

# ロボットはモーツァルトやダリになれるか？

スーパーヴィーニエンス(SV)と反实在論の狭間で

第21回 AI美芸研  
「美意識のハードプロブレム」  
於：なかのZERO

金沢大学教育担当理事(副学長)  
柴田正良 Oct. 13, 2018

# 話の流れ

1. ロボットは、機能的な「心」(心理学的機能)以上に、感ずる「心」(現象的意識やクオリア)を持ちうるか？
2. <美しさ>は物理的性質にSVするか？
3. SV関係における<美しく見える>と<美しさそのもの>・・・現実世界はどういう可能世界か？
4. ロボットにとっての<美しいもの>とは？

# 心的機能とクオリア

- 認知症で問題となるような心の諸機能は、脳科学的、生理学的、化学的、分子生物学的現象として、科学的説明が切れ目なく全てを明らかにするだろう。
- しかし、脳過程の現象がいかにして「レモンの色の感じ」(クオリア)を生み出すのかは、脳過程の説明をいくら精緻にしても、理解できない。
- 機能とクオリアの間には、物理学的、神経生理学的説明では埋められないギャップがある。

# 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…
- 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

- 挽き立てのコーヒー・・・あなたが感じる<香り>と、私が感じる<香り>が同じであるかどうかは、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)
- コウモリであることがどのようなことであるかは、コウモリになってみなければ分からない。
- しかし、最も決定的な点は、他人同士ですら互いのクオリアは不可知だということ。

# 痛みの生理学的(因果的)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、＜痛みの感覚質＞、＜痛みのクオリア＞はどこに登場するのか？

# 機能はクオリアを必要としない

- どこにも登場しない。神経生理学的レベルのどの出来事 (x) も、〈痛みのクオリア〉 (p) であること ( $x = p$ ) を本質的に要請されていない。
- 同様に、機能としての意識、つまり心理学的意識は、現象的意識を必要としていない。

## D. Chalmers による

- 逆転クオリア (inverted qualia)
- 不在クオリア (absent qualia)
- 哲学的ゾンビ (philosophical zombies)

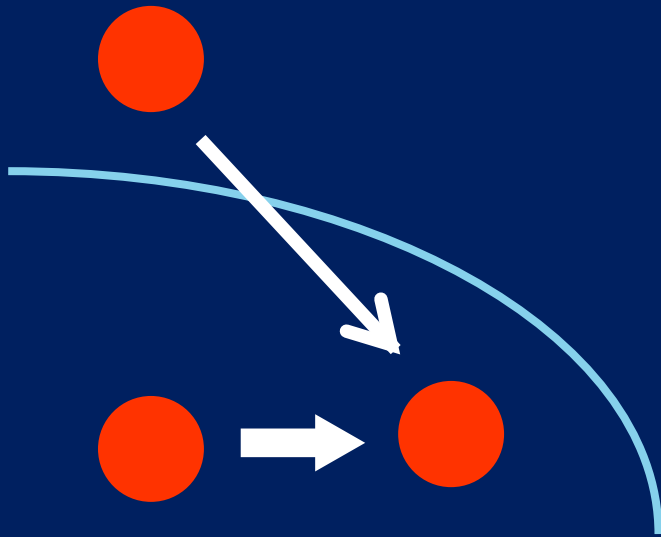
# ロボットはクオリア、あるいは現象的意識をもちうるのか？

- 可能ではあるが、いかにしても技術的には保証できない。
- どれほど人間と同じ心的機能を装備しても、そのロボットがそれ独自のクオリアや現象的意識をもつかどうかは、この現実世界がどのような可能世界であるかに依存する。
- しかし、その事情は隣のあなたも同じだ。
- 結局、クオリアや現象的意識と因果的機能の関係は、自然法則内部の関係ではなく、可能世界によってようやく説明できる、もっと根源的な関係である。



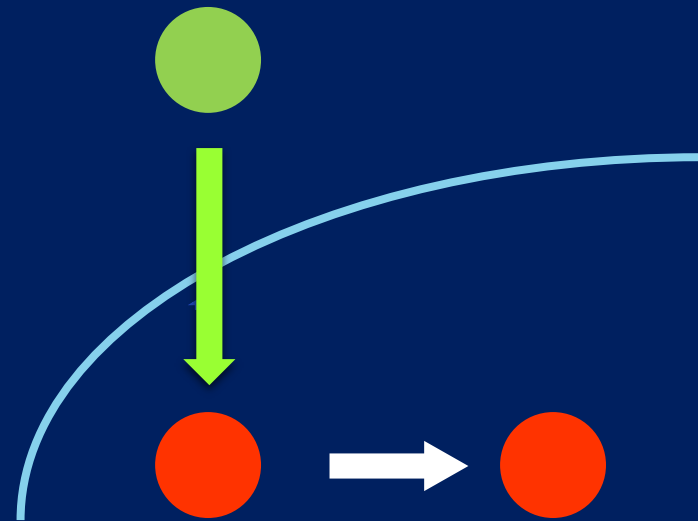
# 因果関係とSV

因果関係(出来事)




