

ロボットの心は やはり冷たいか？

於：西田幾多郎記念哲学館

ロボットは、認知機能を超えて、意識・クオリア(qualia)を持てるか？

金沢大学大学院人間社会環境研究科
／人文学類・・・柴田正良 July 4, '09

話の流れ

1. クオリア (qualia) とは何か？
2. 物理プロセスからの遊離
因果的閉包性
3. ロボットも機能的心ならもちろん持てるが
クオリアと意識は？
4. われわれの世界
ロボットがクオリアをもつ物理的世界



クオリア(qualia)とは何か

- クオリアは<感覚質>と訳されるように、ある感覚経験があなたにもたらす<感じ>である。歯の<痛み>、コーヒーの<苦み>、朝の光の<まぶしさ>、などなど。
- クオリアは、徹底して主観的、一人称的だ。つまり、客観的な現象としては、あなたの脳の中にも見出せない。

クオリアの私秘性

- 強く香るユリの花・・・あなたが感じる<香り>と私が感じる<香り>が同じかどうかは、あなたが私にならない限り原理的に確かめようがない。
- What is it like to be a bat? (T. Nagel)
- 赤い色が見えるとはまさにそのことであるような、そのそれとは何か？

痛みの生理学的(因果的)な説明に クオリアは必要か？

- <痛み>という心的状態は、「入力(傷)→内部状態(信念や欲求)→出力(行動)」の連関を支える一定の因果的役割に他ならない。
- この因果的役割を果たすもの＝痛覚神経の興奮(in human)
- しかし、この説明の中で、<痛みの感覚>はどこに登場するのだろうか？

クオリアは物理的世界のどこにも 居場所がない

- どこにも登場しない。神経生理学的レベルのどの出来事も<痛みの感覚>であることを本質的に要請されていない。
- 逆転クオリア (inverted qualia)
- 不在クオリア (absent qualia)
- 哲学的ゾンビ (philosophical zombies)



物理的世界における因果的閉包性

- いかなる物理的出来事(結果)にも、それを引き起こすのに十分な物理的出来事(原因)が存在する。
- どんな物理的出来事の原因の連鎖をどれほど遡っても、この物理的世界を飛び出すということはない。



物理的世界からの遊離

1. 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。

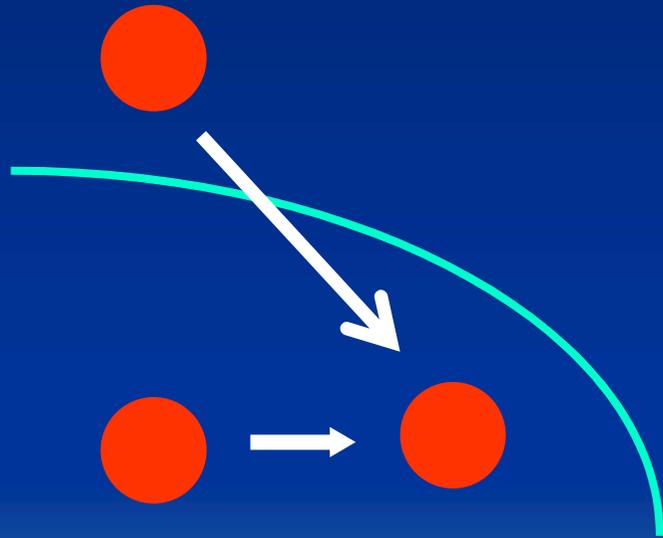
2. それらは、**物理的現象に依存して生じる**
← まったくデタラメな、理解不可能な仕方で生じるのではない。 **法則的ちょうちん**
(nomological danglers・・・ファイグル)

随伴(スーパーヴィーニエンス)

