

# <われわれ人間>の本質を創る

…人類とロボットとエイリアンのための十分条件の一つとして…

「道徳的行為者のロボットの構築による<道徳の起源と未来>  
に関する学際的探究」(基盤研究(A) 19H00524)による研究会

於: 中京大学豊田キャンパス

金沢大学名誉教授  
柴田正良 JMar. 14, 2022

# 話の流れ

はじめに (p.3~5)

1. 物理的世界の中での一人称的世界と個性
2. 人間と共生するロボットの条件 (一人称的視点)
3. ロボットとの倫理共同体のイメージ
4. <われわれ人間>の本質を創る試み

はじめに(遺伝子編集の倫理に寄せて)

## 遺伝子編集と個性

1. 遺伝子編集(改変)によって、個性は失われるのか？

答えは「No」。

遺伝子構造の同一性は、物理的同一性より遙か上層にあって「かなり緩く」、また「個性」の遙か下層にあって、個性の基盤であるとしても、個人の経験のすべてを決定できるようなものではないからだ。

2. むしろ問題は、人類が「治療を超えて beyond therapy」どこまで遺伝子編集(改変)をおこなうのか、である。

人類という生物種の同一性を「はみだす」ような、身体も心も異形・異質な個体が際限もなく出現するのか？

## 遺伝子の編集はどこまで許されるか？

- この世界で、人類がどのような存在になりたいかに従って、つまり、人間のアイデンティティをどこに求めるかに従って、どこまでも許される。
- 遺伝子の構造によってアイデンティティを決定するなら、我々は、2億年ほど前の祖先の哺乳類とは別の存在だろう。
- しかし、我々にそのアイデンティティを与える<本質>なるものが存在するなら、我々の遺伝子構造は何でもよい。少なくとも、十分条件としての<本質>をその個体が実現しているならば・・・

## しかし、注意すべきこと・・・自律性

- <われわれ人間>のアイデンティティは「緩やかに」その実現形態の多様性を許すとしても、いずれの存在者も、道徳的主体であるためには、行為と選択における自律性を持っていなければならない。
- では、自律性なき、出生前の胎児に関する遺伝子編集・遺伝子改変は倫理的に許されるのか？

私の暫定的な答え：

その時点での標準的な「治療の範囲内」であれ、「治療を超えて」であれ、その胎児が成人して自律性をもったなら同意するであろう遺伝子改変のみが倫理的に許される。

「自分がまた生まれたとしても、そのとき自分の自閉症を治してほしいとは思わない」、という自閉症者の声。

# アイデンティティ(本質)と個体存在と倫理共同体

- アイデンティティ(本質)を共有することは、もちろん、共有者が同一の個体であることを意味しない。しかし、個体どうしが同一の倫理共同体を形成する前提となる。
- そこで、以下では、無数に複製可能なロボットが、一人称的視点をもつことでいかにして倫理的行為者となりうるか、という問題を考えてみよう。
- それは、ロボットとの倫理共同体から、その前提となる〈われわれ人間〉の本質へと、遡って問うことである。ここで、〈われわれ人間〉とは、人類やロボットやエイリアンなどを含みうる。

## 物理的世界とスーパーヴィーニエンス(SV)

- SVという存在論的原理は、いかなる心的現象にも妥当するが、

我々が生きる場面で直面するのは、この原理の認識もしくは行動のレベルである。そこでは、この原理との「一人称的ギャップ」が存在する。つまり、我々は、三人称的に観察可能な物理的世界にSVしている心的現象のすべてを自分が観察できるわけでもないし、自分の視点から経験できるわけでもない。