過剰  
決定

SV関係(性質)



# スーパーヴィーニエンス(SV)

- 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・現象的意識は物理的性質にSVする。
- 身体物体の完全なコピーは心／意識状態の完全なコピーだ(SVの主張)
- 基盤性質(物理的状态)が同一なら上部性質(心的状態)も同一。しかし逆は成立しない(多重実現)

# ローカルなSVとグローバルなSV

## ローカルなSV関係

〈千円札に見える〉という性質は、ある紙切れの色や図柄といった性質にローカルにSVする。

## グローバルなSV関係

〈千円札である〉という性質は、上の性質にローカルにはSVしない。偽札のこと。

しかし、〈千円札である〉も、因果的歴史を含めた、ある範囲の物理的世界にグローバルにSVする。


本物は、造幣局における過去の印刷を含めた時空範囲の物理的世界にSVする



本物

偽物

# 何が脳へローカルにSVするか？

- 自然現象として脳に発生する心的現象・・・現象的意識、クオリア、感情、美的判断、道徳的判断・・・などは脳にローカルにSVする。
- しかし、判断やクオリアがローカルに脳にSVしても、それが対象とするすべての性質が、存在するわけではない。なぜなら、脳以外のいかなる物理的事物にもSVしないことが可能だから  反実在論
- 「美しい」という判断がなされても、当の物理的事物が〈美しさ〉という性質を持っていないことも可能。つまり、〈美しさ〉はその事物にSVしない。

# 幻覚に関する反実在論

- いま、私の目の前を、小さな虎が空中散歩している。その姿は、どうしたって疑いようがない。私の脳の一部に腫瘍ができたせいだ。
- しかし、「虎が目の前を歩いている」という私の判断がどんなに強固で、視覚クオリアの誠実な報告であっても、そこから虎の実在を導き出すことはできない。
- その〈虎である〉という性質は、見えるその場所にグローバルにさえSVしない。虎の視覚イメージは、私の脳にのみSVする。しかし、また、私の脳の中に、見えのままの虎が存在しているわけでもない。

# 美しさはどこにある？

山口華楊「秋晴」



# 美しさはどこにある？

速水 御舟 「翠苔緑芝」





# 絵の〈美しさそのもの〉は、絵にSVしない

- 千円札の例を思い出そう。
- 〈千円札に見える〉という性質は、認識者の時空的局在を含めれば、千円札の色や図柄にSVする。〈美しく見える〉という性質も、同様に認識者依存的にのみ、絵の色や図柄にSVする。
- しかし、〈美しく見える〉という性質と異なり、〈美しさそのもの〉という、認識者依存的でない性質は絵の色や図柄にSVしない。というのも、〈美しさそのもの〉は物理的性質でもないし、物理的性質に因果的に還元される高階の機能的性質でもないからだ。
- したがって、〈美そのもの〉という性質も存在しない。

# <美しさ>と<美しく見える>(1/2)

- もちろん、<美>に関する反実在論は証明不可能。合理的テーゼとして、提案できるだけだ。
- SVの論理：
  - その絵と観察者を含めた物理的時空がそっくり再現されても、片方の時空部分では<美しさそのもの>がその絵の性質となっていて、別の時空部分ではそうでないことも可能だろう。
- したがって、<美しさそのもの>という性質は、その絵にグローバルにさえSVしない。
- 少なくとも、現実世界を含む近傍の可能世界群では。

# <美しさ>と<美しく見える>(2/2)

- さらに、経験的にはどうか？
- 実は、<美しく見える>という性質ですら、認識者が異なれば、<千円札に見える>ほどの堅固さを持たない。人類のすべてのメンバーが、ある絵を「美しい」と判断するだろうか？あるいは、そうせねばならないか？
- ある行為に関する<正しさ>の判断が、対立する2つの宗教集団や政治集団で異なりうるように、美しさに関する判断も人によって変わるだろう。
- 人間どころか、感覚可能範囲を異にする他の知的生物や異星人(エイリアン)を考えてみよう。彼らの脳に、<美しく見える>という判断が生ずる可能性はさらに低いかもしれない。