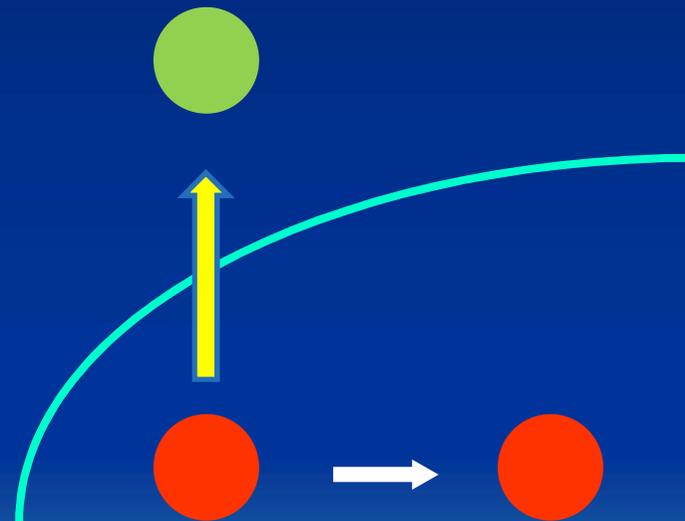
- スーパーヴィーニエンス(SV)とは性質間の連動関係、もしくは連動的な依存関係のことである。
- (非還元的)物理主義の主張・・・心的性質は物理的性質にSVする。
- 分子レベルでの完全な物質複製機 → 身体物体の完全なコピーは意識／心の完全なコピーだ(SVの主張)

因果関係とSV

因果関係(出来事)



SV関係(性質)



過剰決定

機能的性質とクオリア

- 心的性質（信念、欲求、意図、感情など）の多くは機能的性質である。
- 感覚（痛み、視覚、聴覚など）ですらも、情報処理／機能的性質の一面をもつ。
- しかし、痛みクオリア、赤クオリア、音色クオリアなどは、機能とは異なる…機能的性質ではなく、現象的性質である。

性質の同一性ではなぜダメか 機能的性質ですら・・・ましてクオリアでは

- **還元的物理主義** → すべての心的性質は物理的性質と同一。
- **<痛い> = <C繊維の興奮>** (脳細胞 in human のある物理的な性質状態)
↓
- C繊維をもっていない生物(**タコやミミズや宇宙人**)は、痛みももっていないことになる。というのも、 $a = b$ で $b \neq c$ ならば、 $a \neq c$ だから。

多重実現 (multiple realization)

- **機能主義** → 心的性質は、多様な素材において実現される機能的状態だ。
- Pは、〈スペインのグラナダに行きたい〉という心的性質（／心的状態）
- Mnは、それを実現する何らかの基盤性質（／基盤状態）
- **心的状態は基盤状態の選言によって実現される。**
- $P_x \Leftrightarrow (M1_x \vee M2_x \vee \dots \vee Mnx)$



ロボットがなかなか出てこない？

- 機能的性質としての心的性質が多重に実現されるなら、心は、**電子工学的な素材**によっても実現できるだろう。
- そのためには、心と身体が果たしている機能が、**脳神経科学的にも、認知科学的にも**徹底して解明されなければならない。
- 物理主義を取れば、実現性質(素材)と機能的質(心)はSVの関係にある → **機能的心を持つロボットが可能**