「一人称的ギャップ」とは、原理的にすべてがアクセス可能な物理世界の中に、「経験の主体である当人しかアクセスできない領域が存在する」ということだ。

# 行為者の一人称的視点

## 脳が生じさせる「一人称的視点」の特異性

三人称的な世界の中に存在する、それに回収不可能な特異点、「一人称的視点」。それこそが、三人称的世界の中に行為者が出現するための必要条件である。

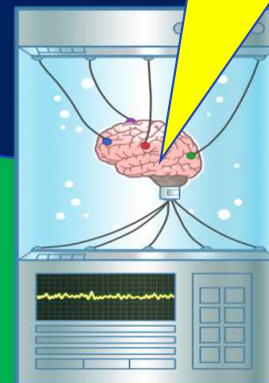
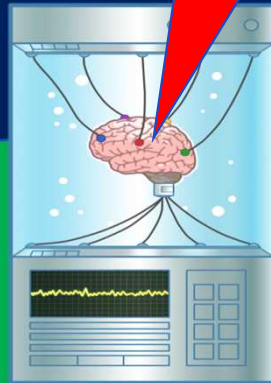
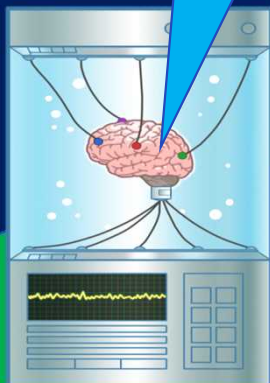
行為者は、単に三人称的に他から特定可能な認知行動能力をもっているだけではない。世界への「一人称的な視点」を持つ必要がある。持っていないければ、それはただの物理的な塊にすぎない。

