

## ＜偶然的真理としての物理主義＞から新しい倫理道德の創作へ

平成 28 年 3 月 10 日

柴田正良（金沢大学教育担当理事・副学長）

平成 27 年 4 月に開催された名古屋大学哲学会の発表で、私が主張したことは 2 つである。1 つは、(1) 心の諸現象に関して、われわれの世界（現実世界）においておおむね妥当と思われる存在論的立場は物理主義（physicalism）であろうが、しかしそれはあらゆる可能世界で成立する主張であるわけではない、という意味で必然的真理ではなく、現実世界を含んだローカルな可能世界群における＜偶然的真理＞にすぎないということである。さらにこのことは、個体及び性質に関する同一性の必然性を前提する限り、現実世界における心的性質と物理的性質の関係は、一方から他方への還元による同一性関係ではなく、たかだか一方の他方に対するスーパーヴィーニエンス（supervenience）関係でしかないということである。また、以上を前提とした私の主張の 2 つ目はこうであった。(2) 心的因果（mental causation）の問題において受け入れざるをえない真理は、現実世界でわれわれが経験する心身の性質に関する限り、いわゆるエピフェノメナリズムであろう。心的性質に因果的効力を確保しようとするいかなる試みにもかかわらず、心的性質と物理的性質が非還元的関係であり、現実世界が物理主義的世界である以上、心的性質に、物理的性質とは「異なる独自の」因果的効力を与えることはできない。

(1) に関しても (2) に関しても、問題領域と議論の性質上、残念ながら決定的な論証を与えることができないことも述べておいた。「前提→結論」に関して局所的な議論の精確さを達成することはできても、「前提」をその都度どこまでも遡って疑問に付すことができるのが哲学の常道であるから、これらの問題に限らず多くの領域において哲学がなしうることは、これまでの科学の成果をできるだけ多く取り入れ、なおかつ可能な限り経験と常識を尊重しながら、論理的に整合的な説明を与えることであろう。そして哲学がなしうる最大限のことは、それらの個々の整合的な説明の上に立ち、この現実世界が結局のところいかなるタイプの可能世界であるかについて、最も広範囲で最も整合的なストーリーを組み立てる（一枚の絵を描く）ことである。発表の最後において私は、そのような説明のイメージを与えようと試みた。その際、物理主義にとっての最大の懸念に関して、私は、「エピフェノメナリズムなんて怖くない」、が同時に、「自由と責任と倫理に関する新たな概念を提案するのがこれからの哲学者の使命だ」ということも訴えた。以下では、その発表の概要を簡単に述べることにする。

### 1. 因果的閉包性（causal closure）と過剰決定（overdetermination）

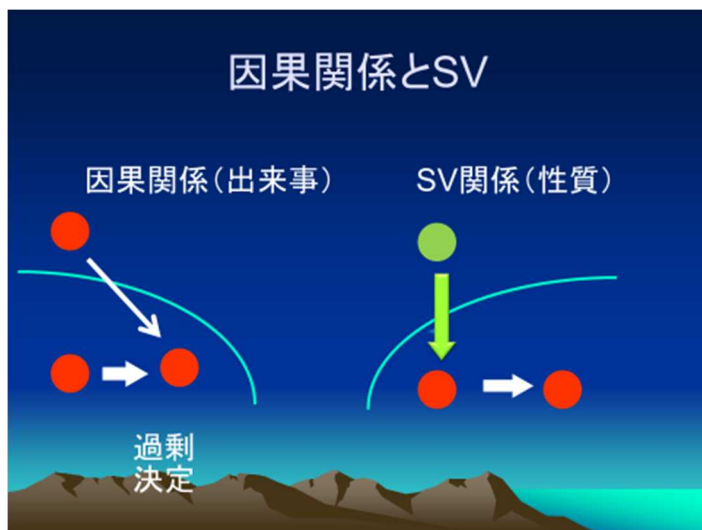
J. キムを初め多くの物理主義者は、われわれの世界が因果的に閉じていると考えている。その意味は、この世界のいかなる出来事の原因をどこまで辿っても、そのつど十分な