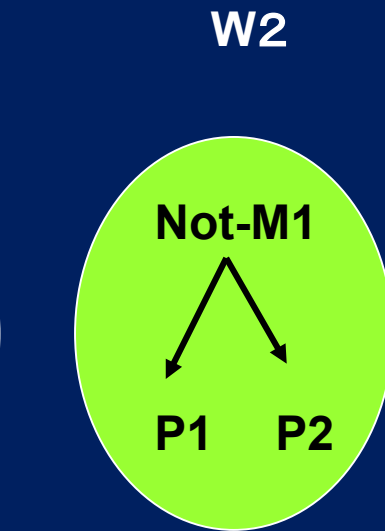
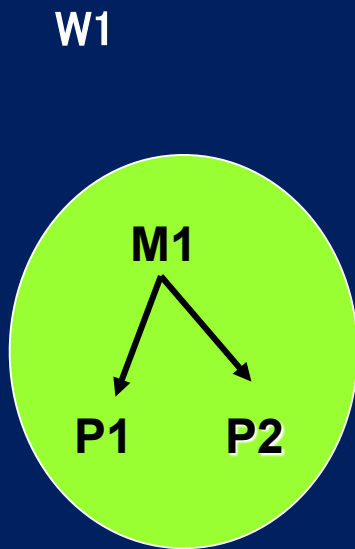
# ＜美しく見えない＞の存在する＜美しさ＞？

ましてや、＜美しさそのもの＞が、＜美しく見えていない＞者たちを含めて、すべての認識者に経験される、ということはないだろう。

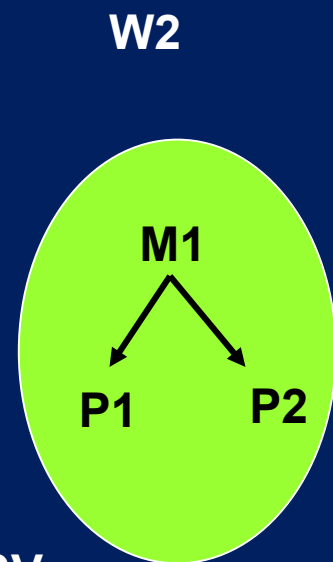
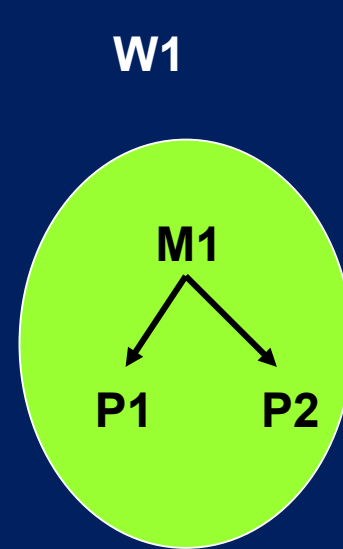
- ＜美しく見える＞を経験できない認識者を超越した＜美しさそのもの＞という性質の存在を、想定する理由が何かあるだろうか？ たとえ、＜美一般＞ではなく、＜この絵固有の美＞だとしても。
- そのような存在は論理的に可能だという理由で、そうした可能世界もあるだろうが、現実世界からは遠い。

# 「弱いSV」と「強いSV」

SVの強さは、必然性の度合いを表す



WSV



SSV

# SVの成り立つ可能世界の範囲

- 「弱いSV」と「強いSV」の違い



SV関係(例えば、C繊維の興奮と痛みのクオリア)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。



- 科学的、すなわち法則的説明は、説明の必然性を要求するがゆえに、「弱いSV」では不十分。
- 自然法則は、複数の可能世界にまたがることで、必然性を得るからだ。

# 「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(たんなる偶然性)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。
- 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、それ以上に遡った法則的な説明は存在しない。

