

新進研究者 Research Notes

日本科学哲学会

科学基礎論学会

第7号 (2024年)

神崎 祥輝	心はどこまで〈拡張〉されるのか	1
吉澤 日芙美	滑り坂が二つ含まれる滑り坂論法の分析	10
清水 颯	Kantian Dog Argumentとロボット倫理 — ロボットの道徳的地位の擁護に関する批判的考察 —	19
杉山 弦	徳帰結主義についての予備的考察	28
KANG KIWON	「科学への信頼」とは何か?	37
白木 啓吾	『哲学探究』において規則論と言語論はどのような関係にあるのか	45
近藤 雅熙	『論理哲学論考』における独我論の整合的解釈に向けて	54
佐伯 優輔	私的言語論におけるウィトゲンシュタインの一つの声： 243節の読みを手がかりにして	63
駒田 珠希	普遍的指令が「適切に」機能するためには何が必要か	70
森内 康太	外界懐疑論に対するカヴェルの応答	78
中川 和彦	根拠づけと帰結の関係の小考	86
檜岡 寛己	文脈主義による無欠陥な不一致への応答とその射程	94
池田 開	問答的理由説にもとづく規範理由の区別へ向かって	103
安東 里沙子	様々な定式化による算術の無矛盾性証明とGentzenの当初の目標	112

JAPANESE STUDENT RESEARCH NOTES OF PHILOSOPHY OF SCIENCE

Philosophy of Science Society, Japan

Japan Association for Philosophy of Science

No. 7

Contents

KANZAKI Shoki	How far can mind be “extended”? ······1
YOSHIZAWA Hifumi	Analysis of slippery slope argument containing two slippery slopes ······10
SHIMIZU Hayate	Kantian Dog Argument and Robot Ethics —A Critical Examination of the Defense of the Moral Status of Robots— ······19
SUGIYAMA Gen	Preliminary Thoughts on Virtue Consequentialism ······28
KANG Kiwon	What is ‘trust in science?’ ······37
SHIRAKI Keigo	What is the relationship the rule theory and the theory of language in <i>Philosophical Investigations</i> ? ······45
KONDO Masaki	Toward a consistent interpretation of solipsism in <i>Tractatus Logico-Philosophicus</i> ······54
SAEKI Yusuke	A Wittgenstein’s Voice of the Private Language Argument: By Using the Readings of §243 ······63
KOMADA Tamaki	What are Necessary for ‘Happy’ Functioning of Universal Prescriptions? ······70
MORIUCHI Kota	Cavell’s Response to External World Skepticism ······78
NAKAGAWA Kazuhiko	A Brief Consideration of The Relation between Grounding and Consequence ··86
NARAOKA Hiroki	Contextualism's reply to faultless disagreement and its scope ······94
IKEDA Kai	Towards a Distinction of Normative Reasons Based on the Erotetic View of Reasons ······103
ANDO Risako	Consistency proofs for first order arithmetic in various formulations and Gentzen's original goal ······112

心はどこまで〈拡張〉されるのか
How far can mind be “extended”?

神崎祥輝

Abstract

To what extent mind can be extended according to the canonical “Extended Mind Thesis” by Clark and Chalmers (1998) is not clear. Although Clark and Chalmers imply that not only cognitive processes and belief but also other types of mental states can be “extended”, the precise criterion has not been given. This paper shows some traits of mental states that are necessary for the logical possibility of extension and provides some external vehicles that can be seen as “extended” cognitive processes or “extended” mental states.

(1) 研究テーマ

「心はどこで終わり、どこから心なしの世界が始まるのだろうか」(Clark and Chalmers 1998, 7)。Clark and Chalmers (1998)が〈心は頭蓋や皮膚を越えて外的環境へと拡張する〉という〈拡張された心〉理論(Extended Mind Thesis)を提唱して以来、心の〈拡張〉は三者三様に論じられてきた。その結果、EMTは反対派から数多くの批判が向けられるのみならず、擁護派の間でも〈拡張〉の概念や〈心〉の概念の内実をめぐる見解が大きく割れるようになってしまった。もはや、EMTについての統一見解は存在しない。

本稿ではまず先行研究の整理として、Clark and Chalmers (1998)におけるオリジナルの EMT——記号計算の EMT と信念の EMT——の議論の流れを確認する。そのうえで本稿独自の主張として、Clark と Chalmers の議論に基づき、〈拡張〉の論理的可能性が認められる、その他の心的状態のタイプを検討する。最後に今後の展望として、〈拡張〉が実現されうる具体的な環境媒体についても軽く言及する。

(2) 研究の背景・先行研究

Clark and Chalmers (1998)において EMT は、まず記号計算の EMT、次いで信念の EMT という順番で導入される。しかし、Clark と Chalmers は記号計算の EMT と信念の EMT で満足しているわけではない。明示的な議論こ

そなされてはいないものの、彼らは同論文内で欲求や情動も〈拡張〉されう
ることを示唆しており、EMT があらゆるタイプの心的状態に適用されること
を期待しているようにも窺える (Clark and Chalmers 1998, 12)。

記号計算の EMT と信念の EMT についての個別の論証に移る前にまずは、
Clark と Chalmers の全体的なスキームを概観しておきたい。H. Putnam や
T. Burge のような (内容) 外在主義とは対極的に、Clark と Chalmers は機
能主義に立脚して心の外在性を訴える。本稿では (内容) 外在主義につい
ての詳細な説明は省略するが、一言でいえば (内容) 外在主義は機能主義、す
なわち行動に対する因果的機能のみに基づく心の説明理論を拒絶する (cf:
Putnam 1975; Burge 1986)。それと対極的であるとはつまり、EMT は因果
的機能に基づきつつも、心が個体の脳内や身体内にとどまることなく、外的
環境中においても実現されうることを主張するのである。

Clark と Chalmers は以下のような思考実験を用いることで、記号計算の
EMT を論じようとする (Clark and Chalmers 1998, 7-8)。

- (1) Inga が、様々な二次元幾何学図形を投影したコンピュータースク
リーンを眺めており、画面上の〈ソケット〉にピッタリと当てはま
る図形を選択するように問われる。Inga はソケットに図形が当て
はまるように、頭の中で図形を回転操作しなければならない。
- (2) Otto が、同様のコンピュータースクリーンを眺めているが、今度
は手元のコントローラーの回転ボタンを押して実際に画面上の図形
を回転させるか、頭の中のイメージを(上の(1)と同様に)回転させ
るかを選択することができる。
- (3) Otto が、同様のコンピュータースクリーンを眺めている。Otto は
しかし、脳内に特殊な神経インプラントを埋め込まれており、脳内
でこの神経インプラントを(2)のケースにおける回転ボタンのよう
に機動させるか、自力で頭の中のイメージを回転させるかを選択す
ることができる。

Clark と Chalmers によれば、以上の 3 つのケースにおいて、脳内での (心
的表象としての) 図形の回転操作・回転ボタンを用いた (画面上の) 図形の
回転操作・神経インプラントを用いた回転操作は全て、〈ソケットに図形が
はまるように回転させる〉という同じ機能的役割を果たしている。そのため機
能主義の観点からは、体表や頭蓋骨といった物理的境界の内側での操作—
心的表象を用いた操作—も、外側での操作—スクリーン上の表象を用い

た操作——も、等しく真正の記号計算とみなすべきである (Clark and Chalmers 1998, 7)。

記号計算の EMT に続いて展開される、Clark と Chalmers による信念の EMT の議論も見てみよう (Clark and Chalmers 1998, 12-4)。

- (4) Inga は、MoMA で開催される展覧会のことを耳にした。彼女はその展覧会を訪れることを決めた。彼女は〈美術館が 53 番街にある〉ということを出し、歩いて美術館へ辿り着いた。
- (5) Otto は認知症に苛まれており、自身の脳内記憶の代わりとしてノートに常に携帯している。彼は新たな情報を得るたび、それをノートに記録する。Otto は、MoMA で開催される展覧会のことを耳にした。彼はその展覧会を訪れることを決めた。自身のノートを参照し、そこに〈美術館は 53 番街にある〉と書かれていたので、歩いて美術館へと辿り着いた。

1 番目のケースにおいて Inga は、〈美術館は 53 番街にある〉という信念を抱いており、その信念が〈美術館に行く〉という行動を因果的に引き起こしたと捉えられる。同様に 2 番目のケースにおいても、〈美術館は 53 番街にある〉というノート上のメモ書きが、〈美術館に行く〉という行動を因果的に引き起こしている。つまり機能主義的な観点からいえば、Inga の内的な信念も Otto のノート上のメモも、同等の因果的機能を担っている。そのため Clark と Chalmers によれば、ノート上のメモは Otto 自身の信念に他ならない。両者の違いは単に、心的状態が個体の内部で実現されているか、個体の外部で実現されているかという、表層的なものにすぎないというわけである。

Clark と Chalmers は、脳内情報と環境内情報へのアクセスの容易さの違いに依拠した、信念の EMT への想定反論もあげている。Inga はいつでも安定した形で自身の信念を参照できる一方で、Otto は彼のもつノートを紛失したりその記述内容を読み間違えるリスクがあるため、同じ機能を果たしていない。Clark と Chalmers はこのような想定反論に対して、以下のように応答する (Clark and Chalmers 1998, 15)。もし Otto にとってノート上のメモが大抵は信頼に足る形でアクセス可能ならば、(紛失や読み間違いによって) 信頼性やアクセス可能性が損なわれるリスクを孕んでいても問題にはならない。なぜなら Inga の信念もまた、外傷や酔酩などに伴い信頼性やアクセス可能性を損なうリスクを孕んでいるからである。Inga の信念も、〈大抵は

信頼に足る形で容易にアクセス可能〉にすぎない点は、Otto のノートと変わりない。すなわち信頼可能性やアクセス可能性という基準からいえば、心的状態と外的環境はやはり機能的に同等であり、信念の EMT が成り立つ。

信頼可能性やアクセス可能性が同等に担保されていれば機能的に同等であるといえるのか、心的状態と外的環境にそれ以外の根本的な違いはないのかなど、信念の EMT に対してはさらなる懐疑も突きつけることができるだろう。しかし本稿では、Clark と Chalmers の論証の妥当性についての検討は重ねない。たしかに、EMT の論証プロセスに対する批判は枚挙にいとまがないが、本稿の目的は EMT 自体の妥当性を問うことではない (cf: Adams and Aizawa 2001; Adams 2010)。本稿の目的は、記号計算の EMT と信念の EMT が正しいと仮定した際に、それらの道具立てを用いて〈拡張〉を示すことができる心的状態のタイプは他に何があるか、という EMT の射程の深さを測ることである。これは、EMT が真であるか否かとは独立の議論だろう。

(3) 筆者の主張

本稿では、記号計算の EMT と信念の EMT が成立していると仮定した上で、同じ説明構図でその他の心的状態にも EMT が可能であることを論じる。

(1) 認知的プロセス。記号計算の EMT を認めるならば、その他の認知的プロセスも同様に EMT を認めることができるだろう⁽¹⁾。認知的プロセスにはたとえば、学習や推論、意思決定などが含まれる。

(2) 志向的内容と機能をもつ心的状態。以下で示す通り、信念同様に志向的内容と機能をもつ心的状態であれば、信念同様に環境への〈拡張〉が想定可能だ。まずは機能主義的な観点から、信念という心的状態が備えている性質を考える。

心の哲学において信念は、〈志向性〉をもつ心的状態のひとつに含まれる (金杉 2007, 93-4)。志向性は〈ついて性 (aboutness)〉とも呼称される通り、〈何か他のものについてのものである〉という性質のことである。信ずるということは、〈何か〉を信ずることであり、内容を欠き空虚に信ずるということとはできない。たとえば、〈家の外で雨が降っている〉という信念は、〈外の天気〉についてのものであり、それを志向的内容として必然的にもつ。

そして、機能主義的に捉えた信念の重要な特徴として、特定の行動を因果的に引き起こすという点を無視することはできない。前節で Clark と Chalmers による信念の EMT を確認した通り、〈美術館は 53 番街にある〉という信念は、〈美術館に行く〉という行動を因果的に引き起こしていた⁽²⁾。

以上をまとめると、(1) 志向的内容をもつことと、(2) 因果的機能をもつ

ことのふたつが、信念の重要な性質である。心的状態の EMT を認める上で、このふたつは必要な条件であるといえる。

では、内容と機能の双方をもつ心的状態として、信念のほかにどのようなものが挙げられるだろうか。欲求や情動はそうした心的状態に含まれる。たとえば〈林檎を食べたい〉という欲求は、〈林檎〉を内容に、〈林檎を食べる〉という行動を引き起こす機能をもつ。同様に、〈蛇が怖い〉という恐怖の情動も、恐怖の対象である〈危険〉を内容に、〈危険から逃げる〉という行動を引き起こす機能をもつ。以上より、欲求や情動、それ以外でも内容と機能をもつ心的状態ならば、信念と同様に EMT を認めることができると考えられる。

(3) 志向的内容や機能をもたない心的状態。心の哲学上の一般的な見解に則れば、あらゆる心的状態が内容をもつわけでもないし、機能をもつわけでもない。はたして、内容や機能をもたないとされる心的状態、たとえば憂鬱なムードや痛みを感じるについて、EMT を認めることはできないのだろうか。

内容をもつことと機能をもつことを EMT の必要条件と捉えるならば、そうした心的状態の EMT を認めることはできないだろう。しかし、憂鬱なムードや痛みを感じるといった、一般には内容や機能をもたないとされる心的状態が、内容や機能を本当にもたないのかを問い直す余地はある。もしこれらの心的状態が内容や機能をもつならば、EMT は論理的に可能である。

〈志向説〉によれば、あらゆる心的状態は内容をもつ（金杉 2007, 95-6）。たしかに一見すれば、憂鬱なムードは、何か特定の事態や対象についての心的状態ではない。それは何にも向けられていない単なる漠然とした感情のようである（金杉 2007, 96）。痛みを感じるについても同様で、それは痛みについての独特な感じであって、何か特定のものに向けられたものではないように思える。しかし志向説によれば、憂鬱なムードは〈世界全体〉あるいはそうした世界を感じる〈自分自身の身体〉に対して向けられた心的状態であり、痛みを感じるもまた〈自分自身の身体状態としての痛み〉に対して向けられた心的状態である。であるならば、これらの心的状態も内容をもっており、内容と関連した行動を因果的に引き起こす機能をもつという考えは不自然ではない。つまり、〈林檎を食べたい〉という欲求が〈林檎を食べる〉行動を、〈蛇が怖い〉という情動が〈蛇から逃げる〉行動を引き起こすように、憂鬱なムードはたとえば〈世界から距離を置く〉行動を、痛みを感じるはたとえば〈ダメージを受けた身体箇所に対して注意を向ける〉行動をそれぞれ引き起こす因果的機能をもつとみなせるかもしれない。

以上をまとめると、内容や機能をもたないと通常解される心的状態も、志向説に基づけば、〈世界全体〉や〈自分自身の身体〉のように内容を広く捉え

ることで、内容や機能を帰属できるようになる。ただし、〈あらゆる心的状態は内容と機能に還元できる〉という強固な志向説を採用しない限り、痛みの感覚や憂鬱なムードが機能的側面に還元されない、その心的状態に内在的な現象的側面を認める余地は残る。本論ではこのような強固な志向説の立場は取らない。しかし、〈あらゆる心的状態は内容と機能を持つが、それらに還元されない内在的性質も持つ〉という弱められた志向説においても、これらの心的状態の現象的側面に関してはともかく、機能的側面に関しては少なくとも、Clark と Chalmers の議論に基づいて EMT を認めることができるだろう。

(4) (自己) 意識。これまで扱ってきた心的状態は一貫して、意識が意以外の別の対象に向けられる構造をしていた。というのも、〈外で雨が降っている〉という信念や〈林檎を食べたい〉という欲求はいずれも自己の外に存在する事態や対象へ向けられた心的状態であり、憂鬱なムードや痛みの感覚も、世界全体や自分の身体の一部といった、やはり自己意識からその他の対象に向けられたものであったからである。

しかし、ほかならぬ自分自身に対して向けられた心的状態、すなわち意識の EMT を認めることはできないのだろうか。意識の EMT を認めるということは、言い換えれば、〈拡張された心 (Extended Mind)〉ならぬ〈拡張された自己 (Extended Self)〉を認めるということではないだろうか。実際のところ、Clark と Chalmers は、同論文の結部で〈拡張された自己〉の可能性を示唆している。

拡張された心は拡張された自己を示唆するのだろうか？ そのように思われる。多くの人は、既に自己というものが意識の境界線を越え出ていることを受け入れている。…Otto のノート上の情報は、たとえば、認知的エージェントとしての彼のアイデンティティの中心的な一部である。このことが導くのは、Otto 自身を、生物的有機体と環境資源の融合からなるひとつの拡張されたシステムとみなすのが最善であるということである (Clark and Chalmers 1998, 18)。

ただし上の引用箇所において注意が必要なのは、〈自己というものが意識の境界線を越え出る〉というように、Clark と Chalmers は自己と意識を概念的に区別している点である。つまり彼らの議論に則る限りにおいて、〈拡張された自己〉は〈拡張された意識〉を含意しない。それどころか同論文内のその他の箇所では、〈拡張された自己〉と〈拡張された意識〉が両立不可能であ

ることを示唆するような記述も窺える。

経験のような、いくつかの心的状態は、内的に決定されるのかもしれないが、他の〔信念のような心的状態の〕事例においては、外的な要因が重要な貢献を果たしている。…これらふたつの事例〔=信念の EMT における Inga と Otto の事例〕において情報は、意識にとって利用可能なように、また行為を導くのに利用可能なように、まさに我々が信念に期待する通りの信頼可能な形でそこに存在しているのだ (Clark and Chalmers 1998, 12-3、〔〕内は筆者の補足)。

これらの箇所では Clark と Chalmers は、経験すなわち意識のような心的状態は内的に個別化され、情報すなわち心的内容は意識が利用するという、極めて内在主義的な描像を打ち出している。つまり EMT は、環境媒体では機能的に代替不可能な意識の存在を前提している。ゆえに、志向説および機能主義的な観点から意識の EMT を論ずるのは、不可能だと考えられる。

(4) 今後の展望

本稿では、さまざまな心的状態のタイプに関して、Clark と Chalmers の議論に基づいて EMT を認めることが論理的に可能か否かを検討した。その結果、現象的意識以外の心的状態の EMT は論理的に可能と結論付けられた。しかし、EMT が論理的に可能であることと、EMT が実際にどのような環境媒体を用いて実現されるのかというのは、全く別の問題である。今後の研究課題としては、〈拡張〉の論理的可能性が示された各種の心的状態に関して、それらの〈拡張〉先となる具体的な環境媒体の候補を検討したい。