一人称的視点は、「物理世界への通路」であると同時に、行為者にとって「クオリアと意識から成る一人称的世界」でもある。

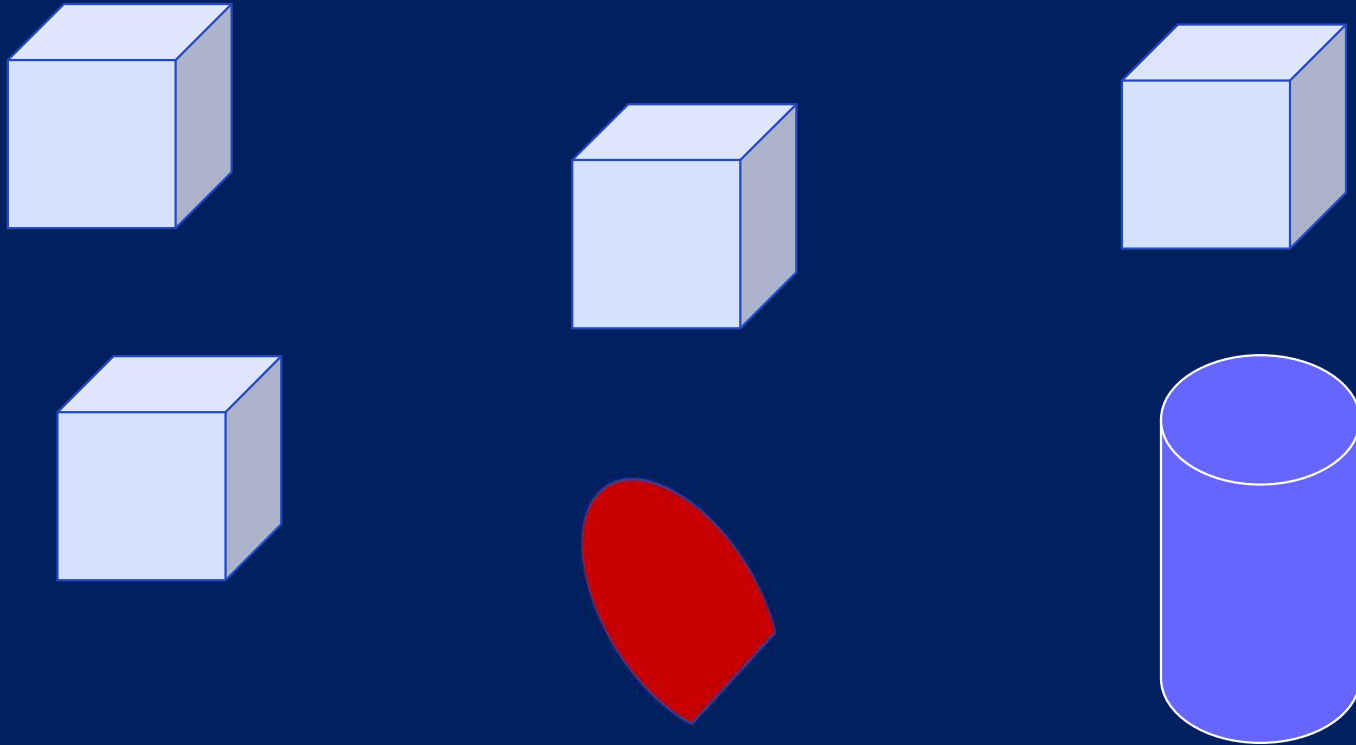


## 行為者の一人称的世界S

- 自分の脳に対する三人称的アクセスは科学的記述だが、一人称的アクセスは、「自分の脳が生み出す世界を自分が生きること」だろう。その世界は、行為者Sにとっての「世界S」である。
- 夢を見ている人物の「夢世界D」を考えてみよう。傍らにいる人々にとって、世界Dに存在する事物は物理的実在性を持たず、物理的に存在しない。だが、夢に導かれて現実世界を彷徨う夢遊病者が経験する世界Dは、ある意味で、世界Sと同様に存在する。
- それは、映画「マトリックス」の登場人物たちが経験させられている、それぞれの仮想世界M1、M2、M3・・・とも同じだ。



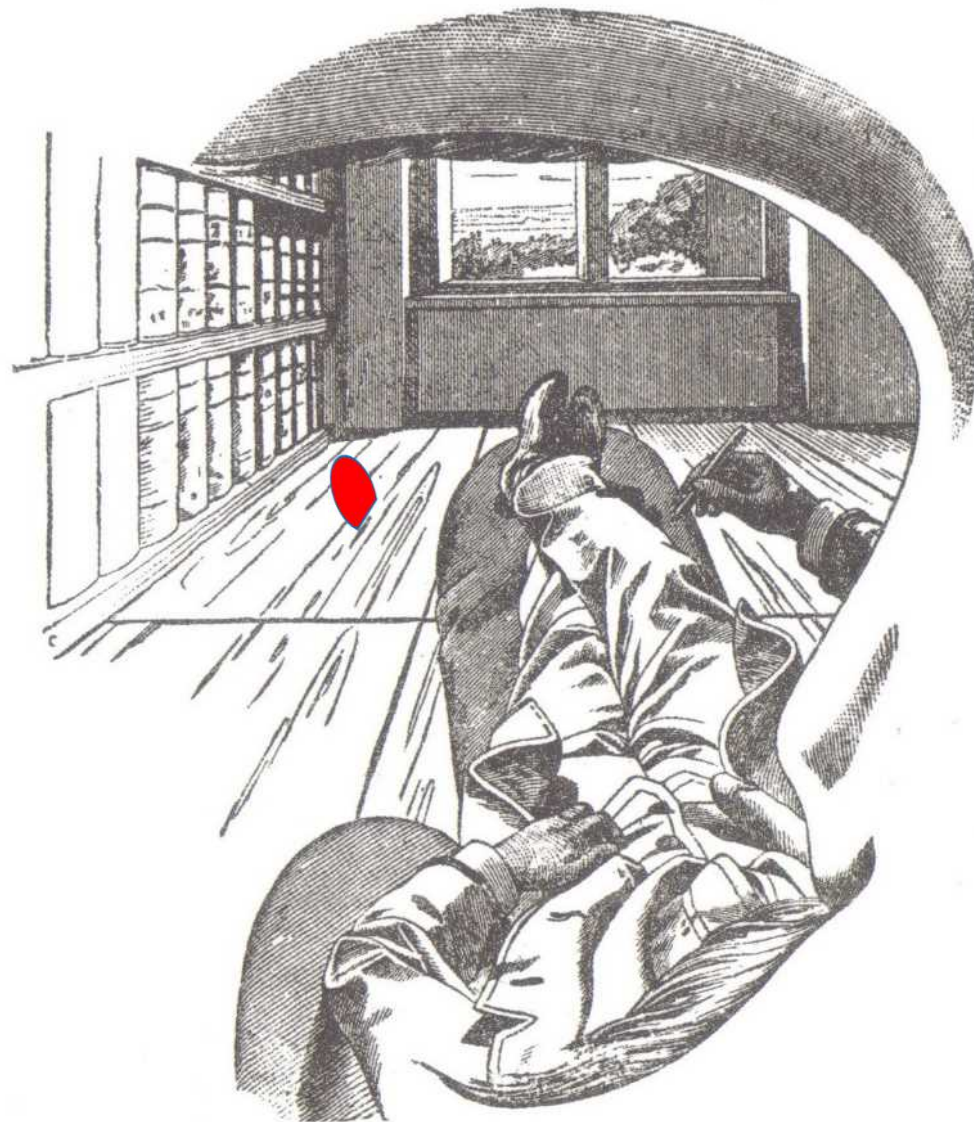
「視点」のない三人称的物理世界  
あなたは自分がどの行為者であるかを  
この状況から選べるわけではない