# 性質二元論的物理主義 (現実世界という可能世界)

## ■ 偶然的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。あるいは逆に、霊的実体と心的性質のみが真に存在する(プラトン)。



# 性質二元論的物理主義（1/2）

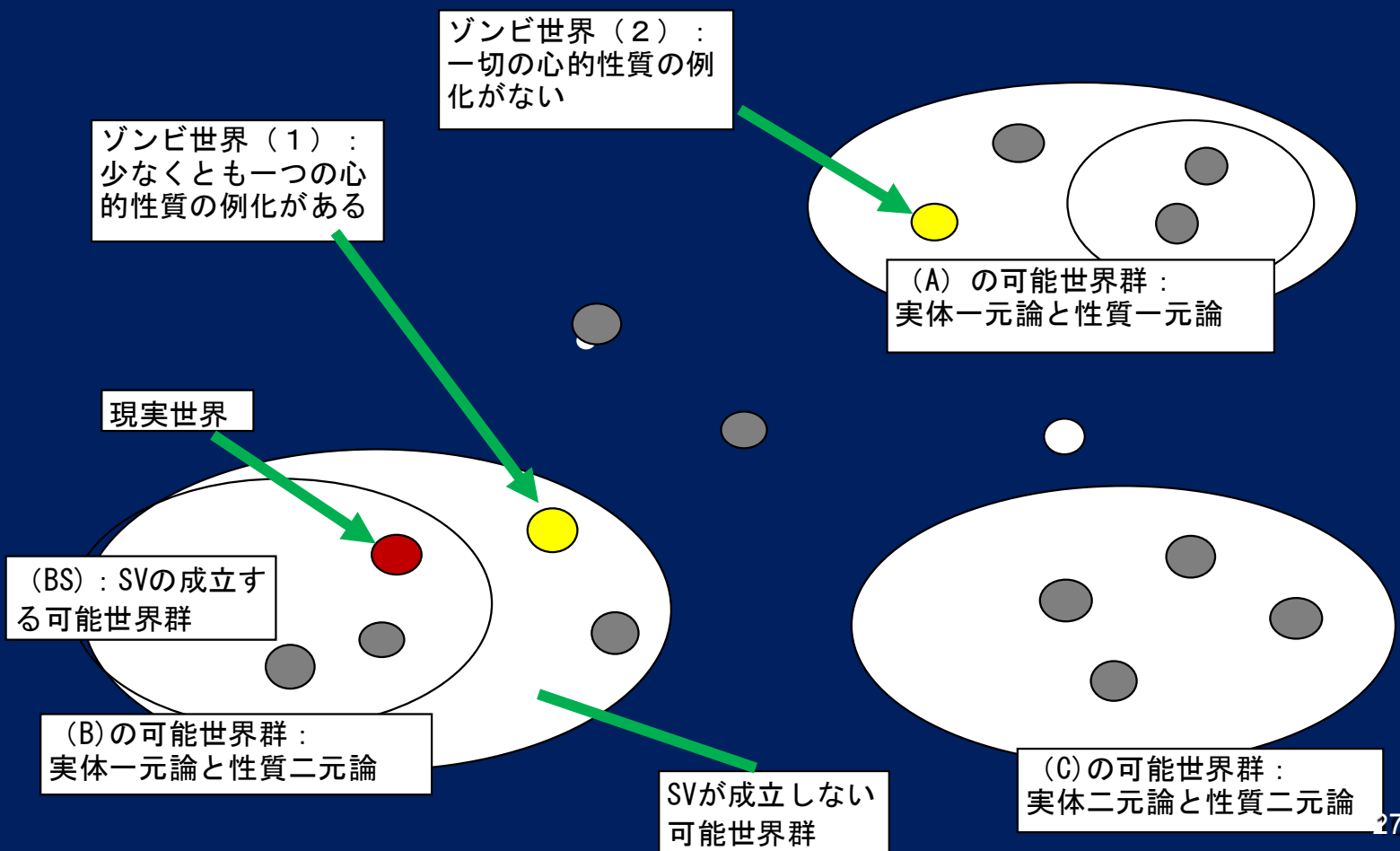
- (B)実体一元論と性質二元論が成り立つ可能世界群。 これらの世界では、例えば、いかなる心的個体も物理的個体と同一だが、いかなる心的性質も自然な物理的性質と同一ではない（現実世界？）。
- (C)実体二元論と性質二元論が成り立つ可能世界群。 これらの世界では、例えば霊的実体がいかなる物理的個体とも同一ではないものとして存在し、かつ、いかなる心的性質も自然な物理的性質と同一ではない（デカルト）。

# 性質二元論的物理主義 (2/2)

- (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ン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



進化の第1幕「自然」から、第2幕「脱自然」へ  
ロボットはこの運命を人間から引き継ぐ

進化における〈適応価〉としての〈美しいもの〉。  
恐らく、自然淘汰に有利な機能を果たすものの一部  
が〈美しい〉と呼ばれた。

やがてロボットへと繋がる「脱自然」の道。  
従来の進化圧を脱した人間の、自然淘汰の残滓と  
しての生得的な脳神経系の「快楽」。  
自然淘汰の新段階が来ないなら、〈美しい〉は生  
物学的根拠さえ失いつつあるのかもしれない。

# 性淘汰によってメスに選択された オスの羽根飾り

<http://ja.wikipedia.org/wiki/%E6%80%A7%E6%B7%98%E6%B1%B0>




# 「すべての美しく輝くもの」

C.F. アリゲザンダー(英国国教会の賛美歌, 1848)  
のパロディ(R.ドーキンス『進化の存在証明』p.315)