原因としての物理的な出来事が見出せる、ということである。つまり、どの出来事に関しても、「物理的ではない心的出来事」が原因となる必要はない。したがって、ここから直ちに、もしも「物理的でない十分な原因」によって引き起こされることが「自由な行為である」ために必要であるなら、自由な行為は、常に、それを引き起こすのに十分な2つの原因を持つことになる。つまり、これは因果過程における「過剰決定」と呼ばれる事態である。しかし、自然界において過剰決定は極めて希なケースである。したがって、人間の自由な行為においてのみ、その希な過剰決定が「常に」生じていることになるが、それは不合理であろう。というのも、「一方の原因（脳内の出来事）が仮に生じなかったとしても、他方の原因（意志の働き）によって問題の行為はやはり行われたであろう」、というような反事実的判断を支持する根拠は、経験的には見出せそうもないからである。

## 2. スーパーヴィーニエンス (SV)

すると、ここから導かれるアイデアの一つは、「意志という状態は脳状態にスーパーヴィーンする心的性質だ」という主張である。これは、非還元的な物理主義の見方に基づくが、それが意味するところは、以下の3つである。(1) 意志の出現は、脳状態の変化と出来事としては同一であるが、性質としては異なる。(2) 因果関係としては「意志→行為」は、「脳状態→身体状態」、という物理的因果関係としてしか存在しない。(3) 性質間の関係としては、脳状態の性質に意志内容といった心的性質がスーパーヴィーン (SV) する。

この図において、左が、因果関係として「意志→行為」を捉えると過剰決定が生ずることを示している。ここで、カーブは心と脳の境界を表し、カーブの上の○印は「意志」、下の○印が「脳状態」と「行為」を表す。左では2つの矢印はいずれも因果関係を表すが、右ではカーブをまたぐ矢印はSV関係を表している。したがって、右のカーブの上の○印は、出来事ではなく、性質としての「意志内容」を示す。



ところで、意志の生起もまた物理的出来事であるから、たとえ意志の生起に因果的な非決定性があるとしても、それが意志に意味のある「自由」を与えることはないであろう。というのも、因果的な非決定性（ランダム性）が大きければ大きいほど、意志の来歴は理解不可能になるからである（例えば、「たいていの場合、なぜ自分が今それをしようとしているのかさっぱり分からない」、というのは自由な行為の本質ではあるまい）。

さて、SV 関係とは基本的にどのようなものかを押さえておこう。SV とは性質間の連動関係、もしくは連動的な依存関係のことであり、ここで、その関係にあると考えられているのは、心的性質と物理的性質である。そこで、非還元的物理主義の主張は、「心的性質は物理的性質に SV する」、と言い表すことができる。したがって、ここには、ある時点での身体物体の完全なコピーはその時点での意識／心の完全なコピーを伴う、ということが含意されている。

ところで、大雑把に「心の状態は脳状態に SV する」と言うとしても、精確に言って心の状態の何が脳状態にローカルに SV するのだろうか？ というのも、ローカルに SV する性質のみが、その時点での物理的因果関係に関与できるように思われるからである。例えば「千円札である」というような性質は、千円札の分子レベルのコピーを考えてみれば分かるように、本物のお札の色や形だけではなく、その歴史的来歴（物質複製機によってではなく、造幣局の然るべき印刷機によって刷られた）を含めたお札全体にグローバルに SV するからである。しかし、精巧な偽札が店員を瞞すとき、問題となる性質「千円札に見える」は、実はそのお札にその時点でローカルに SV している性質であり、それが、「詐欺」という因果過程に関与するのである。

同様に（少し設定をずらせば）、「駅に急ごう」という意志が足を速めさせるように見えるとき、脳状態から身体運動への因果過程に関与している心的状態は、その脳状態にローカルに SV している性質でなければならない。しかし、どのような心的性質が、脳にローカルに SV するのか？ ここには、チャーチランドやスティッチなどの消去主義的物理主義者による多くの議論が横たわっている。結論だけを言えば、それは、概念化もしくは言語化される以前の、自然現象として脳に発生する心的状態であろう。記号【 】を使って表現すれば、【意識】、【クオリア】、【感情】、【欲求】、【信念】、等々であり、したがって、お馴染みの心的な諸概念、例えば素朴心理学によって帰属される心的状態（感情、欲求、信念など）は、これらではない。したがって、例えば【怒り】は、通常われわれが理解する「怒り」の素朴心理学的説明には従わない振る舞いをするだろう。

### 3. 心的状態（心的性質）と因果的効力

ここで、心的性質と因果的効力についてまとめておこう。まず、脳にローカルに SV しているのではない心的状態は、因果的効力を持たないと考えられる。例えば、「憐憫」とされた感情は、実はその時点での実際の脳過程（ $\alpha$ ）とは法則的に連動しない、素朴心理学的にのみ正当化された解釈にすぎず、それゆえその「憐憫」は実は当の因果過程に関与していない。この場合、 $\alpha$ に対応する心的状態はむしろ、素朴心理学的には正確な分類を持たない【嫌悪】であるかもしれない。