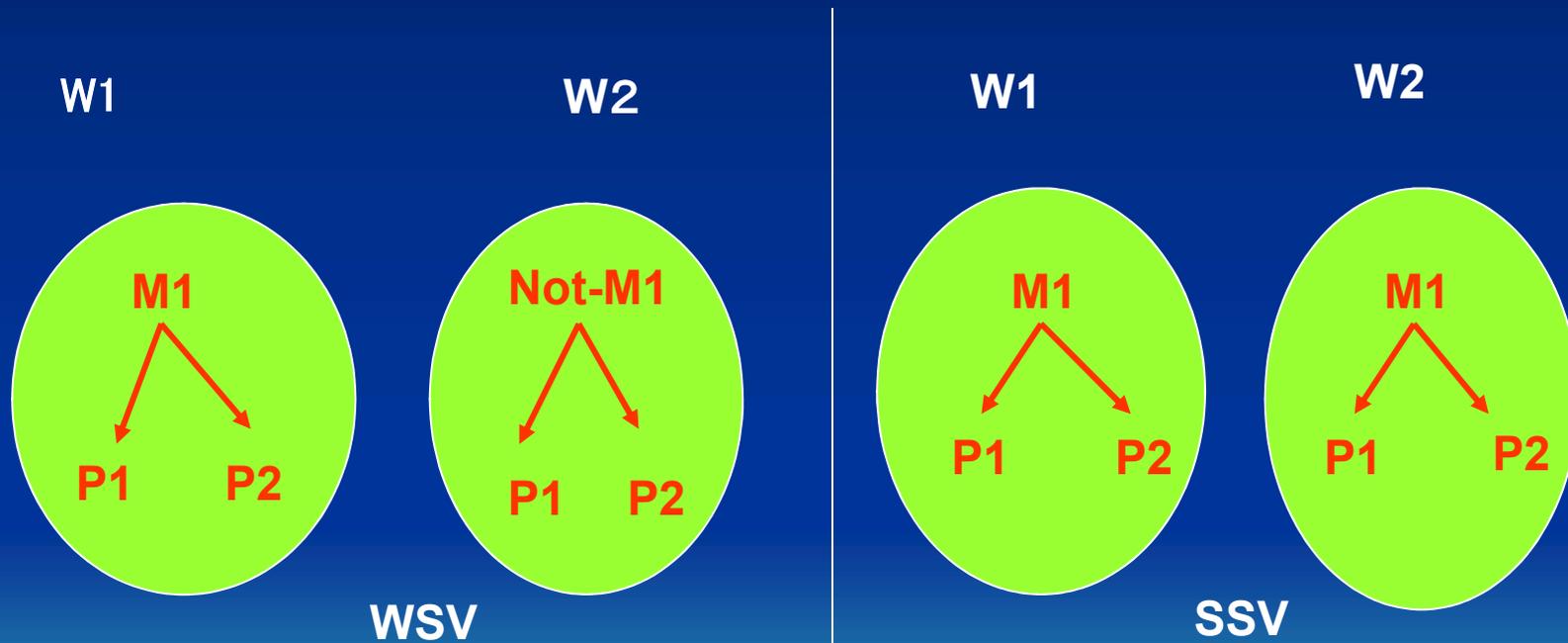
ロボットはクオリアも持つのか？

- クオリアが物理的性質にSVするのは、物理主義を受け入れれば保証される。
- しかし、クオリアが機能にSVすることを示す強力な論証はない。機能は性質の果たす因果的役割だからだ。
- 機能的心を持つロボットが可能だとしても、クオリアを持つロボットは可能なのか？

→ 不可能だとすることの不合理性

<弱いSV>と<強いSV>

SVの強さは、必然性の度合いを表す



SVの成り立つ可能世界の範囲

- <弱いSV>と<強いSV>の違い →
スーパーヴィーニエンス関係(例えば、C繊維の興奮と痛みの感覚)がある可能世界内部で成立すればいいのか、それとも任意の可能世界相互にまたがって成立するのか、という違い。



- 物理主義は、<強いSV>の主張では意味がない



<ほど良く強いSV>

- しかし、<弱いSV>よりも強く、<強いSV>よりは弱いSVが存在するだろう。そして、これこそがわれわれの主張だ。
- それは、ある可能世界でのみ成立する(偶然性)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)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関係(弱い必然性)だ。 → 偶然的真理としての物理主義。
- SVが真に根源的關係なら、そのSVがなぜしかじかの範囲の可能世界で成立するのかについての還元的説明は存在しない。