本稿で取り上げた各種の心的状態の EMT を現時点で網羅することはできないが、たとえば次のような環境媒体を考えている。推論の EMT には、Clark (2008)における(ただし、Clark はこれを記憶の EMT として挙げている)熟練のバーテンダーの例が当てはまるかもしれない (Clark 2008, 62)。Clark によれば、バーテンダーは、様々な特徴的な形状のグラスをカウンター上に順に配置しておくことで、どの客にどのカクテルを提供すべきかをいちいち脳内で推論せずに済むように工夫している。カウンター上のグラスの配列という環境媒体に〈拡張〉された推論とみなすことができるかもしれない。

欲求についてはどうだろうか。欲求の因果的機能を、対応する信念とペアとなって行動を動機づけることと捉えれば、To Do リストのような環境媒体を〈拡張された〉欲求とみなせるかもしれない。たとえば、私の内側に〈原

稿を書き上げたい」という心的な欲求がなくとも、To Do リスト上には〈原稿を書き上げたい〉と書いてあり（＝〈拡張〉された欲求）、私自身が〈原稿の締切は今日だ〉という心的な信念を抱いていれば、これらがペアとして〈原稿を書き上げる〉という行動を引き起こす。

情動の EMT に関しては、たとえば、村山（2022）において挙げられている Munch の『叫び』や Taylor Swift の“Red”のような創造的な芸術作品は、製作者ないし鑑賞者の内的な情動と同等の機能をもつ〈拡張〉された情動とみなせる環境媒体であるかもしれない（村山 2022, 4）。

もともと、ある環境媒体が〈拡張された〉心的状態と認められるためには、心的状態との機能的な同等性に加えて、十分なアクセス可能性と信頼性も必要であることを、Clark と Chalmers は信念の EMT への想定反論に対して言及していた。上で列挙した環境媒体が十分なアクセス可能性と信頼性を有しているのか、そのためにはどのような条件をクリアしていなければならないのかという点も、EMT の実現可能性を考えるうえで重要な論点だろう。

注釈

- (1) 本稿では簡便化のために認知的プロセスも心的状態として取り扱っているが、認知的プロセスを心的状態に含めてよいかは疑問である。Clark と Chalmers は、記号計算は ECT(Extended Cognition Thesis)、信念は EMT (Extended Mind Thesis) と表記を分けていることから、厳密には認知的プロセスを心的状態に含めないと考えられる。
- (2) ここでの主張は、信念が単体で行動を因果的に引き起こしている、というものではない。〈美術館に行く〉行動を引き起こすには、〈美術館は 53 番街にある〉という信念に加えて、〈美術館に行きたい〉という欲求を抱いている必要もある。いずれにせよ、信念が行動を引き起こす因果的機能の一端を担うことが示せばここでは十分である。

(5) 参考文献

- Adams, F and Aizawa, K, 2001, “The Bounds of Cognition”, *Philosophical Psychology*, 14-1, 43-64.
- Adams, F, 2010, “Why We Still Need a Mark of the Cognitive”, *Cognitive Systems Research*, 11(4): 324–331.
- Burge, T, 1986, “Individualism and Psychology”, *The Philosophical Review*, 95(1),3-45.
- Clark, A. and Chalmers, D, 1998, “The Extended Mind”, *Analysis*, 58-1, 7-19.

Clark, A, 2008, “Supersizing the Mind: Embodiment, Action, and Cognitive Extension”, Oxford University Press.

Putnam, H, 1975, “Mind, Language, and Reality”, Cambridge University Press.

金杉武司, 2007, 『心の哲学入門』, 勁草書房.

村山正碩, 2022, 『表出性と創造性：表出説を改良する』, 新進研究者 Research Note, 第 5 号: 11-18.

(一橋大学)

滑り坂が二つ含まれる滑り坂論法の分析

Analysis of slippery slope argument containing two slippery slopes

吉澤 日英美

Abstract

This paper has two objectives: (1) an introduction to Walton's analysis of slippery slope argument (SSA) and (2) suggestions to supplement Walton's analysis. (1) Walton provides a basic argumentation scheme of the SSA. According to Walton, the SSA could be basically a reasonable form of argument only if it fits into the scheme. (2) However, Walton's schema overlooks a phenomenon concerning SSA that contains two sequence elements (the case is that there are two slippery slopes in a SSA). This paper argues that the reasonableness of such a type of SSA depends on how we combine two slippery slopes in presenting the type of SSA.

(1) 研究テーマ

滑り坂論法 (Slippery Slope Argument、以下 SSA) とは、生命倫理のなかでも、中絶や安楽死、遺伝子操作などの是非を問う文脈でよく用いられる論証である (den Hartogh, 2009, p.321)。例えば、児玉は「安楽死を合法化すると、死にたくない人までが意に反して安楽死させられるという濫用が生じる。ゆえに、合法化すべきでない」(児玉, 2017, p.22) を SSA の例として挙げている。多くの SSA は、最初の前提を受け入れた結果、その前提が招く最終的な帰結や将来に波及する社会的影響を受け入れることができないため、最初の前提を実現するべきではないと提示する。

SSA は、ある行為や政策の道徳的な正しさを示すための議論ではなく、ある行為や政策を許容してもよいかの「決定」に関する議論であるとされる¹(van der Burg, 1991, p.64)。児玉が挙げる SSA の例では、「安楽死を合法化する」ことの「決定」について議論をしているのであり、「安楽死を合法化する」ことそれ自体が道徳的に正しいかに関しては、別の議論が必要となる。

SSA はしばしば誤謬推論だとされてきた。その理由として、堆積のパラドックスとの類似が指摘されている (den Hartogh, 2009, pp.323-325)。砂山と平地のように連続した概念の間に線引きをしようとする、その線引きは恣意的なものになる。だから両者を同列に扱わなければならないという結論が導かれるのだが、この結論は直観に反している。例えば、受精卵と胎児は発生過程が連続しているため、両者の間に線引きをしようとするとその恣意

性に悩まされることになる。だからといって、両者を同列に扱うという結論を受け入れると、「受精卵を研究対象にできるのならば、胎児も研究対象にできる」などの主張を認めることになり、これは受け入れがたいように思われる。そのため、SSAは誤謬だと説明される。

他方で、WaltonはSSAについて、「ある行為をとるかどうかが考えている人が、その行為の帰結が破滅的なものであることに合意した場合に、その行為について再考するよい理由を提供できるかもしれない」（Walton, 2017, p.1515）と述べて、「SSAは誤謬推論だ」という一般的な見解を払拭しようとする。WaltonはSSAの基本的なスキーマを提示し、前提がスキーマに合致するか、前提の正しさを示す証拠はあるかなどについて検討できることを示す。なお、後述するように、WaltonのSSAは行為を対象にしており、最終的な帰結には道徳的な不正さや否定的な価値づけが含まれる²。SSAは対話相手を説得しようとする場面や、政策決定について議論している場面で用いられる。ただし、論者たちの間にSSAの統一された定義があるわけではない。以下では、SSAとは行為に関する論証であり、帰結に道徳的な不正さや否定的な価値づけが含まれるものを指すこととする。

しかし、Waltonによるスキーマが対象とするSSAは、滑り坂が一つ含まれる論証である。筆者の見立てでは、一つの論証に二つの滑り坂が含まれるSSAが存在し、そのような論証はWaltonによるスキーマでは検討しきれない。そこで、二つの滑り坂が含まれるSSAの分析についても考察する。

(2) 研究の背景・先行研究

この章では、主に2015年のWaltonの論文に基づいて、SSAの分析を紹介する。多くのSSAは、最初の行為と最終的な帰結だけが特定されており、議論の前提は暗黙的に仮定されていることが多い。児玉が挙げるSSAの例も、一見すると何が連続しているのかが分からない。そこでWaltonは、SSAと否定的な帰結からの論証（Arguments from Negative Consequences、以下AFNC）とを区別する。

AFNCとは、「大麻を合法化すると慢性呼吸器疾患が増加する。ゆえに、大麻の合法化に反対である」のように、受け入れがたい帰結をもとに結論を導く。SSAは、AFNCにおける前提と帰結との間に連続する要素をもつ。SSAとして提示されたものの多くはAFNCの形をしており、提示されたAFNCがSSAに展開できるどうかは、暗黙の前提を明らかにし、Waltonのスキーマを使って評価できる。Waltonによる基本的なスキーマは以下の通りである（Walton, 2015, p.288）。

最初の前提：行為者は行為 A_0 を実行することを考慮している。

連続した前提： A_0 を実行することは A_1 を導き、またそれは A_2 を導き、というように $A_2, \dots, A_x, \dots, A_y, \dots, A_n$ と連続している。

不確定の前提：一連の行為 $A_0, A_1, A_2, \dots, A_x, \dots, A_y, \dots, A_n$ はグレーエリアと呼ばれる A_x, \dots, A_y を含んでいる。 x と y は不明確な点である。

コントロールの前提：行為者は連続している行為の実行を止めることについて、行為者が A_x, \dots, A_y のグレーエリア内のどこか不確定な点に到達するまでは、コントロールを保っている。

コントロールの喪失の前提：一度でも行為者がグレーエリア A_x, \dots, A_y 内の不確定な点に到達すると、行為者はコントロールを失って、 A_n に到達するまで行為し続けることを強制される。

危機的な帰結の前提： A_n は、可能ならば避けるべき危機的な帰結である³。

結論： A_0 は実行されるべきではない。

SSA には、帰結までの一連の流れをそれに沿って推進する要因があり、それをドライバーとよぶ。ドライバーは、「行為が継続することへの抵抗を徐々に困難にさせるもの」(Walton, 2015, p.288) でもある。提示された SSA を検討するには、スキーマに挿入される前提を明らかにして、それぞれの前提が合致するかを確認するだけでなく、連続する要素の連鎖性とコントロールの喪失を示す証拠としてドライバーを特定することが必要となる。ドライバーとして、前例、社会的受容、曖昧さなどが挙げられ、SSA 内に一個以上含まれることもあれば、異なるドライバーが論証内の異なる部分で働くこともある。

Walton は合理的な SSA の例として、ヘロインを摂取しないように娘を説得する父親ボブの論証を挙げる。ボブは「ヘロインを摂取すると、人生が台無しになる。だから摂取するべきではない」(AFNC) や、「ヘロインに手を出すと、繰り返し摂取してしまう」(SSA) と言って娘を説得しようとするが、娘は「飛躍している」、「SSA は誤謬だと習った」と言って納得しない。このとき、AFNC に含まれる暗黙の前提を明らかにし、SSA のスキーマに合致す

ることができれば、父の言葉は合理的な SSA であることを示せるため、娘を説得できるかもしれない。Walton が示すボブとアリスの SSA は以下の通りである⁴ (Walton, 2015, pp.289-290)。この SSA ではドライバーが二つ含まれる。

最初の前提：娘はヘロインを摂取するか否かを考えている。

連続した前提：目に見えた危害がなければ、繰り返し服用し続けるだろう。

▶ドライバー：ヘロインを摂取すると、強い快感が得られるので、再度摂取したくなる（精神依存）。

不確定の前提：簡単に依存症になる人もいれば、そうでない人もいる。

コントロールの前提：いつの時点で薬物の摂取を中止するかについて、当初はコントロールできる。

コントロールの喪失の前提：一度でも依存状態になってしまえば、コントロールを失い、薬物の摂取を中止できない。

▶ドライバー：一度身体が薬物に依存してしまうと、離脱症状によって、薬物の減量や継続中止が困難になる（身体依存）。

危機的な帰結の前提：この結果は、人間関係、経済的信用、健康と名誉を失い、はては人生全体を損なう壊滅的なものなので、避けるべきである。

結論：ヘロインを摂取するべきではない。

この SSA では、連続した行為として薬物の単回摂取、グレーエリアとして薬物依存の個人差、ドライバーとして薬物による精神依存と身体依存を挙げ、ドライバーによってコントロールの喪失が生じることを示している。Walton は、「重要なことは、連続した行為が、ドライバーによって、グレーエリアを越えて最終的な帰結に向かわされるというコントロールの喪失があることである」とし、「このようにモデル化された SSA は、その強力さと正当性が正確に理解できる合理的なものになる」と述べている⁵ (Walton, 2015, pp.297-298)。

(3) 筆者の主張

前章では、Walton による SSA の分析を紹介した。Walton によるスキーマにおける「最初の前提」である「行為者は行為 A₀ を実行することを考慮している」から明らかかなように、連続する要素は行為 A のみである。しかし、連

続する要素は一つとは限らない。例えば、先の父親と娘が日本在住だったとして（父親をヒロシとする）、ヒロシが「危険ドラッグ⁶に手を出したら、違法薬物中毒になってしまう」と発言したとする。このとき、ヒロシが言いたいことは「危険ドラッグを一度摂取したならば、違法な薬物であるコカインやヘロインを繰り返し摂取するようになるだろう。違法な薬物であるコカインやヘロインを繰り返し摂取することは、あなたの人生にとって大変な損失である。したがって、危険ドラッグは1回でも摂取するべきではない」ということだ。このとき、娘が危険ドラッグの摂取を考慮している状況は、薬物の種類が変わることと、薬物の摂取回数が増えることという別々の展開の組み合わせとして捉えることができる。つまり、「最初の前提」は、「行為者は対象物 A₀に対して行為 B₀を実行することを考えている」となり、連続する要素として、対象物 A〈ある危険ドラッグ、…、コカイン、ヘロイン〉と、行為 B〈1回目の摂取、2回目の摂取、…、N回目の摂取〉が挿入されている。したがって、ヒロシが述べていることは、二つの連続する要素が含まれる、いわば、滑り坂が二つ含まれる SSA として分析できる⁷。以下で詳述しよう。

ヒロシの「危険ドラッグを一度摂取したならば」の後には、連続する要素ごとに滑り坂が展開されることになる。すなわち、①「危険ドラッグを一度摂取したならば、危険ドラッグを繰り返し摂取するだろう」(行為 B の展開)、ならびに、③「危険ドラッグを一度摂取したならば、ヘロインも一度摂取するだろう」(対象物 A の展開)を提示することができる。この二つの SSA が、ヒロシの懸念の帰結である「ヘロインを繰り返し摂取するようになるだろう」と同様の主張をするには、①「危険ドラッグを繰り返し摂取するだろう」の後に、②「危険ドラッグを繰り返し摂取したならば、ヘロインも繰り返し摂取するだろう」(対象物 A の展開)が接合することになる。同様に、③「ヘロインも一度摂取するだろう」の後に、④「ヘロインを一度摂取したならば、ヘロインを繰り返し摂取するだろう」(行為 B の展開)が接合することになる。要するに、ヒロシの懸念を SSA として展開しようとする、①から②へ向かう論証である「危険ドラッグを一度摂取したならば、危険ドラッグを繰り返し摂取するだろう。そして、危険ドラッグを繰り返し摂取したならば、ヘロインも繰り返し摂取するだろう」と、③から④へ向かう論証である「危険ドラッグを一度摂取したならば、ヘロインも一度摂取するだろう。そして、ヘロインを一度摂取したならば、ヘロインを繰り返し摂取するだろう」の、二通りを提示することができるのである。便宜的に前者を【①から②への接合型 SSA】、後者を【③から④への接合型 SSA】と呼ぶ。

このとき、両者の接合型 SSA によって説得力に違いが生じるかと問いを立て、前者の方が後者の方より合理的であると結論したい。以下では、その論拠を述べていく。

まず、④単体の SSA は、Walton が示した合理的な SSA の例そのものである。①単体の SSA についても、一般的には危険ドラッグも身体依存と精神依存が指摘されている⁸ことから、④と同様に①も合理的であると言えるだろう。①と④の SSA は、帰結の時点で、危険ドラッグないし違法ドラッグへ身体依存と精神依存の状態が成立していると言える。

①と接合する②単体の SSA は、論証の前提として、①における連続する要素の最終点に到達している。つまり、②における行為者は、危険ドラッグへの依存性があることが推測される。したがって、危険ドラッグよりも更に強い快感をもたらす違法薬物に手を出すことが促進されやすい状態にあり（ドライバーがある）、最終的な帰結に向かうと考えられる。反論として、危険ドラッグとヘロインの間には合法と違法の区別があるので、これらは連続していないと指摘できる（グレーエリアがない）。しかし、依存性というドライバーは、合法か違法かの区別を乗り越えてしまうと考える。なぜなら、〈ある危険ドラッグ、…、コカイン、ヘロイン〉の順序は、最終点に向かうにつれて精神依存と身体依存が強くなる順序であると想定でき、危険ドラッグに耐性がついた状態にある①を前提にすると、②における連続する要素内では、合法か違法かの区別は機能しないと想定できるからである。つまり、当初は連続している対象内にグレーエリアがなかったとしても、依存性をもつという行為者の状態の変化によって、グレーエリアが生じてしまう。以上より、【①から②への接合型 SSA】は、対話相手を説得するうえで合理的である。

問題は、③単体の SSA 「危険ドラッグを一度摂取したならば、ヘロインも一度摂取するだろう」である。〈危険ドラッグ、…、コカイン、ヘロイン〉と連続する要素をもつものの、それらの摂取は1回という前提がある。このとき、摂取する対象が別の対象に移行する連鎖を説明できていない。また、合法と違法の区別ができるので危険ドラッグからコカインまでは連続していない（グレーエリアがない）。つまり、③は合理的な SSA とはならない。したがって、④単体の SSA は合理的であるものの、接合される③単体の SSA が合理的でないために、【③から④への接合型 SSA】も対話相手を説得するうえで合理的ではない。

以上をまとめよう。ヒロシが提示する「危険ドラッグに手を出したら、薬物中毒になってしまう」という主張は、飛躍を指摘されたり、「滑り坂論法は誤謬だ」といった反論を受けたりするだろう。だがここで、ヒロシが提示す

る主張について、何の要素と何の要素が連続しているのかを明らかにし、Walton のスキーマに則って、連続する要素が一つ含まれる SSA に分解することができる。続けて、分解された SSA 単体同士を接合すると、合理的な論証となることがある。ヒロシの例でいえば、【①から②への接合型 SSA】である「危険ドラッグを一度摂取したならば、危険ドラッグを繰り返し摂取するだろう。そして、危険ドラッグを繰り返し摂取したならば、ヘロインも繰り返し摂取するだろう」という論証を示すことができれば、対話相手である娘を説得することができるかもしれない。

(4) 今後の展望

本稿の前半では、Walton による SSA の分析について、スキーマを中心にまとめた。すなわち、Walton は、SSA のスキーマに暗黙の前提が合致すること、連続する要素がグレーエリアを越えて最終的な帰結に向かわされるというコントロールの喪失があること、連続する要素の連鎖性とコントロールの喪失の証拠となるドライバーがあることを示すことができれば、SSA は合理的な論証になると主張している。

しかし、Walton のスキーマは、二つの連続する要素が含まれる（滑り坂が二つある）SSA を見過ごしている。二つの連続する要素が含まれる SSA は、その要素を明らかにし、最初にどの要素を展開するかによって、複数個の SSA に分解できる。このとき示される SSA は、一つの連続する要素が含まれる SSA となる。そのような SSA 同士を接合して再構成すると、接合された SSA によって、説得するうえでの合理性が左右されることがある。