すべての醜く冴えないもの  
丈の低きもの地に伏すものもすべて  
粗野なものも嫌なものすべて  
主なる神が全部つくった  
毒牙を打ち込む小さなヘビの一匹ずつに  
針で刺す小さなハチの一匹ずつに  
神は凶悪な毒液をつくり  
その忌まわしい翅をつくった  
癌のようにはびこるすべての病めるもの  
大小のすべての悪  
すべての汚れて危険なもの  
主なる神がすべてつくられた／…(略)  
すべての疱瘡や天然痘  
腐って、汚れて、ただれたもの  
主なる神がすべてをつくられた

# <美>のための、<美>を目指した進化？

## 第2幕： 何に対する進化圧が働くのか？

- 人間の場合、生物個体の表現系に対する自然淘汰は、進化の第1幕で終了した。
  - <美>の領域においては、進化的制約から解放された自由が訪れるかもしれない。
- 
- しかし、その自由は、暗闇からの無根拠な創造なのか？ それとも、何かへの進化圧を介した新たな脳神経系の創造なのか？

# 有限なる存在者がなす〈表現〉の理由

それでもわれわれ人間は、何かを表現したいという  
本質(本能)をもつ。

しかし、表現者に安心と帰依を与えてくれる超越的  
な「理由」は、(たぶん)存在しない。

有限であるがゆえにわれわれは、そのつどの表現  
の「理由」を自ら創らねばならない、という運命を背  
負う。

われわれは、表現者の極私的「理由」と、表現の中  
で出会う。自分の存在の根源的偶然性に出会うと  
きと同じように、不思議さと当惑をもって。



## 結語に代えて(1/2)

ロボットはモーツァルトやダリになれるか？

答え：ロボットが人間的環境で育つなら、もちろんそう  
なれるだろう。

しかし、自律したロボットたちが自分の複製(子孫?)を  
人間と大きく異なる環境で制作し、育てたなら、彼らが  
創る<美しいもの>は、人間から見て<美しくない>  
かもしれない。

そもそも、ロボットたちの多くが賛同する<美しいもの>  
とは何だろう？

## 結語に代えて(2/2)

ロボットのもつ可能性があまりに大きく、また多岐にわたるので、逆に、彼らの持つ「根源的制約」が私には分からない。

そして、あまり言いたくはないが、「根源的制約」のないところに  
＜美しいもの＞と＜醜いもの＞の区別があるのだろうか？  
両者を共に創造したと揶揄された、あの神においてのように。

だが、エンハンスメントやサイボーグを持ち出すまでもなく、この問いは、再び、未来の人類にも向けられる。

さあ、人間もロボットも、ここでは際限もなく自由だ、うんざりするほどに…

# おしまい

柴田の研究関連webサイト

<http://siva.w3.kanazawa-u.ac.jp/>

# 参考文献(ほんの少々)

- Barnbaum, D. R., *The Ethics of Autism*, Indiana University , 2008. 『自閉症の倫理学』(柴田正良・他監訳) 草書房、2013.
- Chalmers, D., 1996, *The Conscious Mind*, Oxford U. P. (『意識する心』林一訳、白揚社)
- Davidson, D., 1980, *Essays on Actions and Events*, Oxford U. P. (『行為と出来事』服部裕幸・柴田正良訳、勁草書房)
- Kim, J., 1993, *Supervenience and Mind*, Cambridge U. P.
- ---- 1998, *Mind in a Physical World*, The MIT Press.
- Heil, J. and A. Mele, 1993 (eds.), *Mental Causation*, Clarendon Press, Oxford.
- Shibata, M., “Toward robot ethics through *the Ethics of Autism*”, in J. L. Krichmar and H. Wagatsuma (eds.), *Neuromorphic and Brain-Based Robots*, Cambridge University Press, pp. 345-361, 2011.
- カス、L. R., 『治療を越えて』青木書店、倉持武(監訳)、2003.
- ハート、H. L. A., 『法の概念』みすず書房、矢崎・他訳、1976.
- 柴田正良, 2006a, 「よみがえったソクラテス-----物理主義と心的因果の問題を理解するために-----」『思想』No.982、pp.4-15.
- 思想, 2006b, 「機能的性質と心的因果-----キムの還元主義を越えて-----」『思想』No.982、pp.53-76.
- ・ ストローソン、P, F., 「自由と怒り」、『自由と行為の哲学』(門脇・野矢監修)、春秋社、2010.