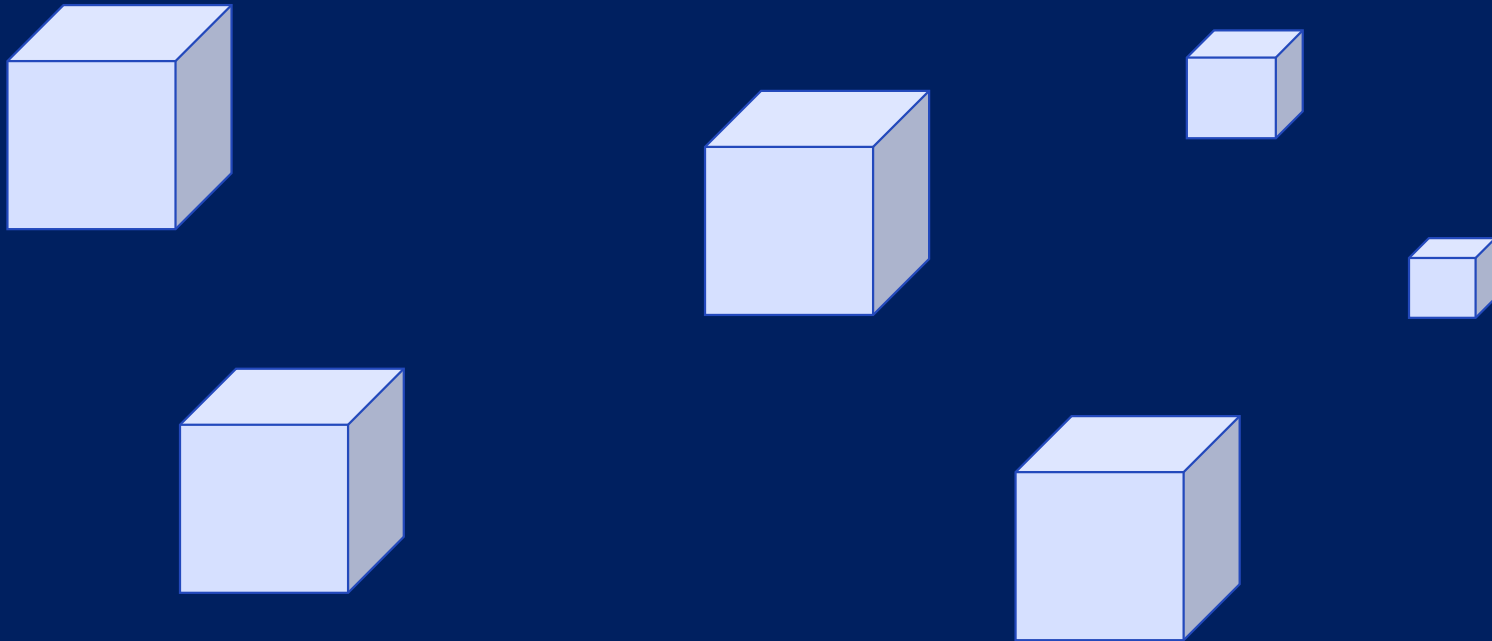
一人称的視点

赤いレモンの  
あるあなたの  
経験世界

E. マツハの自画像



人類の個体がこれほどまでに識別不可能ならば、社会性昆虫のような役割の差異もなく、それゆえ個体間の関係性における進化もなかったであろう。したがって、個体間の差異の調整に必要な自律的「作法」(つまり倫理)も存在しなかったであろう。