ただし、接合された SSA は、合理的か否かの検討手続きが増えた分、その判定も難化するだろう。また、連続する要素として、例えば、誰が、いつ、どこで、何に、何をした等を挙げるができる。これらの連続する要素が複数含まれる SSA は、理論上想定可能である。連続する要素として何が位置づけられるのかという問いを含め、連続する要素が複数含まれる SSA の分析と、そのような複雑な SSA の検討が現実的に可能なのかについては、今後の課題としたい。

注

- (1) ただし、「妊娠初期の中絶は無実の人間を殺す一例である」ことを表明するために SSA が用いられることがあり (Williams, 1995, p.214)、この論証は最初の前提となる「妊娠初期の中絶を許容する」ことの道徳的含意に関する議論となっている。以上から、SSA が必ずしも「決定」に関する議論のみ

に用いられるとは言えない。

- (2) したがって、Walton (2015, 2017) によれば、堆積のパラドックスは SSA には当てはまらない。
- (3) 「危機的な帰結の前提」に含まれる「可能ならば避けるべき」という文言は、基本的スキーマの結論を冗長なものにするため、後に Walton 自身によって適切ではないとされた (Walton, 2017, p.1521)。
- (4) ただし、部分的に筆者が再構成している。具体的には、Walton の元の記述では、「危機的な帰結の前提」の一つ前に、「連続した前提」として「人間関係 (中略) はては人生全体を損なう」が独立に挿入されている。確かに、薬物依存が招く帰結にはグラデーションがあるかもしれない。しかし、SSA の提示者には危機的な帰結に含まれる道徳的な不正さや否定的な価値について対話相手と共有することが要求されており、このグラデーションのうち、どの要素をとっても危機的な帰結であると述べることができる。したがって、「人間関係 (中略) はては人生全体を損なう」を、「連続した前提」として独立にスキーマに挿入する必要はなく、「危機的な帰結の前提」として「人間関係 (中略) はては人生全体を損なう」というあらゆる事態が想定されることを示すことができればよいだろう。
- (5) Walton (2015) によれば、SSA は連続する要素の連続性に注目すると三種類に分類できる。すなわち、因果タイプ、(曖昧な用語に基づく) 言語タイプ、前例依拠タイプである。なお、これらすべてのタイプは、基本的な論証スキーマに合致する。
- (6) 危険ドラッグとは「規制されている薬物と化学構造がわずかに異なる薬物に目を付け、これを未規制の薬物として法律の網の目を巧妙にすり抜けて販売されている製品」(船田, 2016, p.66) のことである。
- (7) 本稿では、「対象物 A (ある危険ドラッグ、…、コカイン、ヘロイン)」を連続する対象として扱う記載をしている。この記載について、査読者からは、対象物が連続する要素となる想定をした場合、SSA が行為についての論証であるとする立場と齟齬をきたすこと、ならびに、行為以外を連続する要素として位置づけることができるのかという指摘を受けた。確かに、Walton によれば SSA は行為に関する論証であるため、A の要素は、対象物 (ある危険ドラッグ、…、コカイン、ヘロイン) ではなく、行為 (ある危険ドラッグの摂取…、コカインの摂取、ヘロインの摂取) として整理するのがより正確だろう。その場合、「最初の前提」は、「行為者は行為 A₀ を実行することを考慮している」(A (ある危険ドラッグの摂取、…、コカインの摂取、ヘロインの摂取) として) と「行為 B₀ を実行することを考慮している」(B (1

回目の摂取、2回目の摂取、…、N回目の摂取)として)という別々の前提に分解可能だろう。もしくは、「最初の前提」で言及される行為を一つの行為として捉えたうえで、二つの部分を含む行為として分析する方針もありうるかもしれない。詳細な検討は今後の課題とし、ここでは、分かりやすさのために、一つの SSA として捉えられるように、A を対象物として扱っている。

(8) 詳しくは船田 (2016) を参照されたい。

(5) 参考文献

den Hartogh, G. (2009). “The Slippery Slope Argument”. *A Companion to Bioethics*. Editors: Helga Kuhse, Peter Singer. Blackwell Publishing Ltd. pp. 321-332.

Walton, D. (2015). “The Basic Slippery Slope Argument”. *Informal Logic*: 35(3). pp. 273-311.

Walton, D. (2017). “The Slippery Slope Argument in the Ethical Debate on Genetic Engineering of Humans”. *Science and Engineering Ethics*: 23(6). pp.1507-1528.

Williams, B. (1995). “Which slopes are slippery?”. *Making Sense of Humanity: And Other Philosophical Papers 1982-1993*. Cambridge University Press. pp. 213-223.

van der Burg, W. (1991). “The slippery slope argument”. *Ethics*: 102(1). pp. 42-65.

児玉聡 (2017) 「第一章 倫理学の基礎」赤林朗編『入門・医療倫理 I [改訂版]』勁草書房 pp.22-23

船田正彦 (2016) 「危険ドラッグの薬物依存性と細胞毒性：基礎研究から探るその正体」『YAKUGAKU ZASSHI』136 巻 1 号 pp.65-72

(北海道大学)

Kantian Dog Argument とロボット倫理

—ロボットの道徳的地位の擁護に関する批判的考察—

Kantian Dog Argument and Robot Ethics

—A Critical Examination of the Defense of the Moral Status of Robots—

清水 颯

Abstract

Do robots have moral status? If they have, should we give them moral consideration for their own sake? This paper focuses on Coeckelbergh's paper, which argues that robots have an indirect moral standing from a relational approach. In particular, I critically evaluate Coeckelbergh's argument, which presents a Kantian dog argument as its starting point. I will then argue that robot ethics, which takes Kantian dog argument as its starting point, is incompatible with arguments in defense of the moral status of robots.

(1) 研究テーマ

ロボットは道徳的地位をもつだろうか。あるいは、ロボットは道徳的配慮に値する存在なのだろうか。人間とロボットの共生が想定される社会においては、我々はロボットをたんなる道具として扱うのではなく、何らかの意味で道徳的に配慮することが必要になるかもしれない。しかし一方で、ロボットは機械であり、人間（や動物）のような心的状態を持つと考えるのは難しい。そうであれば、ロボットは道徳的地位をもちえないようにも思われる。

さて、道徳的地位とは何だろうか。従来、道徳的地位の問題としては、「道徳的行為者性(moral agency)」と「道徳的被行為者性(moral patiency)」の二つが扱われてきた。道徳的行為者性の問題においては、その存在が自律的な意志や意識をもち責任を負うことができる主体であるかどうか問われ、道徳的被行為者性の問題においては、その存在が道徳的配慮に値するがゆえに尊重されるべき存在であるかどうか問われてきた⁽¹⁾。本稿が問題にするのは、ロボットが道徳的被行為者としての道徳的地位をもつかどうかである。

(2) 研究の背景・先行研究

ロボットが道徳的配慮に値する存在であるか、すなわち道徳的地位をもつかどうかという問題へのアプローチは様々である。ロボットや AI の倫理を

先導する倫理学者の一人である M. Coeckelbergh によれば、ロボットの道徳的地位をめぐる見解は、「特性アプローチ(properties approach)」と「関係論的アプローチ(relational approach)」という二つの立場に大別できる。特性アプローチは、例えば意識や痛みを感じる能力などの道徳的に重要な内在的な(intrinsic)特性がある存在は、道徳的配慮に値する存在であるから道徳的地位をもつ、という主張をする。Coeckelbergh は、D. Gunkel とともに、意識や感覚などの内在的な特性に基づいて道徳的地位の要件を定めてきた特性アプローチを批判する⁽²⁾。批判のポイントは、人間であれロボットであれ、その内的状態について確信することはできないという懐疑である。そこで、Coeckelbergh らは新たなアプローチとして関係論的アプローチを提唱し支持する(cf. Coeckelbergh 2021: 339)。

特性アプローチと対照的に、関係論的アプローチは、「機械の特性に基づいて道徳的地位を機械に付与するのではなく、その地位がロボットとの関係の中でどのように形成されるかに焦点を当てる」ものである(Coeckelbergh 2014: 71, Coeckelbergh 2021: 339)。Coeckelbergh は、特性アプローチに基づいた直接的な道徳的地位を問題にせず、関係論的アプローチから帰結する道徳的地位に焦点をあて、それを「間接的な」道徳的地位と呼ぶに至る(Coeckelbergh 2021: 340)。この「直接／間接」の区別は、カントの徳論で展開される「間接義務」の議論を参照し、それをより広義に道徳的地位についても適用することで提案されたものである。それゆえ、Coeckelbergh は、間接的な道徳的地位を擁護するための推論を助けるアイデアとして、「Kantian Dog Argument」という議論を提示し、それを出発点とする(ibid: 341)。

では、Kantian Dog Argument とは何か。Kantian と名がつくように、それはカントに触発された議論であるが、その出自は『道徳の形而上学』の「徳論の形而上学的基礎」(以下『徳論』)や『倫理学講義』の一部で展開された、「動物に関する間接義務」についての議論である⁽³⁾。カントは実際に、人間以外の存在をむやみに残虐に扱わないことや、それに感謝の気持ちをもつことは、間接的に義務であることを認める。間接的である理由は、それがあくまで人間に対する義務だからである。道徳感情を弱め、悪徳な性格が形成されることによって、人間が人間に対して負う義務に違反するため、それを間接的な根拠として、我々は動物などの非ヒト存在を配慮する必要がある。当然、カントはロボットと人間の共生などを想定していないため、ロボットを配慮する必要があるかどうか、答えを与えてくれるわけではない。しかし、Coeckelbergh は上記の Kantian Dog としてロボットを想定することで、同じ理由でロボットも配慮するべきであると提案する。これを Kantian Dog

Argument として定式化したのは Coeckelbergh だが、カントの間接義務を参考にして、ロボットへの振る舞いが人間の性格特性に影響を与えるという理由から、ロボットに危害を加えることを控えるよう主張する論者には、K. Darling や A. Gerdes、R. Sparrow などもある。

Kantian Dog Argument によれば、ロボットはそれ自体として道徳的地位をもつわけではないが、人間との関係を媒介する仕方でも間接的に道徳的地位を獲得できると考える。この主張を擁護するために、Coeckelbergh は四つの条件（一つが満たされればよい）を提示する。その筆頭に挙げられているのが、次のような徳倫理的な条件である。「もしロボットに悪いことをするような人間がいたら、他の人間からは性格が悪い、悪い人間（徳倫理でいうところの徳がない、人間に悪いことをする危険性がある）と見られるだろう」(ibid: 341)。この条件は、上記で述べた Kantian Dog Argument に直接基づいた議論であり、本稿ではこの第一の条件に焦点を当てる⁽⁴⁾。カント的な議論では、理性的存在者のみが直接的な義務の対象であり直接的な道徳的地位をもつが、『徳論』においてカントは、人間への影響を考慮して動物への行為を義務に関連付ける。Coeckelbergh も引用しているように、カントは「動物の苦痛に対する人間のうちなる共感が鈍くなり、そのことによって、他の人間との関係における道徳性に非常に役立つ自然的素質が弱められ、そのうちに根絶されてしまう」ことを理由に、動物への残虐な扱いを禁止する (MS VI: 443)。これは、間接義務として知られる議論の基本的な枠組みである。

この議論を徳倫理的に再構成すると、次のようになる。「もし私たちが犬に対して残虐に振る舞う習慣がついた場合、つまり、この悪徳を培い、性格が凶悪になった場合、私たちは人間に対して凶悪になる可能性が高くなる」(Coeckelbergh 2021: 343)。カントはあくまでも動物についてのみこの議論をしているが、Coeckelbergh は、人間の徳に注目する場合、ロボットにも同じ議論が適用できなければ首尾一貫していないと指摘する。これは人間の有徳な性格についての議論であるが、Coeckelbergh はこの議論を関係論的な仕方 (relational way) で定式化する⁽⁵⁾。それによれば、「何が有徳なのかは、私自身の評価だけでなく、他者の評価にも依存する」ため、「ロボットは徳に損害を与えるような行為をする可能性のある人間の徳に対する懸念に基づいて、道徳的地位を与えられうる」(ibid)。

このように Coeckelbergh は、関係論的アプローチと、人間に対する残虐さにつながるかもしれないという「道徳的リスク」への懸念に基づいて、ロボットを配慮すべき存在として位置付ける (ibid: 353)。しかし、Coeckelbergh も指摘するように、この結論は定言的なもの (categorical) では

なく仮言的なもの(hypothetical)である(ibid: 354)。ロボットを含めた非ヒト存在の直接的な道徳的地位が論争的である場合には、カントを源流とする間接的な議論が有効であると示したに過ぎない。しかし、カントの『徳論』に基づいた **Kantian Dog Argument** が、ロボットの間接的な道徳的地位を擁護する関係論的アプローチを発展させるための出発点になりうるのだろうか。例えば、動物倫理の議論においては、カントの動物に関する義務は動物の道徳的地位を間接的にでも擁護することを意図したものではなかったと評価されることが多い⁽⁶⁾。むしろ、非ヒト存在は人間の道徳的地位のための道具的で手段的な地位しかもたないはずである。また、カントの枠組みでは、人間への影響を考慮して動物などの非ヒト存在への道徳的な振る舞いを規定するが、そこで具体的な他者との関係性が問題になるのだろうか。次節では、人間の徳からロボットへの道徳的配慮の問題を考える徳倫理的なアイデアは、カントを出発点とする場合、どこに向かうべきなのかを提示したい。

(3) 筆者の主張

本節では、Coeckelbergh のように **Kantian Dog Argument** を出発点とする議論は、関係論的アプローチに基づいてロボットの間接的な道徳的地位を擁護するためには有効ではないことを主張する。Coeckelbergh は、**Kantian Dog Argument** を援用して、ロボットが間接的な道徳的地位を持つことを擁護するが、この議論はロボットが人間の徳を促進する手段として保護されるべきであるという点に焦点を当てており、ロボット自体が道徳的配慮の対象となるべき存在であるとは述べていない。それゆえ、**Kantian Dog Argument** に基づいて道徳的地位を擁護しようとするアプローチは、結局のところロボットの道徳的地位を根本的に擁護するものではないように思われる。**Kantian Dog Argument** によれば、ある存在がどのような理由であれ道徳的配慮の対象となる可能性がある場合、その存在に道徳的地位が認められると見なしているように思われるが、この理解は、その存在の固有の道徳的価値を認めるものではない。そこで、**Kantian Dog Argument** を出発点とするならば、ロボットにはそれ自身のために配慮されるべき道徳的地位はないとみなしたうえで、ロボットがそれとは別の理由で配慮される可能性があるのではないか、という問題設定にすべきだと提案する。

まずは、カントの見解をもう少し詳細にみていきたい。これまで見てきたように、Coeckelbergh も正しく指摘しているが、カントの主張の焦点は、人間の有徳さにとって重要である自然の素質、すなわち道徳感情を損なわせることを防ぐための間接的な手段として、動物への残虐な扱いを禁止すること

である。この議論が展開されるのは、「道徳的反省概念の両義性」、すなわち自己に対する義務を別の存在に対する義務と取り違えることについて説明される挿入章である(MS VI: 442)。つまり、カントはここで、動物に対する義務があるように見えるかもしれないが、それは自己に対する義務であることを示している。ここで想定される自己に対する義務とは、「完全性の義務」であり、有徳な性格を完成させる義務を意味する(cf. MS VI: 385-7)。動物を残酷な仕方で扱うと有徳な性格が損なわれる可能性があり、自己に対する義務に違反するため、そのような扱いは控えるべきなのである(7)。

ここで重要なのは、道徳的配慮の対象は誰なのか、という問題である。あくまで動物に関する義務は、自己に対する義務に関連する手段として導出されるため、義務がそれに対して向かうのは人間(とりわけ自己)である。それゆえカントの間接義務の議論では、動物に関する(regarding/in Ansehung)義務はあるが、動物に対する(toward/gegen)義務はない(MS VI: 443)。そしてカントは、「それはいつでもただ人間の自己自身に対する義務にすぎない」と結論づける(ibid)。ロボットについても同様に、人間のみが義務として配慮されるべき対象であり、その手段としてロボットへの残酷な扱いの禁止が導出される。それゆえ、**Kantian Dog Argument**を出発点とする場合、ロボットが道徳的配慮の対象でありうるのは、あくまで人間のためであることがわかる。しかも、ここでは自己の有徳さを完成させるための陶冶プロセスのみが焦点であるから、ロボットとの具体的な関係やインタラクション、それによる他者からの評価はとりわけ関係のないものである。それゆえ、**Kantian Dog Argument**と関係論的アプローチの結びつきは自明ではない。例えば、Coeckelberghが**Kantian Dog Argument**と呼ぶ議論と同じ個所を参照し、徳倫理的な観点からロボットへの残酷な扱いを控えることを主張するSparrowは、関係論的アプローチにはコミットしていないし、それゆえロボットの道徳的地位を擁護することを意図しない(cf. Sparrow 2017)。

Kantian Dog Argumentは、非ヒト存在が、人間の徳を中心とした道徳的配慮の一部に入りうることを正当化するためには有効な出発点だろう。しかし、それは関係論的な観点から道徳的地位を与える議論ではないため、ロボットの道徳的地位を擁護する議論の出発点としては有効ではない。少なくとも、**Kantian Dog Argument**に触発された徳倫理的な議論は、関係論的アプローチの枠内で間接的な道徳的地位を擁護する議論から独立させる必要がある(8)。

また、**Kantian Dog Argument**と呼ばれるカント的な間接義務の議論では、ロボットの道徳的地位を擁護し、それらが道徳的に配慮されるべきであるこ

とを主張することはできない。カント的な議論では、ロボットはそれ自身のために道徳的に配慮されるべきものではない、すなわち道徳的地位をもたないことを前提にしたうえで、人間の徳に依存する仕方での配慮を要求するからである。間接義務の議論では、それ自身が道徳的に配慮されるべき存在ではないとしても、別の理由から配慮される必要性があることを示唆するだけで、それ自身が道徳的に配慮されるべき道徳的地位をもつことを擁護できない。つまり、**Kantian Dog Argument** は、本来道徳的地位のない存在に対する道徳的な振る舞いを規定する議論であり、それらに道徳的地位を与えるというのはややトリッキーである。あくまでカント的な議論を出発点とするなら、その存在が道徳的配慮の対象としての道徳的地位をもつとは別の理由から、すなわち自分の徳を損なわせないという理由から配慮の対象範囲に含まれることは可能だが、それだけで道徳的地位を与えることは避けるべきであると考えられる。なぜなら、道徳的地位が問題にしていることとは、「その存在がそれ自身のために道徳的に配慮されるべきか否か」だからである (cf. Kamm 2007, Jaworska, Agnieszka and Julie Tannenbaum 2023)。

