分かります」と答えるようでは、自らの知覚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、「そこにチョコレートが本当に見えるんだから・・・」と答えるだろう（ibid., p.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、私の双生ゾンビは私と同じように、「そこにビターのチョコレートが見えるのは本当だよ、見るから見えるんだよ」といったとしても、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。まさに物理的な出来事としての神経興奮を第三人称的に観察して報告しているのである。なぜか。彼は心的性質を一切欠いているからである。そして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、つまり「そのロボットであるとはどのようなことか」ということの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。

このようなゾンビ世界の想定が、論理的可能性というきわめて＜薄い＞内容以上のものをも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以前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「チョコレートを見ている」という状態にあるが、実際には何も＜見えていない＞。というのも、チョコレートが見えるということは、少なくとも現象的経験としては、チョコレートやそれを入れている箱の色クオリアを感ずることを含むが、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、しかし現象的経験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、現実世界の私と私の双生ゾンビとでは、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。私には答えうるような一人称的経験は何もない。したがって、何も答えることができない。ある意味では、＜私＞は存在しないのだ。ゾンビ世界では、現象的経験の内容に関して語られ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なストーリーが、世界の実在とかみ合っていないということを意味している。その巨大なスケールのズレは、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、自己挫折的な言明である（Putnam［1981］p.7、邦訳、p.9-10）、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、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＞現象的経験、つまりクオリアや意識が対応して存在する、ということ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、ということを意味している。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。

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(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ない。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。したがって、そのことはまた、ともかく現実世界においては、クオリアを人工的に生成するにはそれに対応する認知

機能を人工的に作り出すことで十分だ、ということを含意しているはずだが、それについての具体的な問題は、稿を改めて論ずることにしよう。

注

(1)　これに関する報告は、2006年8月に米国アトランタで開催されたInternational Society for Research on Emotions、および同年10月に名古屋で開催された中部哲学会において行われた。

A Primitive Emotion and Its Cooperative Function Simulated in Neural Networks: Toworads a Theory of Emotions as Cognitive Functions」(S. Nagataki, M. Shibata et al.), International Society for Research on Emotions, Annual Meeting, Atlanta, Georgia, August 6-10, USA, 平成18年8月

「感情機能と感情のニューラル・ネットワーク」(柴田正良・月本洋)、2006年 中部哲学会、シンポジウム「情念について」、椙山女学院大学、平成18年10月

(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)　しかし、このままでは、この可能世界群 (BS)の内部には、なおもわれわれの<因果的ストーリー>にとって理解しがたい奇怪な可能世界群が存在する。逆転クオリアの世界はまさにそのような可能世界の一つであり、それを<素面の物理主義>から排除するためには、柴田 [2003], pp.51-54 が行ったような法則性への顧慮が必要である。

(4)　<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、訳者に感謝の意を表したい。

参考文献

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

柏端達也，2006，「選言化する心と二元論的世界」，『思想』第 982 号，pp.16-34.

Lewis, D, 1983, “New Work for a Theory of Universals”, in Mellor [1997]. (「普遍者の理論ための新しい仕事」、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.

Putnam, H., 1981, *Reason, Truth and History*, Cambridge University Press (『理性・真理・歴史』、野本和幸・中川大・三上勝生・金子洋之訳、法政大学出版局、1994)

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）、pp.47-66.

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

-------, 2006, 「機能的性質と心的因果-----キム的還元主義を越えて-----」『思想』No.982.

Stalnaker, R. C., 1996, “Varieties of Supervenience” in Stalnaker [2003].

-------, 2003, *Ways a World Might be*, Oxford University Press.