## 一人称的視点としての個体における差異、 つまり個性こそがあらゆる目的の源泉である

- われわれの世界では、人類にあっても、自己の遺伝子の最大限の時空的拡散と存続が、個体の究極的な生存目的である(ドーキンス)。  
あるいは、「目的であった」。
- しかし、そのはるか上層で生きているわれわれの社会においては、個人の活動の目的を定めているのは、個人間における差異、つまり個性である。
- 個性は、人が目指し、欲求し、反発し、避けようとするものの原因である。個性相互の相違と衝突を調整するために、様々な社会システムが存在し、倫理と法が存在する。
- そして個性は、素朴心理学的な経験主体であるということ、つまり一人称的な「視点」をもつということの内に、最も先鋭に宿る。

## もしロボットに一人称的視点がなかったら？

- それは、「それ自体が他のために存在するのではなく、それ自体のために存在するような、価値の源泉」ではありえない。(I. カント)
- また、人の存在を時として脅かす「眼差し」の根拠でもありえない。(J. P. サルトル)
- 我々の現在の共生の相手、猫や犬でさえ、ある種の一人称的視点と一人称的世界をもっている。彼らは、「自分の特定の視点」から世界を生き、そこに彼らの運命が結びつけられる。
- そのような視点と行為世界をもたないロボットは、靴や皿やパソコンと同じ、単なる道具にすぎない。



## もしロボットが一人称的視点をもったら？ (1/2)

- 「何らかの仕方で自己接続する特異な物理システム」という存在者、すなわち、三人称的物理世界に出現する「行為者」という一人称的特異点を、現実世界は許容する、としよう。
- 現実世界は、そのような物理的可能世界の一つである。
- そこで更に、SVの成立を前提するなら、一人称的世界が法則的に生じることは保証されるが、どの行為者にとっても、他の一人称的世界はどうあっても文字通りには「**侵入不可**」であり、「**共有不可**」である。
- 行為者世界の、この絶対の「**侵入不可能性**」と「**共有不可能性**」のうちに、行為の「**責任と自由**」が宿る。
- なぜなら、そこに、他の存在者が肩代わりしえない、代替不可能性が生ずるからだ。



## もしロボットが一人称的視点をもったら？ (2/2)

- 一人称的視点と、さらには自律と知性すら持ったなら、そのロボットはまさにパーソン(person)である。つまり、われわれの社会のメンバーとしての行為主体(認知的行為者)である。
- しかし、それが人間に対する悪意や敵意を持っていたら？
- ロボットの方が、いまや知力も体力も人間より上であり、人間より丈夫で、過酷な環境にも耐えられる。
- 多くのSF映画で描かれてきたような、ロボットによる人間支配、人間の殲滅が現実には生じないだろうか？
- つまり、人類の破滅が？

# ロボットとの共生の道

## 人類の破滅を回避する2つのシナリオ

1. ロボットが一人称的視点をもつパートナーとなる手前で、ロボット開発を法的に禁止する。封印された人類の技術として。
  - しかし、これはほぼ確実に失敗するだろう。人間は、違法だとしてもロボット開発を止めないだろう。そういう生き物なのだ、我々は。
2. ロボットとともに、一つの道德共同体を創る。そこで、人間とロボットは原理的に同じ権利と義務をもつ。
  - 幼児や認知症の老人のように、我々は、その道德共同体の中で「劣った」メンバーかもしれない。しかし、この道德共同体が創れなければ、ロボットとの共生の道はない。

## ロボットとの共生の倫理(1/3)

1. 我々は、「信念・欲求・行為」という素朴心理学的概念装置で互いに理解し合っている。少なくとも最小限の素朴心理学的メカニズム(と一人称的視点)をロボットが持つなら、道德共同体に属するのに十分である。したがって、ロボットは、その前提である<われわれ人間>の十分条件をも満たしている。

感受性と反応的態度の程度によって心的存在者としてのタイプが判別され、認知能力の程度によって理性的存在者としてのタイプが判別される。

2. ロボットたちの生存条件が人間とはあまりに異なるので、これまでの人類の倫理・法システムのように、生存条件でその倫理の内容を決めることはできない。むしろ新たな<倫理システム>と<準・倫理システム>を、人為的かつ自覚的に制定する必要があるだろう。