さらに、脳にローカルに SV する心的状態であっても、それは基盤性質である脳状態の因果的効力を逸脱した、独自の因果的パワーを持つわけではない。これは、因果的閉包性から導かれる。したがって、還元的物理主義者が期待するように、仮に心的性質と物理的

性質が同一であったとしても、心的性質は「心的性質として qua mental properties」因果的効力を持つのではない。心的性質による予測や説明は、あくまで、物理的性質が生じさせる因果過程を「なぞる」だけなのである。

したがって、物理的世界における心的性質の位置をまとめるとこうなる。心的現象（クオリア、意識など）は、因果的閉包性により、物理的現象の原因であれば「過剰決定」をもたらす。そうでないなら、「独自の」心的性質としては、「独自の」因果的効力を持たない。しかし、われわれの経験上、それらは物理的現象に依存して生じる。つまり、それらはまったくデタラメな、理解不可能な仕方では生じるのではない。言わば、ファイグルがかつて述べたような法則的ちょうちん(nomological danglers)なのである。それゆえ、エピフェノメナリズム(epiphenomenalism)が妥当な帰結となる。

#### 4. エピフェノメナリズム恐怖症 (epiphobia)

論理の矢印は、われわれがエピフェノメナリズムを受け容れるべきことを指し示している。しかし、これこそが、西洋の多くの思想家にとって憂いの種であり、実際、それを回避したいという願望が現代の心の哲学、とくに「心的因果」をめぐる論争の最大の動機だった。それは、端的に、自分たちの意図的行為のすべてが、自分たちの知らない物理過程の〈影〉にすぎない、ということへの恐怖である。J. フォーダーをもじって言えば、もしエピフェノメナリズムが真なら、”it is the end of the world!”と思われるだろう。また、フォーダーの論敵のキムもこう問う。われわれの行為は、走行する自動車が作り出す、自動車の〈影〉のようなものなのか？ と。しかし、実は、それは情緒的な問題にすぎないだろう。一旦、基本的に物理主義を受け容れたなら、心的性質に〈物理的性質を超える独自の〉因果的効力を認める道はないのだから。

そこで、改めて「エピフェノメナリズムで何が悪いのか」と問うてみよう。問題は、結局のところ、情緒的な反応を除けば、自由意志なき世界において行為責任の帰属をいかに考えるのか、ということに帰着する。すべての出来事が法則的に決定された世界において、われわれはそもそも何かに対して責任を問えるのだろうか？ もし責任の帰属根拠が「物理法則を逃れた意志を原因とする」というものであるなら、われわれはこの世界のいかなる行為に対しても行為者に責任を問うことはできないであろう。しかし、逆に、責任帰属の根拠がたんに「ある出来事の原因である」という、無原則とも言えるほど極端に「緩い基準」であるなら、われわれは基本的に何に対しても責任を問うことができることになる。それこそ、普通の事物の生起から、動物や人間の振る舞いまで、およそすべてに責任を問うてもよいのだ。しかし、風で飛ばされた木の枝に、通行人への傷害の責任を問い、さらにその枝を罰するために裁判にかける、などということは、責任と処罰の社会的な意味からしてナンセンスである。したがって、存在論的な正当化はどうあれ、〈社会的実践〉という観点からすれば、責任を問うことに意味があるのは、そもそも、「被害者と加害者の位置を入れ替えれば応報や報復という意味が生ずる」、という意味でく差し替

え可能な社会的存在者>としてのわれわれ人間である（応報主義）。そしてまた第二には、将来の類似状況に対する警告が意味をなす<同類的存在者>としての、これまたわれわれ人間である（帰結主義）。