Coeckelbergh は関係論的アプローチと **Kantian Dog Argument**、徳理論的議論をつなげて、道徳的地位をあまりに薄く使用しているように思われる。例えば、この議論は「物を大切にしない人は徳がない」という議論と平行であると考えた場合、椅子にも間接的に道徳的地位が与えられるべきである、という主張がなされるはずである。これは、どこまで道徳的地位に値する存在とみなせるのか、という厄介な問題を抱え込む。しかし、本稿が明らかにしたように、カント的な着想に基づいた徳倫理的な議論はロボットに道徳的地位を与えることなく、別の理由からそれを道徳的配慮の対象に含めることができる。少なくとも、**Kantian Dog Argument** を出発点とする徳倫理的な観点からロボット倫理を検討するならば、その存在に道徳的地位がなくとも道徳的配慮が要求される可能性があるという方向で語る方が望ましいと言える⁽⁹⁾。本稿が論じてきたように、少なくとも **Kantian Dog Argument** を出発点とする徳倫理的な議論であれば、ロボットは道徳的地位をもつから配慮すべきだ、という実践的な含意は導くことはできない。しかしこの議論は、それ自身が配慮に値する存在でないとしても、その存在に対する倫理的振る舞いが規定される可能性を示唆している。

(4) 今後の展望

本稿は主に Coeckelbergh の **Kantian Dog Argument** を出発点とする関係論的アプローチに基づいて、徳倫理的な観点からロボットの道徳的地位を擁

護する試みを批判的に検討してきたが、今後は関係論的アプローチの展望と徳倫理的議論の展望をより広いスコープから見る必要があるだろう。本稿でも多少触れたが、ロボットに道徳的地位はないと考える立場から徳倫理的な議論を採用する Sparrow や、関係論的アプローチからロボットの道徳的地位を擁護しつつ徳倫理的な観点を採用しない Gunkel など、それぞれの立場の主張が独自の仕方で多様に展開しているように見える。ここで論点になるのは、それぞれの立場が考える「道徳的地位」や「道徳的配慮に値すること」の概念的な位置づけである。しかし、現状のロボット倫理の議論では、ターミノロジーが整理されていないこともあって、立場によってそれらの概念がずれており、それぞれの観点が複雑に交叉している状況にある。本稿の貢献の一つは、その立場上の概念的な混乱を整理することにあっただと言えるが、今後は、既に多くの蓄積がある動物倫理や医療倫理における道徳的地位に関する議論などを参考にしつつ、より包括的な観点からロボットと人間の共生についての倫理を考えていく必要があるだろう。

注釈

(1) ロボット倫理において行為者性(agency)と被行為者性(patency)がどのように扱われてきたかについては、Froridi 2013、猪ノ原 2022 を参照した。

(2) Gunkel は、レヴィナス的な他者論を根拠にして、ロボットが他者として現れるのであれば、それを根拠にロボットに権利を与えるべきであると主張する (cf. Gunkel 2018: 95-96)。

(3) カントの著作に関してはアカデミー版カント全集を用い、慣例に従って、巻数をローマ数字、頁数をアラビア数字で示す。MS は『道徳の形而上学』(Die Metaphysik der Sitten) の略称である。

(4) 紙面の都合上、残り三つの条件には触れることはできないが、それらを整理すると、「2. 関係的なケア感情に基づく条件」、「3. 共同行為に基づく条件」、「4. 予防的な議論に基づく条件」である(Coeckelbergh 2021: 341)。

(5) 2010年代の Coeckelbergh は、人間とロボットの具体的な関係性やインタラクションそれ自体を基礎としていたが、2020年代に入り、その関係を通じた人間への影響に焦点を移しているように思われる。これは、立場変更とまでは言えずとも、内在的な展開と評価することは可能だろう。

(6) カントが動物に対する義務はないと述べて動物の道徳的地位を否定したことはよく知られる見解であり、それは多くの動物倫理学者を不満にさせたと評価される(cf. Müller 2022: 59)。

(7) カントは他人に対する共感などの道德感情が損なわれないような扱いができれば、動物それ自体が積極的な配慮の対象であるとは考えない。それゆえ、『徳論』の同個所で、動物を使役することや、目的達成のために遂行される苦痛の少ない動物実験を行うことでさえ許容されている。

(8) カントの間接的な議論がロボット倫理でいかに位置づけられているのかを明確にすることを意図して書かれた Flattery 2023 では、関係論的アプローチが擁護しようとする間接的な道德的地位が議論の前提から外されている。これも、関係論的アプローチと Kantian Dog Argument の相性の悪さを物語っているように思われる。

(9) ある存在が道德的地位を持たずとも、それが私たちの道德的配慮の一部となりうることは、道德的地位に関するより一般的な議論の中でも成立している (cf. Kamm 2007: 299)。

(5) 参考文献

- Coeckelbergh, M. (2010). "Moral Appearances: Emotions, Robots, and Human Morality," *Ethics and Information Technology*, 12, pp. 235-241.
- Coeckelbergh, M. (2014). "The moral standing of machines: towards a relational and non-cartesian moral hermeneutics," *Philosophy & Technology*, 27(1), pp. 61-77.
- Coeckelbergh, M. (2021). "Should We Treat Teddy Bear 2.0 as a Kantian Dog? Four Arguments for the Indirect Moral Standing of Personal Social Robots, with Implications for Thinking About Animals and Humans," *Minds & Machines*, 31, pp. 337-360.
- Darling, K. (2016). "Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior Towards Robotic Objects," In R. Calo, A. M. Froomkin, & I. Kerr (Eds.), *Robot Law*, Edward Elgar, pp. 213-232.
- Flattery, T. (2023). "The Kant-inspired indirect argument for non-sentient robot rights," *AI Ethics*, Online publication date: 05-July-2023, <https://doi.org/10.1007/s43681-023-00304-6>.
- Froridi, L. (2013). *The Ethics of Information*, Oxford University Press.
- Gerdes, A. (2016). "The Issue of Moral Consideration in Robot Ethics," *Acm Sigcas Computers and Society*, 45 (3), pp. 274-279.
- Gunkel, D. (2018). "The Other Question: can and should Robots have Rights?," *Ethics and Information Technology*, 20, pp. 87-99.

- Jaworska, A., and Julie T. (2023). “The Grounds of Moral Status,” *The Stanford Encyclopedia of Philosophy* (Spring 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.). <https://plato.stanford.edu/archives/spr2023/entries/grounds-moral-status/>.
- Kamm, F.M. (2007). *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*, Oxford University Press, New York.
- Kant, I. (1908). *Kants Gesammelte Schriften*, Hrsg. von der Königlichen Preußischen Akademie der Wissenschaften und Nachfolgern, Bd.VI, *Die Metaphysik der Sitten* (MS), Berlin. [樽井正義、池尾恭一訳 (2002). 『カント全集 11 人倫の形而上学』、岩波書店]
- Müller, N. D. (2022). *Kantianism for Animals: A Radical Kantian Animal Ethic*, New York: Palgrave Macmillan.
- Sparrow, R. (2017). “Robots, rape, and representation,” *International Journal of Social Robotics*, 9(4), pp. 465–477.
- Sparrow, R. (2020). “Virtue and vice in our relationships with robots: Is there an asymmetry and how might it be explained?,” *International Journal of Social Robotics*, 13, 23–29.
- 猪ノ原次郎 (2022). 「ロボットの道徳的地位をめぐる近年の議論：道徳的なものの概念についての中間所見」、『応用倫理—理論と実践の架橋—』(北海道大学応用倫理・応用哲学研究教育センター発行オンラインジャーナル)、vol. 13、18-33 頁。

(北海道大学)

徳帰結主義についての予備的考察
Preliminary Thoughts on Virtue Consequentialism

杉山 弦

Abstract

In recent years, some consequentialists have proposed a sophisticated form of consequentialism called virtue consequentialism. Virtue consequentialism generally requires us to develop character traits that lead to good consequences. The purpose of this study is to clarify why the position of virtue consequentialism is necessary and to seek a valid formulation of virtue consequentialism. There are two types of reasons why virtue consequentialism is necessary: reasons extrinsic to consequentialism and reasons intrinsic to it, but most of the previous studies focused on the former. The author will discuss the latter as well, then point out the problems with the existing formulation of virtue consequentialism and attempt to construct a more refined formulation.

(1) 研究テーマ

近年、一部の帰結主義者は、徳帰結主義という洗練された帰結主義の形態を提案している。徳帰結主義は、一般に、善い帰結をもたらす性格特性を身につけることを我々に要求する。本研究の目的は、徳帰結主義という立場が必要とされる理由を明らかにするとともに、徳帰結主義の妥当な定式を求めるということに存する。徳帰結主義が必要とされる理由としては、帰結主義に外在的な理由と内在的な理由とがあるが、先行研究の多くはそのうち前者に焦点を当てていた。筆者は、後者についても議論するとともに、既存の徳帰結主義の定式の問題点を指摘し、さらに洗練された定式を構築することを試みる。

(2) 研究の背景・先行研究

帰結主義者が徳帰結主義という特殊な理論を採用する理由としては、帰結主義に外在的な理由と、内在的な理由とがある。研究の背景として、本節では、先行研究を踏まえつつ、外在的な理由について確認していく。徳帰結主義は、素朴な帰結主義とそれに対する徳倫理学者からの批判を背景として誕生したという側面がある。

外在的な理由については、次の順序で確認を行う。第一に、規範倫理学における三大理論の動向を概観する。第二に、徳倫理学者から素朴な帰結主義に対して寄せられる批判を概観する。第三に、帰結主義者の応答としてドライバーの徳帰結主義の定式を確認する。

(2-1) 規範倫理学における三大理論の動向

規範倫理学とは、行為の正不正や性格特性の善し悪しなどについて理論的に考察する分野である。規範倫理学には、大きく分けて、帰結主義、義務論、徳倫理学という三つの立場が存在する。

これら三つの立場のうち、帰結主義と徳倫理学の伝統的で素朴な主張は次のように整理できる。帰結主義によれば、「行為は、それが最善の結果をもたらす時、またその場合に限って、正しい」(Hursthouse 1999: 26) とされる¹。他方、徳倫理学によれば、「有徳な行為者とは、ある性格特性すなわち徳²をもち、かつ働かせる人のことである」(Hursthouse 1999: 29) とされる。

この整理からわかるように、素朴な帰結主義は行為の正しさに注目している。一方で、徳倫理学は、行為者がどのようなべきか³に注目している。ここにおいて、両者の関心は食い違っていると言える。

(2-2) 徳倫理学者から素朴な帰結主義に対して寄せられる批判

上記のような徳倫理学の主張を踏まえて、徳倫理学者は、大別して二つの批判を行っている。

一つ目には、徳倫理学者は、帰結主義は行為中心の行為者として批判してきた (Anscombe 1958; Hursthouse 1999)。確かに、伝統的な帰結主義はもっぱら行為の評価のみに関心を寄せてきたが、他方で徳倫理学は行為者および性格特性ないし徳に関心を寄せており、その点で徳倫理学は帰結主義や義務論とは根本的に異なる立場であるように見える。

二つ目には、徳倫理学者は、帰結主義は感情などからの行為者の疎外を引き起こすとして批判してきた (Railton 1984: 134)。そうした疎外批判というべきものを行っている論者として、ストッカーが挙げられる。

ストッカーは、感情からの疎外に着目している。感情からの疎外は、行為者のもつ、行為を勧奨する動機ないし感情と、行為の正当化理由との間に、齟齬があるときに生じる。さらに具体的にいえば、ストッカーの主張は次のようなものである。すなわち、あるもの A について内在的な価値を認める考えを持ち、その考えをある行為 X を勧奨する動機ないし感情に組み込んでいながらもかかわらず、もの A について内在的な価値を認めない考えをも持ち、それゆえに行為 X を勧奨する正当化理由をもたない、あるいはもの A について道具的な価値のみを認める考えをも持ち、その考えをその行為 X を勧奨す

る正当化理由としていることは、問題である、というものである (Stocker 1976)。たとえば、行為帰結主義者が友人を病院に見舞いに行くとき、その人が友情というものについて内在的な価値を認める考えを持ち、その考えを見舞いを勧奨する動機ないし感情に組み込んでいるにもかかわらず、友情というものについて世界の善を増進するという道具的な価値のみを認める考えをも持ち、その考えを見舞いを勧奨する正当化理由としているのは、問題であるとされる。ストックカーの言葉を借りれば、その行為帰結主義者は「道徳的統合失調症」³に陥っている (Stocker 1976: 454)。

ここまでで、徳倫理学者から素朴な帰結主義に対して寄せられる批判を概観した。たしかに、もし我々が意思決定の場面において素朴な帰結主義、ここではすなわちいわゆる行為帰結主義を参照するのであれば、こうした問題が引き起こされることは尤もである。しかしながら、より洗練された帰結主義の理論、すなわち本稿で主張される徳帰結主義は、こうした批判を回避できるかもしれない。

(2-3) 帰結主義者の応答

そこで、帰結主義者の応答としてドライバーの徳帰結主義の定式を確認する。

ドライバーの徳帰結主義

Vは、Vが現実世界においてより多くの善を体系的に生み出す性格特性であるとき、かつそのときに限り、(現実世界において)徳である。

(Driver 2001: 68; Bradley 2005: 283)

この定式におけるVとは、注2で述べた徳の特徴づけを参照すると、倫理的に望ましい性格特性を表していることがわかる。つまり、Vは、感情や好き嫌い、欲求、快樂などに関わる、安定的な、倫理的に望ましい傾向性のことを指す。このような徳帰結主義を採用した場合、行為者は徳を持ち、個々の行為の場面では徳を行使することが求められるのであって、個々の行為の場面で素朴な帰結主義、すなわちいわゆる行為帰結主義に照らして正しい行為を選択する必要はない。ドライバーは、このような徳帰結主義を採用することで、上述したような徳倫理学者から素朴な帰結主義に対して寄せられる批判を、少なくとも部分的には、回避できると主張する (Driver 2020: 463-465)。実際、この定式は行為中心な従来の帰結主義を乗り越えるものである。また、感情からの疎外が帰結主義的にも問題含みであるならば、行為者は、感情に沿って行為の正当化理由を構成し、疎外が起こらないような性格

特性を身につけることを要求されるのである。

(3) 筆者の主張

前節では、帰結主義者が徳帰結主義という特殊な理論を採用する理由としては、帰結主義に外在的な理由と、内在的な理由とがあると述べた。本節では、筆者の主張として、独自の用語法を使用しつつ、内在的な理由について議論していく。

内在的な理由については、次の順序で検討を行う。第一に、独自の用語法を提出する。第二に、全域的帰結主義という考え方を導入する。第三に、性格特性の被評価項としての重要性について論じる。第四に、全域的帰結主義を採用して徳帰結主義を論ずるメリットについて述べ、議論を整理する。

(3-1) 用語法

まず、帰結主義の諸定式に関する用語法を整理する。最初に、一般に、帰結主義は直接的帰結主義と間接的帰結主義という二つの定式に分類されることがあることを述べておく。直接的帰結主義とは、ある被評価項をその被評価項の帰結に照らして直接的に評価する帰結主義の定式である。いわゆる行為帰結主義は、直接的帰結主義の一定式である。他方で、間接的帰結主義とは、ある被評価項を、その被評価項とは別の被評価項の帰結に照らして間接的に評価する帰結主義の定式である。いわゆる規則帰結主義は、間接的帰結主義の一定式である。

上記の点を踏まえて、本稿では正確を期するために、呼称を次のように整理したい。

まず、直接的帰結主義の呼称については、直接的帰結主義の諸定式が直接的に評価する被評価項の名前を、「についての」という節を後置した形で、「帰結主義」の前に挿入する。例えば、**行為についての帰結主義**は、いわゆる行為功利主義のことを指す。そして、**性格特性についての帰結主義**⁴は、性格特性を性格特性の帰結に照らして直接的に評価する帰結主義の定式である。

また、間接的帰結主義については、それが間接的に評価する被評価項を第一の被評価項と呼び、第一の被評価項を間接的に評価するために参照される、第一の被評価項とは別の被評価項を、第二の被評価項と呼ぶことにする。ここにおいて、間接的帰結主義の呼称については、間接的帰結主義の諸定式の第一の被評価項の名前を、「についての」という節を後置した形で、「帰結主義」の前に挿入し、その第一の被評価項の名前の前に、第二の被評価項の名前を、「を経由した」という節を後置した形で挿入する。例えば、**規則を経由した行為についての帰結主義**は、いわゆる規則帰結主義のことであり、行為

を規則の帰結に照らして間接的に評価する帰結主義の定式である。そして、**性格特性を経由した行為についての帰結主義**は、行為を性格特性の帰結に照らして間接的に評価する帰結主義の定式であり、これは第一の被評価項を行為とする点において、いわゆる規則帰結主義と平行に理解することができる。

なお先に注で述べたとおり、本稿で「徳帰結主義」という用語を使った際は、それは、「**性格特性についての帰結主義**」の諸定式、および「**性格特性を経由した行為についての帰結主義**」の諸定式をひとまとめにして指している。

(3-2) 全域的帰結主義

全域的帰結主義とは、あらゆる事物や事象を被評価項として、それらの帰結に照らしてそれら进行评估するという発想である (Pettit & Smith 2000: 121; Driver: 2012: 146; Parfit 1984: 24-26)。被評価項は、行為や規則だけでなく、性格特性でも、目の色でも、気候でも、なんでもよい。一方、局所的帰結主義は、全域的帰結主義とは対照的に、限られた数の被評価項のみを妥当なものとして認める立場である。例えば、行為についての帰結主義のみを認め、他の被評価項を考慮しない立場は、局所的帰結主義の一種と言える。

他方で、全域的帰結主義は、さまざまな被評価項を認め、それについての具体的な定式を包含する。**行為についての帰結主義**や、**規則を経由した行為についての帰結主義**、**性格特性についての帰結主義**、あるいは、**性格特性を経由した行為についての帰結主義**は、それら具体的な諸定式の例である。

(3-3) 性格特性の被評価項としての重要性

ただし、**全域的帰結主義**を採りながらも、特定の存在物が被評価項として重要であるという事実は存在するであろう。もし、その存在物の有無や状態の如何が世界の善の増減に大きく関わっていたり、その存在物の有無や状態の如何に我々が介入可能であったりするならば、その存在物は被評価項として重要である。

そして、だからこそ実際、少なくとも行為や規則は、被評価項として重要だろう。それゆえに**行為についての帰結主義**や、**規則を経由した行為についての帰結主義**は実践的に有用なのである。

問題は、性格特性が、被評価項として重要であるかどうかである。性格特性を被評価項として扱うことが可能であることは、性格特性が被評価項として重要であることを含意しない。他方で、もし性格特性が重要であれば、それゆえに徳帰結主義は実践的に有用であるということになる。