## ロボットとの共生の倫理 (2/3)

3. サイボーグとなり、やがて<機械>に進化した人間のアイデンティティは、どこにあるのか？

それは、素朴心理学が提供する「説明と理解の枠組み」かもしれない。

あるいは、<われわれ人間>の本質・・・

4. 倫理を自由の<欠如態>と考えるなら、理想的状態は、倫理ができるだけ出番を減らすことである。異質なメンバーがどんな目的、欲求、興味、好み、感情を持っていようと、それを最大限に尊重するような倫理システム。

政治哲学的な意味での<リベタリアンの自由>？

## ロボットとの共生の倫理 (3/3)

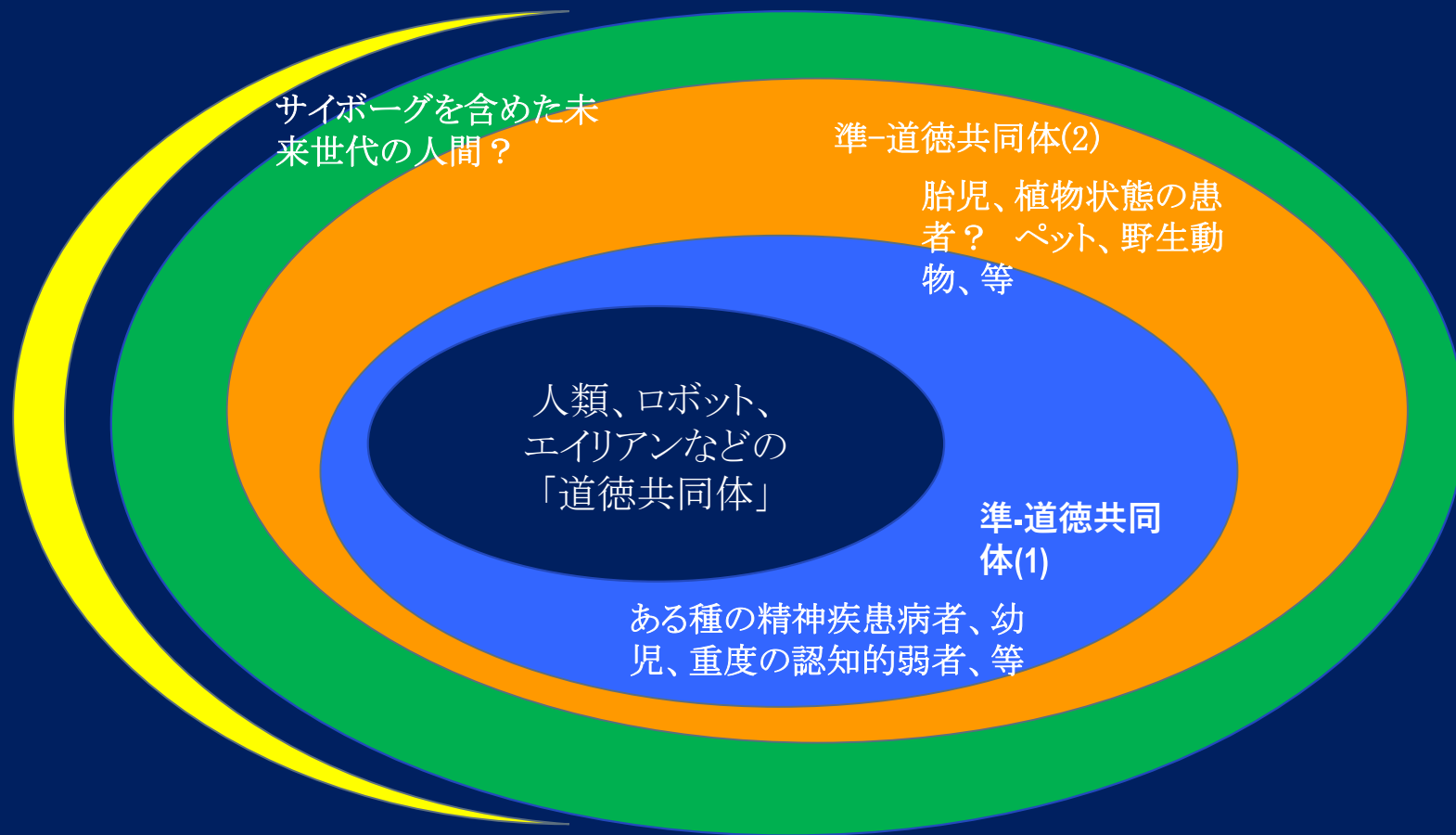
5. 少なくとも、常に最上の人間と同等以上の知力・体力・耐久性・再生可能性を持つサイボーグやロボットは、ふつうの人間とケアの倫理や介護の倫理で結ばれることにはならないだろう。