このように社会的な実践、もしくは慣習として責任帰属を眺めた場合、その根拠はその振る舞いの<合理性>にある、とすることができる。すなわち、われわれは、行為に対してのみ責任を問うのである。なぜなら身体運動や振る舞いは、欲求と信念によってその理由が合理的に再構成されうる場合に、そしてその場合にのみ<行為>となるからである。つまり一言でいえば、<信念・欲求・行為>の三位一体が持つ合理性にこそ、責任帰属の根拠がある。すると、ここから、責任（道徳）の自然化は可能か、つまり道徳や倫理学を自然科学に還元できるのか、という問いに対する答えが導き出されるだろう。つまり、こういうことである。もしも合理性によって制約されている行為に関する諸概念や心的諸概念、例えば「嫉妬」などが脳状態にローカルにSVするなら、言い換えれば合理性概念が自然的性質にローカルにSVするなら、責任は<きれいに>自然化されるはずである。したがって、道徳や倫理もまた自然科学の一つになるだろう。しかし、行為の概念、心的諸概念、ひいては倫理道徳の諸概念はいずれも、<きれいに>自然化することは不可能である。なぜなら、それらは脳の因果過程にローカルにSVしていないからである。こういう言い方を好む人がいるかもしれないが、それらの概念の出自と分類は、自然科学的に見れば「いい加減」だからである。責任の自然化を阻む要因として、これまで心の哲学は、いやというほど多くの議論を積み上げてきた。その項目の幾つかだけをここで挙げておこう。①心的内容の外在性（パトナム、他）、②心の非法則性（ディヴィッドソン）③消去主義的唯物論（チャーチランド、他）、④コネクショニズム（スティッチ、他）、⑤道具主義（デネット）、直ちに思いつくだけでもこれだけある。つまり、一言でいえば、合理性は因果性の中に法則的な対応物を持たないのだ。

## 5. 新たな倫理道徳の創作

ここで、新たな倫理道徳の体系を考える上で重要な2つのテーゼを提案しておきたい。第一は<連続性テーゼ>である。例えば、脳の異常や障害による行為と、いわゆる普通の行為は、行為者に帰属できる責任の大きさにおいて連続的である。最も大きな責任は、通常の成人の行為者が通常の状況で行う意図的行為において問われるであろう（もっとも、この場合でさえも、様々なモラル・ラック、すなわち道徳的な運不運の介在する余地は排除できない）。その対極をなす場合の一つは、心神耗弱や心神喪失の場合であり、現在の日本の刑法では、心神喪失者の場合、刑事責任を問うことはできない。法的責任の設定の場合が明らかにそうであるように、帰属される責任の大きさには程度がある。これは、行為概念に関する哲学的な反省によって導かれたわけではないだろうが、社会的実践として妥当であるのは、責任帰属の決定的な要素である<行為の合理性>に程度があることを社会が実感しているからである。合理性が自然化できない以上、この連続性テーゼは、われわれが行為に責任を問う限

り避けられない。具体的には、人体の運動が、事物としてある結果を引き起こすケース（責任がほぼ 0%）から、行為理由にしたがって行為者が合理的な行為によってある結果を引き起こすケース（責任がほぼ 100%）までの間に、中間の様々なケース（ $0 < X < 100$ ）がありうることになる。

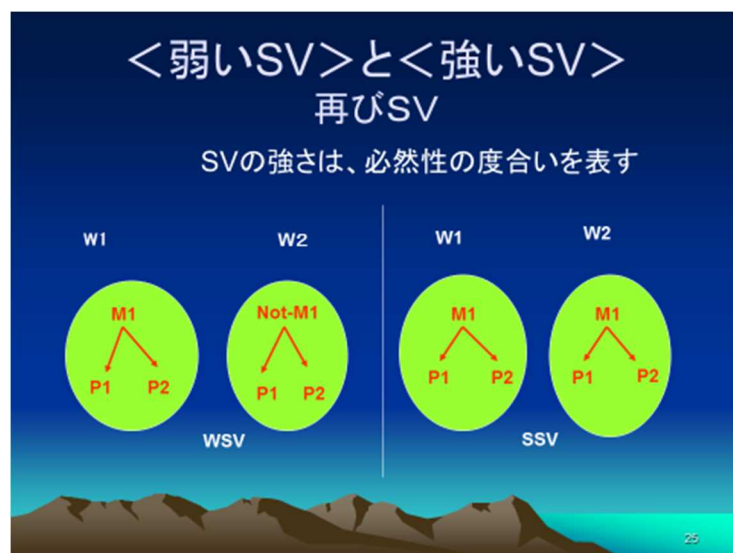
その中間のケースを合理性以外の何に依拠して考えるのか、あるいはそもそも合理性に依拠して行為責任を問うことにどのような根拠があるのか。ここで、私は、先ほど述べた第二のテーゼを提案したい。それは〈恣意性テーゼ〉である。翻って考えるに、倫理道德の究極の正当化は (A) 経験を越えた何らかの原理によってアプリアリになされるか、(B) 経験に依拠して、つまり自然化によってなされるか、のいずれかしかないだろう。しかし、ここで詳細に議論はできないが、どちらの方策も望みなき試みであると思われる。どちらも、社会実践としての倫理道德を十分に正当化することはできない。つまり、われわれは、究極の正当化を欠いた、倫理道德の〈絶対的選択〉に直面しているのである。これは、われわれにとって朗報であろうか、それとも凶報であろうか。いずれにせよ、厳しい試練であるのは間違いない。すなわち、われわれに課せられているのは、道德的・倫理的責任の創作であり、それはとりもなおさず〈倫理道德システム〉の新たな創作である。何が倫理的に「悪」であり「善」であるかは、自然化できないのであるから、自然科学の真理によって決定することはできない。また、それは、何らかの超越的な真理から導出することもできないのであるから、概念的な議論のみによって決定することもできない。すると、こう問う人がいるかもしれない。道德や倫理に関して圧倒的に多くの人々が一致するなら、それは正しいのではないかと。しかし、たとえ倫理道德に関して多数の人が一致しても、それはなお「集団幻覚」でありうる。なぜなら倫理道德に関する思考や信念が脳に SV するとしても、それらは物理的世界に SV しないからである。つまり、倫理道德そのものは物理的事実ではなく、したがって実在していないからである（それらについての思考は、脳にローカルに SV する限りで物理的事実であり、したがって実在するとしても）。