SVは根元的関係か？

- **SV ≠ 同一性関係、SV ≠ 因果関係**
→ SVは、他の関係によって説明できない世界の原初的、根元的関係である。

2つの疑問

- 本当にSVのような原初的關係はあるのか？
- あるなら、SVの成り立つ可能世界の範囲はどれくらいで、なぜそうなのか？

可能世界とは？

- 可能世界とは、現実世界と少しでも違った状況を考えたときに、われわれが呼びだしている**概念装置**である。←反事実的想定
- 可能世界の4つの特徴
整合性、完全性、飽和性、独立性

これらを満たす**可能世界は、無限に存在する**(不計算無限個)。



可能性と必然性

- **可能性と必然性の関係** (偶然とか法則的必然などはすべて可能世界における真理に還元できる)

pは可能である... $\diamond p$

pは必然である... $\Box p$

$\Box p \equiv$ あらゆる可能世界でpが真である。

$\diamond p \equiv$ 少なくとも一つの可能世界でpが真である。

ゾンビ・ロボットの不可能性

- **ゾンビ・ロボット**・・・人間とまったく同じ認知機能を持ちながら一切のクオリア・意識を欠いている。
- ゾンビ・ロボットの不可能性は論理的不可能性ではない → **認知機能と主観的経験の不整合**である。

クオリア報告から一人称的判断へ

- 認知機能ロボットがクオリア体験の報告をする → 「まぶしい鮮やかな青が見える」

- その報告がわれわれと同じなら



- 彼に不在クオリアや逆転クオリアを想定することには、「それが機械である」という抵抗感しか理由がない。

- 認知と経験の不整合…理由なき想定

一人称的判断から経験の主体へ

- 認知機能ロボットが自分の一人称的体験を報告する → 「仲間がいなくて寂しい」
- その報告が、内部メカニズムの三人称的な観察ではなく、一人称的な経験内容にもとづく



彼はすでに「**経験の主体**」である。



ロボットがクオリアをもつ世界

認知機能ロボットがゾンビ・ロボットだとすることは、自己挫折的な想定である。



「私はゾンビ・ロボットである」 → この言明
が真なら、経験の主体は存在しない ←

われわれの世界がロボットにクオリアを授ける



性質二元論的物理主義 (現実世界という可能世界)

- ローカルな真理としての物理主義

→ 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。



性質二元論的物理主義(2)