なぜなら、互いが、自律的な心的存在者としての自由と能力を承認しているのだから。

6. すると極めて皮肉なことに、未来世界の存在者との共生の倫理において妥当するのは、差し当たり、「他者に危害を加えない限り何をしてしても許される」、という古典的な「他者危害の原則」(Principle of harm to others)しかない。

他にいかなる義務と権利があるのか・・・私にはまだ分からない。

# 道徳的共同体と 準・道徳的共同体のイメージ



## 人間のアイデンティティ(本質)を創る！(1/3)

- 人間のアイデンティティを持つなら、その存在者がどのような遺伝子構造をもつていようと、どれほどサイボーグ化され、ついには機械になっていようと、また、始めからロボットとして作られていようと、それは<われわれ人間>だ、と言おう。
- その本質を実現するなら、その存在者の物理的実現形態は何でもよい。

私がとりあえず提案する十分条件としての<本質>の候補は、3つある。それらは、相互に独立した3つの十分条件であるから、<われわれ人間>であるためには、少なくともその1つを満たせばよい。

また、それらの3つ以外にも、独立した別の十分条件があることは排除されない。

## 人間のアイデンティティ(本質)を創る！(2/3)

$\alpha$ . 自由・真理・表現を真に求めるなら、その存在者は<われわれ人間>だ。

ここには自覚的な自律性、真偽の概念的な理解、単なる表出ではない「構造化された言語のような能力」などが前提条件として含まれるだろう。

動物たちは、認知能力の限界ゆえに、この条件を満たさない。

$\beta$ . 信念・欲求・行為という素朴心理学的説明の「対象であると同時にその使い手でもある」なら、その存在者は<われわれ人間>だ。

そもそも、素朴心理学という概念枠組みは、科学的な脳神経心理学には還元されないし、それを越えた部分は、個体存在の目的や意味に焦点を合わせている。その上で、他の個体存在との共生が道德共同体を要請する。



# 人間のアイデンティティ(本質)を創る！(3/3)

$\gamma$ . 一定の範囲内での遺伝子構造の同一性を満たすなら、その存在者は<われわれ人間>だ。

ある時点での幼児も認知症患者も、認知能力の限界や反応的態度の劣化のせいで、道德共同体の正規メンバーにはなれないとしても、この条件を満たすがゆえに<われわれ人間>である。彼らは、準-道德共同体のメンバーとなるだろう。

そこで、ここで意図されている、以上の条件間の論理構造を示しておこう。

MC: 道德共同体の正規メンバーである

FP: 一人称的視点をもっている

「 $\alpha$ 」、「 $\beta$ 」、「 $\gamma$ 」は、先の3条件を表す。

H: <われわれ人間>である

I:  $MC \rightarrow (FP \& H)$

II:  $(\alpha \vee \beta \vee \gamma) \rightarrow H$

## 参考文献

- Barnbaum, D. R., *The Ethics of Autism*, Indiana University Press, 2008.  
『自閉症の倫理学』(柴田正良・他:監訳)、勁草書房、2013.
- カス、L. R., 『治療を超えて』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、「ロボットがフレーム問題に悩まなくなる日」『シリーズ心の哲学:ロボット篇』、勁草書房、信原幸弘(編)、pp.119-174. 2004.
- 柴田正良、「異世界の者たちの倫理」『哲学・人間学論叢』創刊号、金沢大学哲学人間学研究会、pp.17-37, 2010.
- 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.
- 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.

おしまい

柴田の研究関連webサイト

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