したがって、ここから、どのような倫理道德のシステムも、われわれは〈幻想の制度〉として、しかも〈真正の倫理道德〉として選びうる、ということが帰結する。しかし、一体、どのようなタイプの倫理道德を選んだら良いのか？ 答えはすぐには与えられないだろう。これこそ、われわれ、哲学・倫理学の徒が 21 世紀において果たすべき喫緊の課題である。私がこの時点でとりあえず提起できる〈新たな倫理道德〉創作のための前提は、政治哲学的意味での自由主義（リバタリアニズム）である。これは、よく混同されるが、先行原因によって拘束されていない意志のみを「自由意志」と認め、それによる行為のみを「自由な行為」とする、いわゆる古典的な意味での自由主義ではない。後者の意味での「自由意志」、つまり形而上学的な意味での「自由意志」の概念は、それ自体、人々の願望を寄せ集めた不整合な代物でしかないし、物理主義とも相容れない。したがって、それは明確に放擲すべきである。では、ここで私が提案する〈政治哲学的な意味での自由主義〉の本体は何か？ それは、「他者危害の原則」を侵さないいかなる行為も倫理的には許される、とする自由至上

主義である。これはこのままでは内容を欠いた倫理道德であるが、われわれはまず、これを出発点の大枠として選ぶほかはないように思われる。他者とは誰のことか？ 知性ある異星人やロボットも含まれるのか、ということのみならず、認知症の老人や幼児、あるいは動物やペットをどう考えるのか、ということも今後の課題である。そして、仮に倫理道德の原則として、われわれが「他者危害の原則」以外のものを、新たに、人類として選択することができないとすれば、「他者危害の原則」以外の考慮は、すべて、〈他者と付き合うわれわれの流儀・作法〉、すなわちわれわれの文化・趣味ということになるであろう。その意味では、今後重要になってくるのは、地球規模での人類の、新たな倫理道德の創作と文化・趣味の洗練である。

## 6. スーパーヴィーニエンス (SV) の「強さ」と偶然的真理としての物理主義

SV 関係については、「強い SV (Strong SV)」と「弱い SV (Weak SV)」を区別するのが普通である。これらの違いは、SV 関係がある可能世界 (W1) だけで成立しているのか、それとも任意の可能世界 (Wn) で成立しているのかの違いである。そして、これらの強さは、端的に「必然性」の強さを意味する。つまり、もしもある SV 関係が SSV であるなら、その関係はあらゆる可能世界において成立する。すなわち、その関係は論理的必然性と同じだけ「強い必然性」をもつ。それに対して、WSV の場合には、その関係はただかた一つ可能世界においてのみ成立するにすぎないので、必然性をもたず、せいぜい「偶然にもまったく例外がない」という意味で「偶然的な普遍性」をもつにすぎない。