そして実際、性格特性は、その有無や状態の如何が世界の善の増減に大きく関わっているという意味において、重要である。筆者は、その根拠として、

我々は意思決定の場面において性格特性ないし徳を行使しているし、せざるを得ないという事実を指摘する。さもなくば、問題が生じる。我々は、一瞬ごとの微細な行為について、それが**行為についての帰結主義**に照らして正しいものかどうか、自ら功利計算して確かめたり、**規則を経由した行為についての帰結主義**に照らして正しいものかどうか、規則を参照して確かめたりして、その行為をなすということは、不可能である、というのも、我々には時間・情報・認知能力という資源が不足しているからである（Hare 1981: 35-36; Crisp 1992: 147-148）。そうではなくて、我々は習慣化された性格特性ないし徳を行使することによって、一瞬ごとの微細な行為を行っているのである。我々はむしろ、**性格特性についての帰結主義**が指定する望ましい性格特性すなわち徳を身につけ、それを行使することが必要である。そのような徳の行使によってなされる行為は、**性格特性を経由した行為についての帰結主義**に適うものとなる。

(3-4) 全域的帰結主義を採用して徳帰結主義を論ずるメリット

全域的帰結主義を採用して徳帰結主義を論ずることのメリットは、行為の正しさの基準を**行為についての帰結主義**に求める一方で、それとは異なる評価軸として性格特性の有徳さや行為の有徳さも導入できる点にある。性格特性の有徳さは**性格特性についての帰結主義**によって、行為の有徳さは**性格特性を経由した行為についての帰結主義**によって、それぞれ基準が与えられる。

このように、全域的帰結主義は、**行為についての帰結主義**という帰結主義の原点にあたる発想を維持しつつ、新たな評価軸も取り入れることができるのである。この包括的なアプローチにより、より多角的な視点から帰結主義を構築し直すことが可能になる。

(4) 今後の展望

筆者が今後、第一に行うべきことは、既存の徳帰結主義の定式の問題点を指摘し、さらに洗練された定式の構築を試みることである。上述のドライバーの**徳帰結主義**は、現代の徳帰結主義に関する論争の嚆矢となったものの、定式自体が問題を含んでいたり、あるいは重要な論点を棚上げにしていたりすると指摘されている（Hartman & Bronson 2021）。既存の論点に関しては、客観主義と主観主義の対立（Bradley 2018: 403）や、最大化問題（Bradley 2017: 82）、閾値の設定問題（Bradley 2005）、個人主義と普遍主義の対立（Bradley 2005）、現実主義と反事実主義の対立（Bradley 2005）などを解決しなければならない。また、新規の論点として、性格特性の集合における個々の性格特性の相互作用の問題や、集団における性格特性の多様性が最善

の帰結をもたらす可能性があるという問題についても、議論が必要である。

また、第二に行うべきことは、第一の作業によって得られた洗練された徳帰結主義が、実際に徳倫理学者からの批判に耐えうることを示すことである。それによって、帰結主義そのものが、徳倫理学に伍する、あるいは優る理論であることを示すことができるかもしれない。

注

[1] ここで注意すべきことは、ここにおける帰結主義の典型的な主張は、一般に行為帰結主義と呼ばれる帰結主義の一形態に過ぎないということである。しかし、20世紀の学界における帰結主義に関する認識としては、行為帰結主義こそが帰結主義の本流であったことは事実である。

[2] 「徳」という概念は、徳倫理学の立場にコミットせずに、次のように特徴づけられる。すなわち、徳とは、倫理的に望ましい性格特性のことである。つまり、徳を持つ人は、倫理的に望ましい感情を持ち、倫理的に望ましい好き嫌いのあり方を持ち、倫理的に望ましいことをなすことを欲求し、そして、倫理的に望ましいことをなすことで快樂を得る。そして、徳は、自己の安定的な状態である。また、徳は、習慣づけに基づいて獲得されるものであり、また、ある種の知識によって成立するものである。なおこの特徴づけは、アリストテレスおよび現代の彼の支持者の記述を参考としている（アリストテレス 2014: 1103a–1105a; Urmson 1988: 26–27; Hursthouse 1999: 92–93）。

[3] 「道徳的統合失調症」という用語は統合失調症患者へのスティグマを助長する点で、望ましいものではない。とはいえ、ストッカーにとっては実に多くの人々が「道徳的統合失調症」に陥っているのである。見舞いの例は常識的にありふれたものであるが、それこそをストッカーは問題視している。

[4] 「**性格特性についての帰結主義**」は、「**徳についての帰結主義**」と書き表すべきかもしれない。というのも、諸々の先行研究においてはすでに、そのような**性格特性についての帰結主義**にあたる、いわゆる性格特性帰結主義とでも呼ぶべきもののことを、徳帰結主義（Virtue Consequentialism）と呼ぶことが一般化しているからである。しかし、正確を期するのであれば、**性格特性についての帰結主義**は、そのように書き表した方が、**行為についての帰結主義**や**規則についての帰結主義**と同様に、名前が実態を表すことができる。すなわち、**性格特性についての帰結主義**は、性格特性を性格特性の帰結に照らして直接的に評価する帰結主義の定式である、と言える。そのため、本稿ではやむをえず、「**徳についての帰結主義**」と一般に書き表されるべきものを、「**性格特性についての帰結主義**」と書き表すことにする。なお、本稿

で「徳帰結主義」という用語を使った際は、それは、「**性格特性についての帰結主義**」の諸定式、および「**性格特性を経由した行為についての帰結主義**」の諸定式をひとまとめにして指しているものとして理解していただきたい。

(5) 参考文献

- Anscombe, G. E. M., 1958, “Modern Moral Philosophy,” *Philosophy* 33(124):1–19.
- Bradley, B., 2005, “Virtue Consequentialism,” *Utilitas* 17(3):282–298.
- Bradley, B., 2017, “Character and Consequences,” Fileva, I. eds., *Questions of Character*. Oxford University Press, 78–88.
- Bradley, B., 2018, “Contemporary Consequentialism Theories of Virtue,” Snow, N. eds., *The Oxford Handbook of Virtue*, Oxford University Press, 398–413.
- Crisp, R., 1992, “Utilitarianism and the Life of Virtue,” *Philosophical Quarterly*, 42(167):139–160.
- Driver, J., 2001, *Uneasy Virtue*, Cambridge University Press.
- Driver, J., 2012, *Consequentialism*, Routledge.
- Driver, J., 2020, “Consequentialism, Virtue, and Character,” Douglas W. P. eds., *The Oxford Handbook of Consequentialism*, Oxford University Press, 463–473.
- Hare, R. M., 1981, *Moral Thinking: Its Levels, Method, and Point*, Oxford University Press.
- Hartman, R. and Bronson, J., 2021, “Consequentialism and Virtue.” (<https://philpapers.org/archive/HARCAV-4.pdf>、2023年5月4日取得)
- Hursthouse, R., 1999, *On Virtue Ethics*, Oxford University Press.
- Parfit, D., 1984, *Reasons and Persons*, Oxford University Press.
- Pettit, P. & Smith, M., 2000, “Global Consequentialism,” Brad H., Elinor M. & Dale M. eds., *Morality, Rules and Consequences—A Critical Reader*, Edinburgh University Press, 121–133.
- Railton, P., 1984, “Alienation, Consequentialism, and the Demands of Morality,” *Philosophy and Public Affairs* 13(2):134–171.
- Stocker, M., 1976 “The Schizophrenia of Modern Ethical Theories,” *Journal of Philosophy* 73(14):453–466.
- Urmson, J., 1988, *Aristotle’s Ethics*, Blackwell.

アリストテレス、神崎繁訳、2014、『アリストテレス全集 15：ニコマコス倫理学』岩波書店。

(東京大学)

新進研究者 Research Notes

「科学への信頼」とは何か？

What is 'trust in science?'

KANG KIWON

Abstract

This study explores the concept of 'trust in science,' addressing the challenges arising from public skepticism towards scientific findings. It argues that trust in science is not merely about accepting the outcomes of scientific consensus, but rather about trusting the process of reaching that consensus. The research highlights the role of emotions in shaping public trust and emphasizes the importance of transparent and inclusive scientific processes. By examining cases like vaccine hesitancy and controversies in evolutionary theory, the study suggests that involving third parties in the scientific process can enhance trust. It concludes that fostering trust in science requires acknowledging scientific uncertainties and promoting a culture of open and critical discourse within the scientific community.

(1) 研究テーマ

科学への不信による諸問題をより厳密に議論する後続研究の土台として「科学への信頼」の概念を具体的に整理し、科学への信頼を増進させるための考察を行う。

(2) 研究の背景・先行研究

近年、科学への信頼 (Trust in Science) は、パンデミックといった世界規模の災難などの脈絡で言及されることが増えている。例えば、気候変動における温室効果ガスの問題を指摘する科学的主張を信じない人々に対し、科学への信頼を欠いている (Mistrust、以後「信頼を欠く」もしくは「信頼の欠如」は「不信 (Distrust)」とは区分される比較的中立の意味として用いる) と評するのは代表的な事例といえるだろう。COVID-19 期間中に見かけられたワクチン接種の躊躇 (vaccine hesitancy) の問題を考えてみよう。ワクチン躊躇はワクチン反対や忌避とは区分される概念であり、接種をするか (させるか) どうか「悩んでいる」状態を含意する (町田・井上 2023)。ワクチ

ンがリスクよりベネフィットが上回っている点から接種を推奨するという科学的立場からして、ワクチン躊躇は科学への「不信」ではないものの、科学への信頼が「欠けている」とはいえるだろう。

科学への信頼を欠く要因は概ね、信念 (beliefs)、価値 (values)、感情 (emotions) の3つが挙げられる (Furman 2023)。これらは相互的に作用しているが、その中でも核となりうるのは「感情」である。というのも、自身なりの信念や価値観を持っていない人は珍しくないが、個人差はあるとしても感情そのものを持ち合わせていない人は極めて特殊な状況を除けばいないからだ。上述したワクチン接種の躊躇 (vaccine hesitancy) においてもリスクよりベネフィットが上回るという点を理論的には理解しているが、それでも極端な状況(リスク)を恐れてしまうという点から感情の問題が大きい。そのため、ワクチン躊躇の解消にはこれを取っ払う必要があるとされている (Zhou et al., 2023)。

もちろん、科学はその性質上、リスクを完全に排除することは不可能に等しいが、社会において最も信頼に値する基準を提示できるのは依然として科学である。不確実な状況下で常に選択を迫られている我々は、客観的な基準を見出す方法を科学的方法以外に持たないからだ。しかし、感情の問題を乗り越えるにはこういった理屈だけでは有効ではないことが、昨今の風潮を鑑みる限り明らかに思える。科学への不信や不安が漂っている現状を乗り越えるには、市民へ盲目的な信頼を要求するのではなく、信頼される科学を如何に営むかの工夫ではないだろうか。

上述した通り COVID-19の流行や気候変動といった世界規模の危機が相次ぐ最中「科学への信頼」をめぐる議論は増えてきているが、研究者間の明確な合意はなされておらず、いささか迷走しているかのようにも思える。例えば Price & Lindsey (2016) は、経済的に豊かな社会ほど寧ろ科学への支持 (support for science) が低いと結論付けているが、論文中に飢餓や医薬品の不足といった極度の貧困を経験した場合は科学への信頼 (confidence in science) が低いという矛盾したような結果を提示している。また、科学リテラシーも科学への信頼の要因として挙げられることがあるが、両者は関係がないか (横山 2019)、関連が大きくない (山本 2019) という論者も少なくはない。このような不明瞭な状況に問題意識を覚え、ここにて検討することが本研究の目的である。

まず、本研究で想定している「科学」とは何かを先行研究の検討を通じて整理しておく。本稿は科学に対する市民の信頼をテーマとしている点から、自然科学を指し示す狭い意味での科学ではなく、広い意味として捉えている。

とりわけ、科学と技術の一本化が進んでいる（文部科学省 1999）もしくは実質的な意味で結びついている（神里 2016）とされる昨今の認識を背景とした上で、米国の科学史家・科学哲学者の Oreskes の「科学は専門家による議論を経たうえでの合意（*science is simply the consensus of relevant experts on a matter after due consideration.*）」という定義を採用する。Oreskes は著書 *Why Trust Science?* にて、科学への不信には合意を導き出す主体である科学者への不信が潜んでおり、科学者自らの行いが重要という脈絡で上記の定義を提示している。

本稿では採用しないが、一連の議論を理解するためには Kitcher (2011) の観点も参考に値する。Kitcher は一部の倫理的に間違っただ科学者がいるとしても、市民と科学者の価値観の調和に基づいた方法論を選択していくことで科学への信頼を増進（boost）させられると提言している。しかし Gürol & Faik (2021) は、Kitcher が市民と科学者の調和に基づいた方法論を常に選択できるかのように述べている点に加えて市民がどのように選択を行えばよいか明確な説明がないことを指摘している。*Why Trust Science?* の共著者である米国の科学哲学者 Lange の見解も Kitcher の立場に近い。Lange は Oreskes へのコメント（第 4 章）として、科学への不信を退けるには、科学の合理的根拠（*rational basis of science*）をどう伝えるかを工夫する方が重要であると強調している。Oreskes は再反論（第 7 章）として科学における合意は認識論的条件ではなく社会的条件であり、我々は真実が何かを見定める方法がないため合意を用いると反論している。

いささか議論が平行線を走っているように思われるが、本研究では社会的な脈絡の上で「科学」を捉えている点から、基本的に Oreskes の観点を擁護する形で展開していく。

(3) 筆者の主張

背景を整理しよう。ワクチンの事例のように、「感情」の問題は人々が科学技術に不信感を抱く根本的な要因として作用しうる。また、科学は科学者の合意の産物という点を併せて鑑みると科学への信頼を確保するには、科学者自らの合意が公正に行われていることを信じてもらえるよう、説得を繰り返さないといけないように思われる。

説得という点から、古典的ではあるがアリストテレスの説得の 3 要素を参考にできる。アリストテレスはロゴス（論理）、パトス（情熱）、エトスの 3 つを説得において必要な要素として挙げている。論理的な構成、情熱的な言い方は間違いなく重要であるが、中でもアリストテレスが最も重要に扱った

のはエトスである。特性や人柄などと翻訳されるエトスは、話者が普段信頼に値する行いをしていたか、もしくは専門知識を備えているか否かを聴者の主観に基づいて、エトスがどうか低いかを判断することを意味する。科学への信頼のために必要な要素として科学者の態度を強調した Oreskes の観点は、まさしくエトスの重要性を指し示しているともいえるだろう。

しかし科学者は時々、市民の目からしてエトスが低いと思われるような態度をとることがある。というのも「科学者たちは、利害を離れた科学的発見が人類によって利用されたことに対して直接の責任はない」と、科学の利用において「無責任態勢」を貫いた過去があるからだ（内井 1998）。このような歴史が ELSI (Ethical, Legal and Social Issues) の発足に繋がり、科学者の倫理観も過去より強化されてきているものの、いまだに科学が客観的で価値中立であると主張し、批判を受け入れないような風潮は完全に消え去っていない。

特定研究不正に該当する 2005 年の ES 細胞事件(韓国)や 2014 年の STAP 細胞事件は、それぞれ事前に外部からの懸念の声があったにもかかわらず「足の引っ張り合い」としてみられ黙殺されている。事件発覚後、内部の科学者たちは、権威に屈服していたか (Hong 2006) 盲目的に指導を行っていた (佐藤ら 2019) ことが明らかになる。このように、批判を許容しない一部の風潮は、外部からもエトスが低い態度としてみられ不信を誘発しているのではないだろうか。

科学が「科学者」の合意に基づいた知識蓄積の営みという観点に立てば、「認定された者」として科学者は該当分野において権威を持つことになる。権威を持つ点から必然的に科学者同士にも序列が発生し、批判を許容しないような風潮になりかねない。とはいえ、権威の問題を取り払う試みがないわけではない。その一つは研究における市民参画であり、もう一つは科学者を観察の対象とする人類学的アプローチが挙げられる。両者とも部外者が間接的ながら研究に加わることで、科学者の合意が公正に行われているかを観察する手段になりえる。市民参画の研究が市民からの信頼の増進に繋がる (植田ら 2022) ことは兼ねてから知られているが、参画できる分野に偏りが見られる。反面、人類学者による長期間のラボでのフィールドワークは、分野の制限が少ない点から多くの可能性を示唆する。

さて、ここで本題である「科学への信頼」とは何かを論じたい。Oreskes の観点を採用し「科学」を「科学者の合意の結果」と定義づけるのなら、信頼の対象は「合意」そのものではなく「合意のプロセス」だと考えられる。合意のプロセスにおいて、不安定要素となる科学者のエトスが問題として挙げ

られている点から、そもそも合意が正しくなされないというプロセス上の不信が潜んでいるからだ。言い換えれば「科学を信じる」ということは、科学とされた知識を盲目的に受け入れることではなく、知識の生産過程が正当に行われていることを信頼することになるといえるだろう。

ここで問題となるのは、合意のプロセスが正当に働いていることを科学者自らが証明することは難しい点である。裁判だけでなく、個人同士の言い争いにおいても本人の発言だけで他人を説得させることは難しい。しかし、第三者の証言がある場合、当事者の発言は信頼されやすくなる。そして科学において「第三者」を招き入れる制度は、上述した市民参画の研究や人類学的アプローチが挙げられるであろう。部外者として、科学者の合意の過程で利益相反や特定の結論を強要するハラスメントなどがなかったどうかを観察できるからだ。

査読が科学者同士で審査するという点で問題があるという Lange の指摘のように、この試みも既存の問題を完全に取り払えるわけではない。果たして市民や人類学者といった第三者が正当に観察を行っているかを確認できるだろうか？ 極端な例を挙げると、不正を行う科学者が観察を行っている人類学者に賄賂を提供する可能性もある。しかし、限界をもっているにもかかわらず、査読という制度は科学の発展に貢献している実績は実在している。同じく本研究の考察が完璧なものではないとしても、科学への信頼を高めることには繋がると期待している。目前の課題を乗り越えるには、完全な解だけを追い求めるより、地道にでも改善をもたらす可能性のある選択肢を考慮すべきではないだろうか。

先行研究として引用した Oreskes の主張との比較としては、科学者の合意が科学であるという主張から更に一歩進み、「科学への信頼」は「科学者の合意形成のプロセスへの信頼」であるとした点に本研究のオリジナリティがある。この成果により、科学者の合意形成のプロセスを観察・証言する第三者の存在が必要であるという結論を見出している。これは科学への信頼のために研究者の多様性が重要と力説した Oreskes の提言と相反するわけではないが、重点の置き方には違いがある。

(4) 今後の展望

本研究は先行研究に基づき「科学への信頼」とは「科学者の合意形成のプロセスへの信頼」のことであることを示し、信頼を増進させるために必要な要件を明確に提示した。ワクチン接種の忌避といった科学技術への不信による社会的な課題に取り込む際に、本研究の成果が貢献できると考えられる。