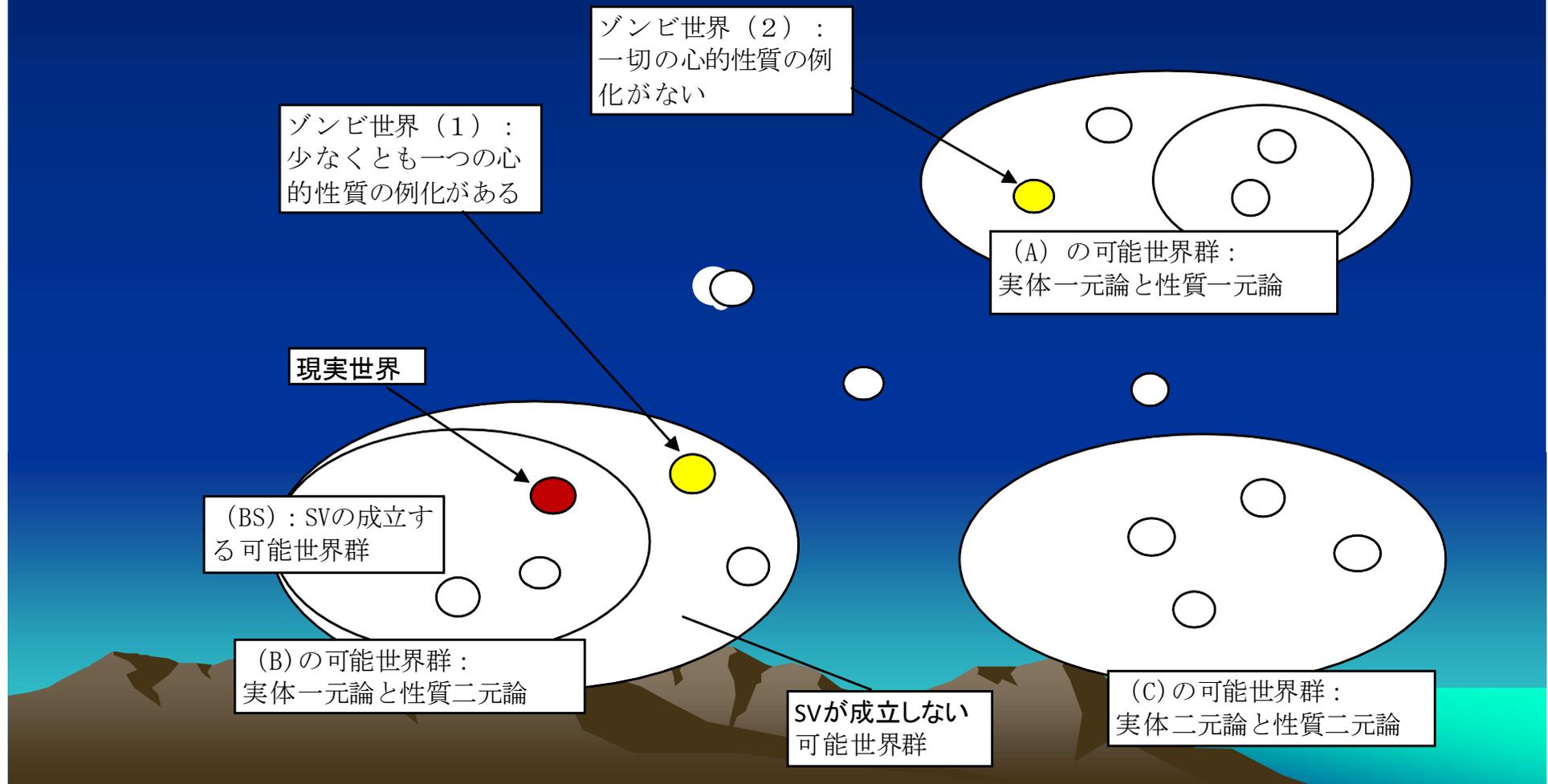
- (B)実体一元論と性質二元論が成り立つ可能世界群。それらの世界では、例えば、いかなる心的個体も物理的個体と同一だが、いかなる心的性質も自然な物理的性質と同一ではない。
- (C)実体二元論と性質二元論が成り立つ可能世界群。これら世界では、例えば霊的実体がいかなる物理的個体とも同一ではないものとして存在し、かつ、いかなる心的性質も自然な物理的性質と同一ではない。



性質二元論的物理主義(3)

- (B)の可能世界群の内部には、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。
- 私が提案している性質二元論的物理主義→**現実世界はこの(BS)の可能世界群に属する**

性質二元論的物理主義から見た現実世界の位置



ゾンビ・ロボット世界は現実世界から 超遠い

- 可能世界群(BS) → 物理主義

「心的個体に関しては唯物的」と「物心のスーパーヴィーニエンスが成立」

ゾンビ・ロボット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。



クオリア・ロボットの作り方

- われわれは物理的なものを作ることしかできない。しかし、クオリア・意識は物理的なものではない。
- 認知機能ロボットを作ること、それがクオリアロボットを作るただ一つの方法である。
- われわれの世界は、認知機能ロボットにクオリアを授けるような可能世界である。



では哲学は何を明らかにするのか？

- 哲学的考察は、すべての科学の真理と、すべての常識を最大限整合的に説明する道を探る。
- 哲学は、現実世界がいかなる可能世界であるか、についてのもっとも合理的な説明をわれわれに提供しようとする。

おしまい