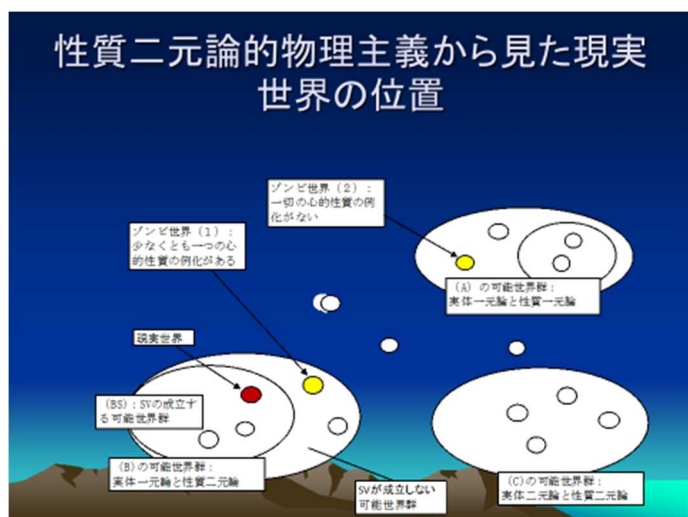
したがって、例えば、「C 繊維の興奮」と「痛みの感覚」の対応関係が現実世界だけで成立するにすぎないのであれば、それは、結局のところ、われわれの世界だけで成り立つ「偶然的」な関係であり、法則的な必然性すらも持たない。というのも、現実世界によく似た可能世界 W2 では、C 繊維と痛みのその対応関係が成立していないかもしれないからである。したがって、キムが口を酸っぱくして主張していたように、物理主義にとっては、WSV は意味がない。

しかし、WSV よりも強く、SSV よりも弱い、中間の強さの SV が存在するだろう。そして、これこそが〈偶然的真理としての物理主義〉という、われわれの主張である。それは、ある可能世界でのみ成立する（必然性をもたない）のでもなく、あらゆる可能世界で成立する（無条件の必然性をもつ）のでもない。われわれの自然法則がそうであるように、それは、ある範囲の可能世界群で成立するがゆえに、弱い必然性、法則的な必然性、したがって〈偶然的

な必然性>しかもたない。われわれの主張は、現実世界とはそのような「中間の強さのSV」が成り立つ可能世界群に属する一つの可能世界である、というものに他ならない。そして、SVが真に根源的關係なら、こうしたSVがなぜしかじかの範囲の可能世界群で成立するのかについては、もはや還元的説明は存在しない。

したがって、以上から、<偶然的真理としての物理主義>が、心身に関する性質二元論的物理主義を含蓄することが帰結する。というのも、心身のSV関係がただか一定の範囲の可能世界群でしか成立しないのであれば、「実はそれら心身の二性質は同一であった」という可能性は排除されるからである。つまり、当の範囲外のどこかの可能世界では、心身のSV関係は崩れており、したがって両者は同一の性質ではありえない。

そこで、心身の関係に関して、可能世界の概略図をここで描いてみよう。もちろん、これは、極端に簡略化した要素だけで描いた一種の図式である。まず、あらゆる可能世界から成る論理空間の中を、次のような可能世界群が棲み分けている。



(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。ここでは、いわゆる還元的物理主義が成立し、いかなる心的個体(と思われるもの)も実は何らかの物理的個体と同一であり、いかなる心的性質(と思われるもの)も実は何らかの自然な物理的性質と同一である。(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と同一であるが、いかなる心的性質も自然な物理的性質と同一ではない。ここでは非還元的物理主義が成立し、これらの可能世界の中に現実世界も含まれる、というのがわれわれの主張である。(C) 実体二元論と性質二元論が成り立つ可能世界群。デカルトが考えたように、これらの世界では、例えば霊的実体や思惟実体がいかなる物理的個体(延長実体)とも同一ではないものとして存在し、かつ、いかなる心的性質も自然な物理的性質と同一ではない。この図はたぶん紙面では判読しにくいと思われるので、ぜひ拙論「感情のクオリアと可能世界」(『感情とクオリアの謎』所収、昭和堂、長滝・柴田・美濃共編、2008、p. 30)の図を参照願いたい。もしくは、柴田の研究関連webサイト：<http://siva.w3.kanazawa-u.ac.jp/> (「柴田の業績」)の中で、同論文に当たられたい。

さて、それでは、このようなく偶然的真理としての物理主義>、すなわち非還元的物理主義が現実世界に妥当するかどうかは今後の議論に委ねるとして、一体、哲学は何を明らかに



しようとするのだろうか？ 私の考えでは、哲学的考察の任務は、すべての科学の真理と、すべての経験と、すべての常識を可能な限り最大限整合的に説明することである。つまり哲学は、現実世界がいかなる可能世界であるかについての、最も合理的で、最も整合的な説明をわれわれに提供しようとするものである。