本研究の適用による展望を、似通っている事例として予測してみることはできるだろう。韓国では医療人の資格を持っていない医療機器メーカーの職員が手術に同席するなどの問題で、手術室内の監視カメラ設置義務化が進み2023年9月25日付けで実施されている。これに対し医者団体は「録画が気になり手術への集中が困難になる」と強く反対していた。だが、今まで医者団体の自浄作用が期待できなかった経緯から、韓国市民はむしろ監視カメラ導入に賛同を示している。実際、監視カメラが導入されることで医者は医療ミスの疑いから身の潔白を証明でき、患者側も安心して手術が受けられるという側面から、寧ろ医療に対する信頼の増進に繋がっている。なぜなら、2021年の時点で試験的に監視カメラの導入された病院において、カメラの導入により安心できると答えた患者は80%を超えているからだ。また、カメラの設置が不満と答えた医者は17%にとどまり、患者の反応のよさもあって満足していると答えた医者が40%を超えている(Ko 2021)。この事例を本研究に対応させると、監視カメラは市民や人類学者のような「第三者」に、監視カメラの記録は第三者の「証言」に該当するだろう。また、最初は気まずい雰囲気であっても慣れてくると寧ろ当事者(医者)の満足度も高くなったという点から、本研究の提言は市民と科学者両方の満足に繋がると期待できる。

日本における科学技術への不信を要因とする問題としては、子宮頸がんを予防するHPVワクチン接種の躊躇問題が挙げられる。2013年当時、副反応による被害の訴えがあったことと、それにより厚生労働省がワクチンの積極的な推奨を一時中断したことが原因で、当時70%程度の接種率は激減し、2022年度も7.1%に留まっている(3回接種)。約10年をかけて厚生労働省とWHOが接種を推奨するメッセージの発信を続けているが、本稿で議論したように当時者(科学者)のみを中心とした発信では、市民の信頼を得るのに限界がある。そこで、本研究の成果を適用すれば、HPVワクチンが信頼のある科学技術の一つとして認められるためには何が必要なのかを検討できるのではないだろうか。

本研究の成果を適用する際に想定される問題は、セキュリティ上の懸念による反発である。近頃の科学研究は、先に成果を発表したグループが栄誉を独占することは稀ではない。そのため、第一線の研究室では部外者を同席させることに躊躇すると思われる。研究上の競争が全てオープンに行われることは理想論であり、また科学技術の軍事利用といった国家間の問題も絡んでくると、本研究の適用にも限界があるかもしれない。しかし、本研究による成果はCOVID-19といった世界規模の危機に対処する研究や、国家間の利害関係のない研究には貢献できる。また「科学への信頼」とは何かを定義し、

信頼のためにどういう手段が考えられるかを提示できているため、少なくとも理論的な寄与はなしていると考えられる。

(5) 参考文献

- Furman, K. (2023) "Listen Research Article Beliefs, values and emotions: An interactive approach to distrust in science", *Philosophical Psychology*, DOI: 10.1080/09515089.2023.2266454
- Gürol, I. & Faik K.(2021) "Distributive Epistemic Justice in Science", *British Journal for the Philosophy of Science*, <https://doi.org/10.1086/715351>
- Hong, S. (2006) "Special Issue : Analysing Hwang Scandal to the root| What is STS position toward the Hwang Scandal?", *Historical criticism*, vol 74, pp.55 – 81
- Kitcher, P. (2011) *Science in democratic society*. Prometheus Books.
- Ko.J(2021) Operating room CCTV 'positive for restoring trust in healthcare'...patient satisfaction up 80 percent, *Youth Doctor*, <https://www.docdocdoc.co.kr/news/articleView.html?idxno=2013486>
- Oreskes, N. (2019) *Why Trust Science?* Princeton University Press.
- Price, M.A. & Lindsey P.(2016)"Scientific progress, risk, and development: Explaining attitudes toward science cross-nationally", *International Sociology*, Vol. 31, Issue 1, <https://doi.org/10.1177/0268580915614593>
- Zhou,Y., Li, R. & Shen, L.(2023) "Psychological profiles of COVID vaccine-hesitant individuals and implications for vaccine message design strategies", *Vaccine X*, doi: 10.1016/j.jvaex.2023.100279
- 植田舞、松尾薫、加我宏之、武田重昭 (2022) 「市民参画型公園づくりにおける参加者に対するエンパワーメントの方法に関する研究」『ランドスケープ研究』85 巻 5 号 pp. 457-462
- 内井惣七 (1998) 「科学者の責任を考えるために」『大学の物理教育』98 巻 3 号 pp. 4-8
- 神里達博 (2016) 「情報技術における ELSI の可能性歴史的背景を中心に」『情報管理』vol. 58, no. 12, pp. 875-886.
- 田町征己・井上茂 (2023) 「Vaccine hesitancy (ワクチン躊躇) の現状、関連要因、評価、対策」『日本公衆衛生雑誌』70 巻 8 号 pp.474-482

佐藤洋一、相澤純、田島克巳、伊藤智範（2019）「科学の不正と倫理教育」
『岩手医誌』71 卷 3 号 pp. 75-87

文部科学省（1999）「「科学」と「技術」、「科学技術」について」『21世紀の社会と科学技術を考える懇談会（第3回）配布資料3-6』

山本耕平（2019）「疑似科学への態度の規定要因に関する諸仮説の検証」『年報科学・技術・社会』28 卷 pp.25-46

横山輝雄（2019）「ブタペスト宣言と科学リテラシーの転換」『科学・技術研究』8 卷 1 号 pp. 3-4

（大阪大学）

『哲学探究』において規則論と言語論はどのような関係にあるのか
What is the relationship between the rule theory and the theory of language
in *Philosophical Investigations*?

白木啓吾

Abstract

The subject of this paper is Wittgenstein's main work, *Philosophical Investigations*. The purpose of this paper is to clarify the relationship between the rule theory and the theory of language in *Philosophical Investigations*. The conclusions of this paper are as follows. His rule theory concerns the foundations of mathematics, not language in general. But he thinks that following the rules is analogous to meaning something by words. In this sense, the rule theory and the theory of language are closely connected in his philosophy.

(1) 研究テーマ

本研究の研究テーマは、Ludwig Wittgenstein（以下 L. W. ）の主著、『哲学探究』（以下 PU）における、規則論と言語論との関係である。本稿では 185 節から 242 節までを「規則論」と定義する。本稿の目的は、PU において「規則論」と言語論がどのような関係にあるのかを明らかにすることである。

(2) 研究の背景・先行研究

様々な主題を扱う PU のなかでも、「規則論」はその中心的な箇所だと見なされている。そのことをいち早く指摘したのは、Kripke（1984）である。

Kripke はまず、「68+57」という数式で規則のパラドックスを再構成する。前提として、私たちはこれまで 57 より大きな数で加法を行ったことがないとしよう。私たちは普通、この数式に対して 125 という答えを与えるだろう。しかし、私たちの前に一人の懐疑論者が現れる。その懐疑論者によれば、「68+57」の正しい答えは 5 であるというのだ。懐疑論者の言い分は以下である。すなわち、私たちが「+」でこれまで意味していたのは、その実プラスではなくクワスであった。今、プラスを「+」という記号で表して、クワスを「+」という記号で表すとすると、クワスは次のような関数だと説明できる。

もし $x, y < 57$ ならば $x+'y=x+y$

そうでなければ $x+y=5$

仮定から、私たちは実際に 57 よりも大きな数で加法を行ったことがないのだから、私たちが実はプラスではなくクワスを行ってきたのだという主張を否定するような事実は存在しないのである。

一度、懐疑論者の言い分を受け入れてしまえば、「68+57」の答えは 125 や 5 だけでなく、どのような数であってもいいことになる。というのも、確かに私たちは「68+57」の答えが 125 でしかありえないと主張するためのいかなる事実も持っていない。しかし他方で、懐疑論者もまた、その答えが 5 でしかあり得ないと主張するためのいかなる事実も持っていないからである。

結局、何らかの解釈を通すことで、どのような答えであっても「68+57」の答えとして受け入れられてしまうように思える。これが、L. W. の発明した規則のパラドックス (PU 201) である。

次に、Kripke はこの議論を言語一般に拡張させる。Kripke は例としてテーベアという奇妙な概念を挙げる。テーベアとはエッフェル塔の外ではテーブルを意味するが、エッフェル塔の中では椅子を意味する概念である。議論の流れはクワスと同じである。私たちがこれまで「テーブル」で意味していたのは、その実テーブルではなくテーベアであったのだ。こうして、規則のパラドックスは言語全般へと拡張される。

結局、このパラドックスが意味するところは、私たちが言葉で何かを意味するという事実は存在しないのだから、言葉が意味を持つということはありません、というものになる。Kripke は L. W. がある新しい懐疑論を、すなわち意味についての懐疑論を発明したのだと述べている。

さて、こうした Kripke による解釈の是非については、本研究の関心ではない。本研究にとって重要なのは、数式という規則について生じたパラドックスを言語全般へと拡張させる議論である。Kripke はここで明らかに、「規則論」は言語ゲーム全般の規則を主題としているという解釈を採っている。そして、Kripke による L. W. 解釈は多種多様な批判に曝されてきたにもかかわらず、この部分を批判した者は、管見の限り、存在しない。Kripke は規則に関する議論を、言語一般に関する議論へと拡張した。この拡張には、言語が規則に支配されているという前提が隠れている。言語を用いることは規則に従うことであるという前提に立つからこそ、規則に従うことが規則のパラドックスによって危機に曝されたとき、言語一般も危機に曝されるのである。

言語が規則に支配されているという前提は、Kripke だけでなく、極めて多くの論者によって採用されている。例として、Hacker と Glock を挙げる。

「規則に従う」ことは言語制度の基礎である。ある言語を学ぶということは、規則に支配されたその言語に含まれる表現の使用に関する技術をものにすることである。知られていない、あるいは理解されていない、そうした規則に従うことは不可能である。したがって、表現の意味を決定し構成する規則が、将来の発見を待っていて、知られていない、そうしたことはありえない。(Hacker 2001, p. 59)

ある表現の言語的意味は、一般的な規則に依存している。(中略)これらの規則は、表現を正しく使用するための基準を提供する。それによって、これらの表現を使って何を言うことが有意味であるのかが決まるのである。したがって、L. W. 的な意味の規範性とは、言語的意味とあるタイプの言語的規則との関連に関する特定の主張である。L. W. によれば、ある表現の意味は、その正しい使用に関する規則によって決定される。(Glock 2009, p. 159)

このように、言語が規則に支配されているという前提は、様々に解釈を異にする多くの論者にとって、数少ないコンセンサスの一つとなっている。

「規則論」と言語論との関係に関する標準的な解釈は、この前提によって答えられるだろう。すなわち、もしこの前提が正しいのであれば、規則について問うことはそのまま言語について問うこととなる。言語が規則に支配されている以上、規則を主題とする「規則論」は言語論の一部なのである。

対して、言語が規則に支配されているという前提を否定する論者も少数ながら存在する¹。そして、この前提を採用しない論者は、上記の標準的解釈を採ることができない。そうした論者の一人である大谷(2020)は、「規則の問題」の主題は——その名称に反して——規則ではない」(p. 126)という驚くべき主張をする。大谷の解釈では、「規則の問題において問題となっているのは、言葉の使用の正しさである」(p. 109)というのである。したがって、「規則論」とはその呼ばれ方に反して、言語論なのである。大谷の解釈から導かれる、「規則論」と言語論の関係はこのようなものになるだろう。

大谷は大きく分けて三つの論点から、標準的解釈を批判する。第一に、大谷による「規則論」の解釈では、L. W. はここで、「何かが表現の使用の正しさを「決定する」とするような哲学的像から我々を解放しようとしている」

(大谷 2020 p. 123) のである。したがって、言語が規則に支配されているという前提に立つ、「規則論」の主題は規則であるという標準的解釈は、「ウィトゲンシュタインの議論の趣旨に合わない」(同上)。第二に、言語が規則に支配されているという前提を L. W. に帰すテキスト的な根拠がない。第三に、185 節以下の議論において「規則」という語が考察の中心にあるわけではなく、この議論は「規則」という語に触れることなく再構成できる。

第一に、言葉の使用の正しさを決定しているものが何もないように思われることが規則の問題であると、大谷は述べる。そして、この問題に対する L. W. の応答は、「決定」という語を明確化することによって、この問題に欠陥があることを示すというものである。L. W. は「移行が代数式によって決定されている」という表現の実際の使われ方を思い起こす (PU 189)。例えば、「計算という行為に関する人々の教育による決定」や、「計算体系における変項の値の決定」のモデルが提示される。そして、このどちらのモデルも、「「+2」という表現の使用の正しさは何によって決定されているのか」という哲学的な問いを成立させるものとはならない」(大谷 2020, p.119) と主張する。こうした疑似問題が生じた原因は、「決定」という語の像が不明確なまま使用されてしまったことにある。「決定」という語にはいくつかの使用モデル、すなわち像があるⁱⁱ。そうした「様々な像が互いに交差し合う結果として」(PU 191)、私たちはこのような「過度な表現を使うように誘惑され」(PU 191) てしまったのである。詳述すると、人間のエラーや限界を無視する「計算体系における変項の値の決定」のモデルとそれらを考慮に入れる「計算という行為に関する人々の教育による決定」のモデルとが交差し合った結果、「数列は計算間違いや人間の計算能力の限界を超えて、神秘的な仕方で決定されている」(大谷 2020, p.116) ように思われてしまったのであり、そこから規則の問題が生じたのである。したがって、何であれ、表現の使用の正しさを決定するものの存在を、L. W. が許すはずがないのだ。

第二の論点について、大谷は標準的解釈が論拠として引用する 199 節を分析して、この節がその実、標準的解釈を何ら支持していないことを主張する。

一度だけ、ただ一人の人間が、ある規則に従ったということはありません。一度だけ、ただ一つの報告が行われたということ、一度だけ、ただ一つの命令が与えられたということ、あるいは、一度だけ、ただ一つの命令が理解されたということ、などはありえない。一規則に従うこと、報告を行うこと、命令を与えること、チェスの対局

をすること、これらは**慣習**（習わし、制度）である。／ある文を理解することは、ある言語を理解することである。ある言語を理解することは、ある技術をものにすることである。（PU 199, 強調原文）

大谷はこの節において、規則に従うことが伝達や、命令、チェスの対局をすることと並列に位置付けられていることを強調する。つまり、規則に従うことは他の「言語使用と区別され、その上で両者の類似性が問題とされているのである」（大谷 2020, p. 124）。したがって、199 節は標準的解釈を支持するどころか、それに反対しているようにさえ読めるのだ。

第三の論点について、大谷は実際に「規則論」を規則について言及することなしに再構成して見せているⁱⁱⁱ。また、185 節以下ですぐに「規則」という語は現れず、そこではむしろ「+2」という「命令」が主題であると主張する。「規則」という語が登場するのは、197 節からなのである^{iv}。したがって、「規則論」の主題が規則であるとするのは誤りであるのだ。

(3) 筆者の主張

本研究では標準的解釈への大谷による批判に対しては同意する。しかし、大谷が代わりに提案する解釈に対しては、同意できない。大谷による議論は、言語が規則に支配されているという前提への反論にはなっていない。「規則論」が規則を主題としているという主張自体を否定できているわけではない。そして、大谷による解釈は以下の二つの点で不十分である。一つ目は、規則論の大部分が実際に規則に関する考察にあてられていることを説明できない点である。二つ目は、数学の基礎を主題の一つとしているという PU の序文と齟齬が生じる点である。以上を踏まえ、本研究では、言語が規則に支配されているという前提を否定しつつも、「規則論」が規則を主題としているという新しい解釈を提案する。

はじめに、大谷に反して、「規則論」はやはり規則を主題としている。その理由は二つある。

第一に、「規則論」の大部分が実際に規則に関する考察にあてられている。「規則論」は PU185 節から 242 節までを指している。そして、確かにこのうち、はじめの 185 節から 196 節までは「規則」という語が表れていない。しかし、その後の 197 節から 242 節までは「規則」という語が頻出しており、その内容からしても、規則に関する文法的考察がほとんどである。したがって、事実として「規則論」の大部分が規則に関する考察にあてられており、

このことは「規則論」が規則を主題としていることを示している。

第二に、「規則論」は数学の基礎論であると考えられる。PUの序文において、L. W. はPUで扱う主題として、数学の基礎を挙げている。そして、この「規則論」が、数学の基礎に該当すると考えられる^v。また、PUの内部では、この「規則論」の他に数学の基礎に該当すると言えそうな議論はない。ところで、大谷の解釈するように、「規則論」の主題が規則ではなく、言葉の使用の正しさであるのであれば、「規則論」は数学の基礎を主題としているとは言えなくなる。そうすると、PU内部に数学の基礎を扱っている箇所がなくなってしまう。したがって、PUの序文と本文との間に整合性を保つためには、「規則論」が数学の基礎を主題としていなければならない、つまり「規則論」は規則を主題としていなければならないのである。

さて、次にPUにおける「規則論」と言語論との関係について述べる。規則と言語は以下の二つの仕方で関係している。

第一に、一部の言語使用は実際に規則に従って行われている。もっとも身近な例を考えれば、外国語の習得が挙げられる。外国語を習得する際は実際に、教科書に載っている文法規則や辞書に載っている語同士の対応規則を参考にして、言語使用が行われる。こうした意味で、一部の言語使用は実際に、規則を基にして行われる。また、あらゆる言語使用は規則によって評価・訂正・説明・教示される可能性を有している。こうした意味であれば、規則は言語使用の基礎にあると言えることができる。ただしこの表現には注意が必要である。ここで述べているのは、あらゆる言語使用が規則によって評価等をされる**可能性**を有しているということであり、**実際に**あらゆる言語使用が規則によって評価等をされているわけではない。また、より重要なこととして、すべての言語使用を一挙に規則によって評価等することは不可能であるという事情がある。規則によってある言語使用を評価等することはそれ自体が一つの言語ゲーム、あるいは言語使用であり、したがって、このゲームを営むためには（少なくともこのゲームを営んでいる間は）規則による評価等から逃れている言語使用が必要なのである。

第二に、規則に従うことは言葉で何かを意味することとアナロジーの関係にある^{vi}。これは、言語活動をゲームになぞらえるPUの核となる方法論の延長線上にあるものだと言える。PUの中で明示的にこのアナロジーを語っているのは、238節である。

私にとって規則がそのすべての帰結を前もって生じさせていると思

えるためには、それら規則からの帰結が私にとって**自明**でなければならない。それは、この色を「青」と呼ぶことが自明であるのと同じように自明であるのでなければならない。（PU 238, 強調原文）

L. W. は規則に従うことを分析するなかで、規則はそのすべての帰結を事前に生み出している、と私たちが思う傾向があることを指摘した。例えば、「+2」という規則は、私たちが実際にその規則に従って数を記述する以前からすでに、その規則からどの数を書くことが導かれるのかが決まっている、そのように考える傾向が私たちにはある。そして、私たちがそのように考えるということは、ある規則からある行為をすることが、私たちにとって自明である必要があると、L. W. は指摘する。

ここでは、規則からその帰結が生じることの自明さと、この色を「青」と呼ぶことの自明さとが比べられている。この色を「青」と呼ぶことは言語使用の一つであるから、ここで L. W. は規則に従うことと言葉で何かを意味することとを比較しているのである。このことから、規則に従うことと言葉で何かを意味することとがアナロジーの関係にあると、L. W. が考えていたことが分かる。そのため、規則に関する考察は、類推を通して、言語に関する考察の一部となる。また、この節は言語が規則に支配されているという前提に対する反論にもなっている。というのも、もしこの前提が正しいのであれば、言語使用は規則に従うことの一部なのだから、この色を「青」と呼ぶことも規則に従うことの一部となる。しかしそうになると、ここで行われているアナロジーが成り立たなくなってしまう。規則に従うことと言葉で何かを意味することとが並列の関係にあるからこそ、アナロジーは成立するのである。

(4) 今後の展望

今後の展望として、以下の二つが挙げられる。第一に、言語が規則に支配されているという前提をより詳しく検討する。本研究ではこの前提を否定したが、PU の一部分しか扱っていない点で、この論証は不十分である、この前提は L. W. の後期哲学全体に関わるものであるため、PU 全体や、その他の著作にまで分析対象を広げて考察する必要がある。第二に、規則に従うことと言葉で何かを意味することのアナロジーは、今後一層注目されるべきである。既存のほとんどの解釈は、言語が規則に支配されているという前提に立っているため、このアナロジーは見逃され続けてきた。今一度、このアナロジーに着目して後期ウィトゲンシュタイン哲学を捉え直す必要がある。

注

ⁱ このような論者としては、野矢（2022）、大谷（2020, 2014）、Glüer & Wikforss（2009）、Hanfling（2002）などが挙げられる。

ⁱⁱ 大谷は像とモデルを使い分けている。しかし、ここでは簡単のために、これらを同一視して議論を進める。（大谷 2020, pp. 57-65）参照。

ⁱⁱⁱ （大谷 2020, pp. 109-121）

^{iv} この指摘は、（Hanfling 2002, p. 62, note. 20）でもなされている。

^v 鬼界彰夫（2016）参照。

^{vi} Glüer & Wikforss（2009）もこのアナロジーを指摘している。

(5)参考文献

Glock, H-J, 2009, "Meaning, rules, and conventions." In E. Zamuner and D. K. Levy. (ed.), *Wittgenstein's Enduring Arguments*, London and New York, Routledge, 156-178.

Glüer, K. & Wikforss, Å, 2009, "Es braucht die Regel nicht: Wittgenstein on rules and meaning." In D. Whiting (ed.), *The Later Wittgenstein on Language*, Hampshire, Palgrave Macmillan, 148-166.

Hacker, P. M. S, 2001, "Wittgenstein and the autonomy of humanistic understanding." In R. Allen and M. Turvey (ed.), *Wittgenstein, Theory and the Arts*, London, Routledge, Reprinted in P. M. S. Hacker, 2001, *Wittgenstein: Connections and Controversies*, Oxford, Oxford University Press, 34-73.

Hanfling, O, 2002, "Does language need rules?" In O. Hanfling, *Wittgenstein and Human Forms of Life*, London and New York, Routledge, 51-65.

Kripke, S, 1982, *Wittgenstein on Rules and Private Language*, Harvard U. P.

Wittgenstein, L, 2009, *Philosophical Investigations revised 4th ed*, Hacker, P.M.S. & Schulte, J. ed, Wiley-Blackwell.

大谷弘, 2014, 「言語は規則に支配されているのか」, 『哲学』第 65 号, 日本哲学会, pp. 135-150.

大谷弘, 2020, 『ウィトゲンシュタイン 明確化の哲学』, 青土社

鬼界彰夫, 2016, 『『哲学探究』の「序文」を真剣に読む』, 『哲学論叢』43 卷, 京都大学哲学論叢刊行会, pp. 14-26.

野矢茂樹，2022，『ウィトゲンシュタイン『哲学探究』という戦い』，岩波書店

（中央大学）

『論理哲学論考』における独我論の整合的解釈に向けて
Toward a consistent interpretation of solipsism in the *Tractatus Logico-Philosophicus*

近藤雅熙

Abstract

The subject of this paper is the issue surrounding the interpretation of solipsism in Ludwig Wittgenstein's *Tractatus Logico-Philosophicus*. This paper argues that the two seemingly conflicting interpretations of solipsism in the *Tractatus*, namely the “Russellian reading” and the “Schopenhauerian reading”, can both hold true at the same time.

And, based on above discussions, this paper presents the possibility that the discussion of solipsism in the *Tractatus* can be analyzed as a relationship between two sets of “solipsism-subject”. According to this interpretation, these two “solipsism-subject” relationships can be analyzed as “Russellian solipsism-the thinking and representing subject” and “Schopenhauer's solipsism- the metaphysical subject”.

(1) 研究テーマ

本稿が研究の対象とするのは、ルートヴィッヒ・ウィトゲンシュタイン『論理哲学論考』（以下『論考』と略記）における独我論の解釈に関する問題である。

(2) 研究の背景・先行研究

『論考』における独我論の議論が何であったのかという問題は、『論考』解釈上の重大な争点の一つであると言える。この問いは、基本的には独我論が議論されている命題 5.6 番台全体をどのように解釈するべきであるかという問題として現れる。『論考』解釈として既に古典の位置を占める Pears(1987)は、「独我論は、ウィトゲンシュタインの様々な思考の流れにおける結節点に位置する」(p.153)と述べ、また、本稿で中心的に取り上げる Tejedor(2015)も、『論考』の独我論に関する議論がこの書物全体の屈折点を表しているこ

とは疑いようがない」(p.73)と主張している。ゆえに、『論考』における独我論の議論を分析することは、結果的には『論考』全体の解釈を確定させる試みに直接的な形で寄与することが期待されるのである。

さて、Tejedor(2015)は、『論考』の独我論に関する先行研究を、ラッセル的な読み、ショーペンハウアー的な読み、そしてマッハ的な読みという三つの型に分類している。本稿はこの分析にもとづいて、ラッセル的な読みとショーペンハウアー的な読みは、『論考』の 5.6 番台に関する解釈として共存可能であると論じることを目標としている。ゆえに、まずは Tejedor に従って『論考』の独我論に関する先行研究を整理することから始めたい。

すでに述べた通り、Tejedor は『論考』の独我論に関する先行研究を三つの読みに分類している。ここでは、マッハ的読みについては本稿の議論との関係を持たないため省略し、「ラッセル的読み」と「ショーペンハウアー的読み」という二つの解釈について紹介する。

第一に、ラッセル的な読みとは『論考』における独我論に関する議論とラッセルの認識論との関係を強調する解釈上の立場のことである。この読みは更に大別されて、ウィトゲンシュタインが『論考』でラッセルの認識論に由来するような独我論の概念を批判したという観点からの読みと、反対に『論考』はラッセルの考えるような「独我論」をある意味では貫徹しようとしたという観点からの読みが存在する(Tejedor(2015), p.48)。このうち、前者は Diamond(2000)に、後者は Pears(1987)に代表される立場であるとされる(cf. *Ibid*, n.8)。だが、そもそも「ラッセルが考える独我論」とはいったい何であるか。

注意しておくべきは、ラッセル自身は独我論的な立場を全くとっておらず、むしろ独我論を反駁することの方に強い動機を持っていたという点である。ゆえに、ここで問題とされているのは、ラッセルが「独我論」ということで具体的にいかなる独我論を意味しようとしていたか、という点である。ここで重要となるのが、ラッセルの認識論における基礎的な概念の一つである「面識 acquaintance」である。面識とは、認識主体とセンスデータの間に成立する関係であるとされるが、このセンスデータは各人に私的なものである。ゆえに、この面識の私秘性に由来する独我論的世界像の可能性は、たとえば『哲学の諸問題 *The Problems of Philosophy*』におけるラッセルにとって重大な問題として立ち現れているように思われる。そして、この面識の私秘性を前提した時に導かれるタイプの独我論を、Tejedor は「ラッセル的独我論」として次のように定式化する(pp.47-48)。

- (a) 面識による知識のみが存在する
- (b) 他我 *other selves* は面識において与えられない
- (c) 私は他の対象と同様に自我 *my self* に対して面識する

結論：

私は面識において私に与えられる、自我と他の対象についての知識のみを有することができる。

ここで対象とはセンスデータのことであると理解してよいと思われるが、ラッセルが自我に関して常に上で定式化されているような見解を採用していたと考えるのは誤導的である。だが、この構図はラッセルがある時期に採用していたタイプの議論を指すというよりは、上の諸前提のどの部分をラッセルが否定し、それによって独我論的な帰結を乗り越えようとしたかという点を明確にする上で役に立つものであると考えられる。

たとえば、Tejedorによると、ラッセルは「面識による知識と記述による知識」において前提 (a) を否定し、『知識の理論』では自我についての知識を「記述による知識」によって知られるものと措定することで (c) を否定している。この (c) の見解は主に『哲学の諸問題』の時期のラッセルに帰される立場であると思われるが、この時期のラッセルはセンスデータからなる各人の私的世界をいかにして公共的世界と接続するかという点に頭を悩ませていたように思える (cf. ラッセル(2005), pp.26-27, McGinn(2006), p.260)。

既に述べたように、Tejedorによれば、『論考』がこのようなラッセル的独我論を攻撃したという解釈 (Diamond) と、反対に『論考』はこのような独我論を貫徹しようとしたという解釈 (Pears) が存在する、とされる。本稿はこのうち、Diamond に代表される『論考』の独我論解釈を採用するが、Diamond(2000)を直接参照する形でこの解釈を擁護することは行わない。本稿の議論は「『論考』5.6番台の議論はラッセルの認識論的諸前提に対する批判である」というテーゼ (この点はDiamondや、彼女の見解を部分的に引き継ぎながら、ラッセルの認識論に対する批判として『論考』の5.6番台を読もうとするMcGinnに共通している) のみを引き継いだうえで、仮にこのテーゼが正しいとするならば、それはショーペンハウアー的な読みとも両立し得ることを論じるものである。

McGinn(2006)が注目するのは、『哲学の諸問題』においてラッセルが行っている、いわば自我の面識可能性に関する議論である。ラッセルは、「太陽を見ること」を面識するという場面を想定し、次のように述べる。

「太陽を見ること my seeing the sun」を面識しているとき、私は明らかに、互いに関係しつつも異なる一つのものを面識している。一つには、私に太陽を提示するセンスデータがあり、もう一つ、そのセンスデータを見ているものがある。(中略)したがって、自分が太陽を見ていることを私が面識しているとき、面識されているのは「自我-面識-センスデータ」という事実全体である(ラッセル(2005)、62頁)。

その帰結として、ラッセルはためらいながらも「私たちはある意味で、個々の経験と対比されたものとしての自我を面識していると思われる」(同書、63頁)と説いている。McGinnによれば、ウィトゲンシュタインは『論考』6.6番台における命題 5.6331 (視野の比喩) および 5.634 (経験のア・プリアリ性の否定) において、このラッセルの「自我-面識-センスデータ」という認識論的な枠組みを批判している(cf. McGinn(2006), p.264)。この議論の詳細には立ち入らないが、ここで重要なことは McGinn が『論考』5.6番台の議論をラッセル批判の文脈において捉えているという点である(この点は(3)で後述される議論と関わる)。

次に、Hacker に代表されるショーペンハウアー的な読みについても、先述の Tejedor にしたがって簡潔に整理する(p.48)。Tejedor によれば、ショーペンハウアー的な独我論とは「表象—あるいは表象において与えられるものとしての世界—は、超越論的主体の世界である」というテーゼで表現される形式の独我論である。ここで超越論的主体は(ラッセルの自我とは異なり)面識可能な対象や、世界の中に存在するような主体ではない¹。それは、いわば「世界の超越論的条件」としての主体である。この解釈によれば、表象としての世界は、その超越論的条件としての主体を必要とするのである。『論考』期のウィトゲンシュタインがショーペンハウアーの『意志と表象としての世界』に影響を受けていたことは周知の事実であるが、このことを踏まえ、Hacker に代表されるショーペンハウアー的な読みをとる解釈者は、ウィトゲンシュタインが『論考』においてこの形式の独我論を主張していると解釈している。

以上が Tejedor による説明であるが、Hacker(1986)の解釈は、文献学的な傍証という点から見て十分な根拠を持つように思われる。また、Hacker はショーペンハウアー的な読みを踏まえた上で、『論考』の独我論を「超越論的独我論 Transcendental Solipsism」と規定する(Hacker(1986), p.104)。

(3) 筆者の主張

先述した通り、本稿はこれらの独我論解釈が両立可能であることを示すことを目標としている。二つの解釈の大きな違いは、一方が『論考』はラッセル的な独我論を否定している」と主張するのに対し、他方が『論考』はショーペンハウアー的な独我論を肯定している」と主張するという点にある。だが、この二つの議論は両立させることが可能であるように思われる。

仮に、これらの解釈が『論考』は独我論を否定している」という立場と『論考』は独我論を肯定している」という立場であるとすれば、この両立は端的に不可能である。しかしながら、先に挙げた二つの解釈はそのような対立構造を持っていない。すなわち、これらの解釈は（実際に両立するかどうかは別として）形式的には両立させることが可能であると述べることができる。ゆえに、本稿はこの形式的な両立可能性を元手にして両解釈の具体的な共存可能性を探りたい。

第一に、ラッセル的な読み（Diamond, McGinn など）によれば、『論考』は独我論の議論においてラッセル的な独我論を否定している。そして、この解釈は文献学的な傍証は不足しているように思えるものの、『論考』の議論が（ウィトゲンシュタイン自身は意図していないものであったとしても）結果的にラッセルの認識論に対する反駁となっている、ということは十分に成立可能であるように思われる。

たとえば、McGinn は命題 5.6331、5.634 をラッセルの認識論に対する批判として読もうとしているが、仮にこの見立てが正しいとするならば、『論考』の命題番号の振り方を踏まえることで、命題 5.63 番台が全体としてラッセル批判に向けられていると解釈することは可能であると考えられる。命題 5.63 は「私は私の世界である」²と述べているが、これは『論考』における独我論テーゼ「世界が**私**の世界である」(5.62)とは異なる。ラッセル的な独我論は、世界と私を分離したものとして捉えた上で、「私が世界を面識する」という認識論的構造を作り出すものである。したがって、少なくとも『論考』の独我論テーゼと形式的には等しい（だが、独我論の概念はそこにおける「我」の性格に応じていくつものヴァリエーションを生じる）。それに対して、「私は私の世界である」というテーゼは、それ自体が「世界を認識する」という性格をもつ私を排除するものとなる。ゆえに、5.63 番台はこのように捉えられた（世界を認識するものとしての）主体という概念の否定により始められていると考えられる。

命題 5.63 に続く 5.631 は、「思考し表象する主体は存在しない」と主張す

る。そして、ここで否定される主体をラッセルが面識において考えていたような主体と同一視することは、続く 5.631 の記述からも可能であるように思える。そこでは、「私が見出した世界」という本を私が書くという想定の下、「ある重要な意味において主体が存在しない」(5.631) と述べられている。この 5.631 の記述と 5.6331 の視野の比喩とは、5.63 番台がラッセル批判の文脈において捉えられるならば、相互に関連したものとして読むことができる。

ゆえに、5.63 番台において、「ある意味において主体が存在しない」ということが主張されており、この主体が（偶然か意図されたものかは明確でないにしても）ラッセル的な主体を指すと考えることは許されるように思われる。また、この視野の比喩がラッセル的な認識論を批判しているという主張は、例えば McGinn によって既に行われている。McGinn は P. Sullivan による議論を踏まえつつ、この視野の図（限界づけられた視野の左端に視点としての眼が描かれている）は二つの点で誤っていると述べる。すなわち、第一に、主体を実体としたこと（つまり、視野の中に書き込んだこと）、そして、第二に視野を限界づけられたものとして描いたことである (McGinn(2006), p.265. Cf. p.264)。

しかし、『論考』5.63 番台は「独我論」という表現を用いてはいない。ゆえに、この解釈が十全なものとなるためには更に別の根拠を必要とするように思われるのである。そこで、その候補として本稿が提案したいのは命題 5.64 の記述である。5.64 は「ここにおいて、独我論を徹底すると純粋な実在論と一致することが見てとられる」と述べている。この命題は『論考』の命題番号の順序によれば 5.6 の注釈であり、またここで現われる「独我論」の表現は命題 5.62 において「(その) **言わんとする**ところはまったく正しい」と主張される「独我論」と同一の独我論を指すと考えられてきたように思われる。

だが、これについては幾つかの疑問を立てることが可能である。第一に、仮に 5.64 の「独我論」が 5.62 のそれと同一の概念であるならば、それを（「**言わんとする**ところはまったく正しい」とされるにもかかわらず）「徹底する」ということの意味が不明となる。第二に、「ここにおいて **Hier**」という指示語が単に命題 5.6 を指すと考えるのは、この解釈が『論考』の命題番号のシステムにもとづいているために反駁困難であるとはいえ、文脈上はやや突飛であるように思われる。そして第三に、命題 5.64 は「純粋な実在論」に関して述べており、その点では 5.63 番台と同様に認識論的なことがらを問題にしていると解釈することには一定の整合性があるように思われる。

以上から、本稿は命題 5.64 における「独我論」がラッセル的な読みにおけ

る「否定される独我論」を指示しているという可能性を提示したい。この解釈の正しさについては、なお議論の余地がある。しかしながら、このように仮定することは次に述べるショーペンハウアー的な読みと、ここでのラッセル的な読みとの共存可能性を高めることにつながるのである。

さて、ショーペンハウアー的な読みは『論考』において、ウィトゲンシュタインはあるタイプの独我論を正しいものとして主張している」と解釈する。そして、この見立ては命題 5.62「すなわち、独我論の**言わんとする**ところはまったく正しい」という記述から正当化されるように思われる。だが、このショーペンハウアー的な読みはラッセル的な読みと両立可能である。

第一に、命題 5.63 番台は「思考し表象する主体」を否定しているとともに、「形而上学的な主体」(5.633)という新たな主体概念を導入している。そして、この「形而上学的主体」は「思考し表象する主体」とは異なり「世界の限界」であるとされており、存在しない主体概念として措定されているわけではない（思考し表象する主体は「存在しない」(5.631)のに対し、形而上学的主体は「世界に属さない」(5.632)）。また、5.632の「主体」が「形而上学的主体」を指していることは、5.633や5.641などの命題から明らかであると思われる(cf. Hacker(1986), p.86)。つまり、5.63番台はHackerが「世界の超越論的条件」とするような主体である「形而上学的主体」についても論じていると考えられる。換言すれば、5.63番台はラッセル的な読みとショーペンハウアー的な読みの双方に（一方の読みを否定することなく）関係していると思われる。

第二に、5.64における「(徹底された場合には)純粋な実在論と一致する」とされる独我論を、ショーペンハウアー的な読みにおける独我論の理解と重ねることは可能である。「[その] **言わんとする**ところはまったく正しい」と言われる独我論は、当の5.62の記述から見てとられるように、「**私の言語の限界**が私の世界の限界を意味する」という命題5.6のテーゼと密接に関係している(cf. 5.62)。そして、独我論が常に主体（「我」）の概念を伴うことを考えあわせるならば、この独我論における主体が「形而上学的主体」であると考えすることは自然であるように思われる。

また、上述の想定には別の根拠づけを与えることも可能である。5.64は続けて「独我論の自我は広がりやを欠いた点にまで縮退し、自我に対応する実在が残される」と述べている。この5.64に対する注釈である5.641は、「それゆえ、哲学において、自我について心理学的にではなく論じうる意味が、確かにある」と述べており、この「自我」は「世界の限界」としての「形而上学的主体」であるとされる。ここで5.641の「それゆえ」という箇所は（命

題番号のシステムから) 明確に 5.64 の記述を受けたものであると考えられるため、5.64 の「自我」は「形而上学的主体」のことを指すという解釈が成立可能であるように思われる。ゆえに、この見立てが方向として正しいのであれば、5.64 は 5.62 における「([その] 言わんとするところはまったく正しい) 独我論」と同じ独我論を意味すると述べることができる。

以上より、ラッセル的な読みとショーペンハウアー的な読みは対立することなく共存可能であることを本稿は主張する。そして、この解釈におけるポイントは、5.64 の「独我論」が二種類の独我論、すなわち「否定される独我論」と「肯定される独我論」を二重に意味しうると読むことにある。なお、5.64 では「独我論」の表現が二度現れる。これを踏まえ、一方を「否定される独我論」に、他方を「肯定される独我論」に結び付ける解釈を提示することも(今後の議論として)可能であるように思われる。

(4) 今後の展望

これまでの議論が認められるならば、なお細部を詰める必要はあるものの、今後の展望として『論考』の 5.6 番台における独我論に関する議論を「思考し表象する主体」と「形而上学的主体」という二つの主体概念に着目する形で、二組の〈独我論-主体〉のペア、すなわち、〈ラッセル的な独我論-思考し表象する主体〉および〈ショーペンハウアー的な独我論-形而上学的主体〉として分析するという解釈が成立し得る。この場合、前者は『論考』において論じられているものの、全体として否定され、後者は『論考』において論じられており、かつ全体として肯定される、と整理することができる(これら二つの「独我論」部分には他の独我論解釈を対応させることも可能である)。そして、このように分析することは、『論考』の多様な独我論解釈を〈独我論-主体〉の組をめぐる解釈という形で再構成することを可能とするように思われる。

注

- (1)ここでは、原文におけるイタリック体の箇所を太字で表現する。
- (2)以下、『論考』からの引用は邦訳(野矢(2003))から行い、邦訳の傍点部を太字により表現する。なお、引用箇所は命題番号で示す。

(5) 参考文献

Diamond, C. "Does Bismark Have a Beetle in his Box?", in *The New*

Wittgenstein, ed. by Alice Crary & Rupert Read, 262-292 (London & New York: Routledge, 2000)

Hacker, P.M.S. *Insight and Illusion*, Revised (Oxford: Oxford University Press, 1986)

McGinn, M. *Elucidating the Tractatus: Wittgenstein's Early Philosophy of Logic & Language* (Oxford: Oxford University Press, 2006)

Pears, D.F. *The False Prison*, Vol.1 (Oxford: Oxford University Press, 1987)

ラッセル、バートランド『哲学入門』高村夏輝訳、ちくま学芸文庫、2005年

Tejedor, C. *The Early Wittgenstein on Metaphysics, Natural Science, Language, and Value* (London & New York: Routledge, 2015)

Wittgenstein, L. *Tractatus Logico-Philosophicus*, tr. by David Pears & B. F. McGuinness (London: Routledge & Kegan Paul, 1974) / ウィトゲンシュタイン『論理哲学論考』野矢茂樹訳、岩波文庫、2003年

(千葉大学)

私的言語論におけるウィトゲンシュタインの一つの声：243 節の読みを手がかりにして
A Wittgenstein's Voice of the Private Language Argument : By Using the Readings of §243

佐伯優輔

Abstract

Ludwig Wittgenstein's text, *Philosophical Investigations*, has been interpreted in a variety of ways. In particular, the private language argument has attracted much attention. Orthodox interpreters hold that the argument leads to the impossibility of private language. This article examines two readings of §243, "substantial reading" and "resolute reading". In these readings, there is a difference in drawing Wittgenstein's respective positions. But the difference is not essential, if it does not affect the conclusion of private language argument. Finally, in respect of two readings, this article indicates where Wittgenstein's voice should be heard in *Philosophical Investigations*.

(1) 研究テーマ

ウィトゲンシュタインが著した『哲学探究』（以下、『探究』）は、文法・規則・規準と徴候・生活形式といった様々な概念を考察しているので、数多の議論をうみ、多様な解釈がなされてきた。たとえば、私的言語論と呼ばれる 243 節からなる一連の節についての議論がそれにあたる。これまで私的言語論は、その帰結が私的言語という概念の不可能性を示していると解釈されてきた。換言すれば、心を私的領域として物象化し、その物象化した対象を指示するような言語が想像できないことを示していると解釈されてきた。この解釈により、認識論や心の哲学などの諸哲学に影響を与えると主張されるだけでなく、「私的言語は可能か否か」という論争もなされた。

本稿は、私的言語の可能性も不可能性も論じない。ここで主題とするのは、243 節に関する二つの読みを検討し、私的言語論におけるウィトゲンシュタインの立場を明確にすることである。そのために、243 節についての「実質的な読み [substantial reading]」と「断固たる読み [resolute reading]」の違いを確認したのち、その違いが私的言語論に影響を与えるのかを考察する。そして本稿では、「断固たる読み」の帰結が私的言語の不可能性を覆すものではない限りにおいて、両者の読みの違いが本質的なものではないと結論づける。他方、「断固たる読み」が採用する方法論を参考にして、243 節を対話形式ではなく自己問答としてみる見方も示す。それにより、ウィトゲンシュタインが私的言語の不可能性を主張していないとする読みを提示する。

(2) 研究の背景・先行研究

私的言語という概念は、243 節における定義から始まる。そこが私的言語論の導入である。243 節に対して異なる二つの解釈がなされており、そうした解釈間の相違点を確認するために、以下で該当箇所を参照する。

甲：ところで、ひとが自分の内的体験—自分の感情や気分などを自分用にかきとめたり、発したりできるような言語は考えられないのだろうか。

乙：そうしたことは、我々のいつもの言語でできるのではないか。

甲：いや、そういうことではない。この言語に属する語は、その話し手のみが知りえるもの、つまり話し手の直接的で私的な感覚を指示するのだ。それゆえ、他人はその言語を理解できない。ⁱ

引用文中には、二つの異なる立場（ここでは、〔甲〕と〔乙〕とする）が登場しているⁱⁱ。どちらをウィトゲンシュタインの主張だと捉えるかで、見解の相違がある。

第一に、〔甲〕をウィトゲンシュタインとみなせば、その問いかけから私的言語論が始まる。まずウィトゲンシュタインは、当人の内的体験をその人自身のためだけに表すような言語の想定可能性を問う。この仮定について対話者は、日常的に使われる言語で可能だと反論する。たとえば、誰にも内緒の手記に「痛い」と書くなどのことが考えられるだろう。だが、ウィトゲンシュタインは厳密な定義を示すことでその反論を退ける。というのも、ここで私的言語として想定する言葉とは、その話者の直接的で私的な感覚のみを指し示すものであるためだ。その言葉が指し示すものは当人しか知りえないので、いかなる相手もその言葉を理解することができない。すなわち、その言語使用のうちに他人が登場することは決してない。そのため、その言語使用が人々の間で一致しているかどうかを問うこともない。したがって、もしそのような言語が想定可能であれば、「定義や判断における一致（厳密には一致の可能性さえ）、それは言語の概念そのものとは内的関係にない」（Hacker 2019, p. 15）ことになるⁱⁱⁱ。

このような読みは、「実質的な読み〔substantial reading〕」と呼ばれる。なぜなら、243節以降の議論が、言語の有意味性について実質的な見解を導出すると読み解くからだ。端的に言えば、「私的言語は不可能である」という見解を導くために、「私的言語は可能である」ことを仮定し、その概念上の矛盾を指摘しているとする読みのことである。「もし我々が243節の最初の（実質的と呼ばれる）読みに従うならば、最も有名な後続の諸節—244、246、253、258節—とは、対話者の直前の文における語の意味を考慮して、対話者がそれらの語から組み立てようとする私的言語という考えは無意味か支離滅裂でなければならず、文法違反であるのを、ウィトゲンシュタインが示す箇所となる」（Mulhall 2008, p. 18）。マルハルも言うように、後続する各節は「指示する」や「私的」といった語を我々がどのように使用するかを示して、それらの語に関する文法を浮き上がらせる^{iv}。たとえば、「感覚は私的である」とは経験命題ではなく、文法命題であり、「感覚は当人のみが知っている」とは「知る」という文法に違反した一例である（246-8節、251節）。そして、私的な直示的定義では指示対象の同一性を保証するための正当化ができないことを論拠として、その言語を正しく使用できていないことが示される（258節）。したがって、「私的」や「指示する」などの語を頼りにして私的言語を想像しようと目論むと、言語としての使用の場を失うので、私的言語は不可能であると結論づけられる。

このように私的言語とは、243節の定義に沿って構成される限り、概念上の困難に陥る。もし私的言語が有意味な言葉だとすれば、その語は他人にも理解可能であるため、その定義と齟齬が生じる。だが、内的体験の私秘性を保持し続ける限り、無意味な言葉の羅列にすぎない。なぜなら、そこには有意味性を保証するような基準が存在しないためだ。そうというのも、その使用の正しさは同一の対象によって正当化されており、「私に正しいと思われるものは何でも正しい」（258節）とされるからだ。それにもかかわらず、我々が内的体験の私秘性を言語で表現可能だと考えてしまう一因は、現実の言語使用の場から離れて特別な使用の場を想像してしまうところにある。「実質的な読み」において、ウィトゲンシ

ユタインが文法的諸条件を明らかにしたことを強調するならば、文法違反へと我々を誘う言語観からの解放を試みるような議論として、私的言語論を読むことになる。

第二に、〔乙〕をウィトゲンシュタインとみなせば、対話者の問いかけに対して、ウィトゲンシュタインがそのような想定は普段使用されている言語で可能だと応答していると読める。むしろ自分自身の内的体験に言葉を当てはめることの成立可能性が疑われるなどあるのかと問い直している。だが、対話者は私的言語に形式を与え、その試みを説明する。

このような読みは、「断固たる読み〔resolute reading〕」と呼ばれる。なぜなら、ウィトゲンシュタインが私的言語という考えを断固として拒絶していると読み解くからだ。換言すれば、「私的言語は可能か不可能か」とは議論するまでもなく、「その概念自体がナンセンスである」という答えが一貫して示唆されているとする読みのことである。というのも、その読みでは、「私的」という概念と「言語」という概念とは両立不可能だという前提が暗に置かれているためだ。「もし我々が 243 節の第二の（断固たると呼ばれる）読みに従うならば、これらの後続する諸節とは、対話者の定式化を構成する用語に意味を与える方法を、ウィトゲンシュタインが想像し、試行しようとする箇所として立ち現れる」（Mulhall 2008, p. 18）。マルハルによれば、後続の各節にて、ウィトゲンシュタインは私的言語を定式化しようとする試み、その構成要素としての語が意味をもつのかを検討している。そうというのも、243 節で対話者は日常の言語で表現するという方法を否定するのみであり、私的言語を表現するための具体的な方法に関して一切言及していないからである。そこで、後続する各節にて感覚語を発明する天才児（257 節）や、いわゆる「感覚日記」（258 節）が想定される。けれども、いずれの試みも対話者を納得させるに至らない。「そうした試みに対する対話者の暗黙の応答が、243 節における最初のウィトゲンシュタインの試みに対する対話者の応答を、すなわち『いや、そういうことではない』を繰り返す限りにおいて」（Mulhall 2008, pp. 18-9）、その定義に内容を与えることはできない。

しかしながら、そもそも対話者の問いは成立するのだろうか。対話者は自分自身の言葉によって特別な何かを表現可能だと確信している一方で、現に使用している自分自身の言葉を疑問視しているからだ。つまり、私的言語という形式において特別な意味をもつ言葉を使用可能だと考えていると同時に、その想定例が既存の文法に依拠した言葉では表現不可能だとするからだ。あたかも対話者は、世界のどこかに既婚の独身者がいることを信じているようだ。だが、ある人が既婚者であり、かつ独身者である場合などありえようか。同様に、いかなる文法にも属していない言語というものを我々は想像できようか。

(3) 筆者の主張

これまで 243 節についての二つの異なる読みを検討してきた。この違いは私的言語論の解釈にとって注意を払うべき関心事なのであろうか。すなわち、その違いは私的言語論に関する本質的な争点だと言えるのか。このことを考察するにあたり、私的言語論の本質的な特徴をみていくとしよう。以下は、正統派解釈者たちが私的言語論の前提であると受容してきた点である（Stern 2011, p. 335、括弧内引用者）。

- (1) その議論（＝私的言語論）は、私的言語の性質についての前提もしくは諸前提から始まる。
- (2) その議論は、そのような言語が不可能であるという結論を導いている。
- (3) その結論は、哲学全体に対して広範囲な影響がある。

- (4) その議論は、『哲学探究』で完全にも明確にも述べられていないとはいえ、演繹的な背理法の議論として理解すべきである。

上記のうち、前提(4)を拒絶するような解釈をスターンは示し、その解釈を「ピュロン主義的な読み」と呼ぶ (Stern 2011)。

「ピュロン主義的な読み」によると、「ウィトゲンシュタインの主要な狙いとは、私的言語に関する非常に特殊な概念が矛盾を導くという議論を与えることではなく、むしろ私的言語という考えそのものが首尾一貫して定式化できないことを読み手に分からせること」 (Stern 2011, p. 340) にある。言い換えれば、その読みでは、私的言語という想定は矛盾を生むゆえに不可能だという議論が展開されていると解釈するのではなく、むしろ私的言語という考え自体が断固として定式化できないものだとして解釈する。さらにスターンは、「正統的な読み」と「ピュロン主義的な読み」という対比を、「実質的な読み」と「断固たる読み」の対比へと当てはめる (Stern 2011, pp. 346-7)。そして、「断固たる読み／ピュロン主義的な読み」のように私的言語論を読み解くことが、通説である「実質的な読み／正統的な読み」に取って代わる解釈であると指摘する (Stern 2011)。

スターンによれば、「断固たる読み／ピュロン主義的な読み」とは簡潔に、「正統的な方法論を拒絶するだけでなく、ウィトゲンシュタインの主要な狙いが、**私的言語という考えが矛盾を導くというのを演繹的に証明することにあるという前提も拒絶する**」(Stern 2011, p. 335、強調引用者) とある。それゆえ、スターンは次のように言っている可能性がある。すなわち、「断固たる読み／ピュロン主義的な読み」とは、先に引用した正統派解釈者たちが受容した諸前提のうち、(4)以外の前提を拒絶するものではない、と。もしそうであれば、「断固たる読み／ピュロン主義的な読み」は私的言語論の結論についての前提を拒絶していないことになる。したがって、あくまでスターンは、結論に至るまでの過程の違いを強調しているかもしれない。いずれの読みも、私的言語と呼びたくなる何かに使用の場を与えられないという結論へ到達するのには変わらないからだ。マルハルも指摘するように、「その結論に到達する過程をどちらの方法で我々が劇的に表現するか重要な違いがあるかどうか」 (Mulhall 2008, p. 20) が論点となる。しかし、そこには重要な違いなどあるのだろうか。無論、私的言語論の核心をその結論にみてとるならば、その違いはまったく重要ではない。そして、これまで258節を含む特定の節に対する解釈が集中してきたのは、その議論の核心が私的言語の不可能性にあると考えられてきたからだろう^{vi}。そうだとすれば、過程における差異とは些末なものだと言える。

また、そうした方法論の違いに注意を向けたとしても、どちらの読みがウィトゲンシュタインの主張に忠実か否かを決定づけることはできない^{vii}。そのため、一方の読みを受け入れることが、もう一方の読みを直ちに拒絶するわけではない。そもそも「ウィトゲンシュタインと対話者」という構図それ自体も確立すべきでないかもしれない。というのも、ウィトゲンシュタインの自問自答として、いわば連続した一つの声として再解釈できるかもしれないからだ^{viii}。このことは、二つの異なる読みのどちらに従うとしても、その議論の核心である結論部分に影響を与えないことが裏付けている。それゆえ、243節の〔甲〕と〔乙〕のどちらの立場にウィトゲンシュタインを見出しても誤りではない。いずれもウィトゲンシュタインの主張の一側面にすぎないので、どちらの読みが忠実であるか否かという見方も成立しえない。問いかけることと、その問いに答えることとは必ずしも他人を必要とするわけではない。したがって、ウィトゲンシュタインが問いかけて、その問いにウィトゲ

ンシュタインが答えていると考えてもよいはずだ。

他方、ウィトゲンシュタインの連続した一つの声という読みを支持するためには、「断固たる読み／ピュロン主義的な読み」が採用した方法論に着目し、両者の読みの違いを再認識すべきかもしれない。それだけではなく、二つの異なる読みのどちらに従うとしても、私的言語の不可能性は主張されているという見解を破棄する必要もある。というのも、「断固たる読み／ピュロン主義的な読み」の「定式化できない」(Stern 2011, p. 340) という文言をどう捉えるかで、これまで検討してきた二つの読みは本質的に同じかどうかが決まるからだ。仮に、その文言を私的言語の不可能性の主張として受け取るならば、「不可能である」と「定式化できない」は単なる表現上の違いにすぎないため、いずれの読みも同等である。だが、その文言を字義通りに受け取るならば、私的言語の不可能性は主張されていないこととなる^{ix}。なぜなら、定式化できていないものに問いを立てることはできず、ゆえに、肯定も否定も回答できないからだ。つまり、私的言語は可能であると言えず、ましてや私的言語は不可能であるとも言えない。もしそうであれば、その読みは「(2) その議論は、そのような言語が不可能であるという結論を導いている」(Stern 2011, p. 335) という正統派解釈者たちの前提も拒絶しており、二つの読みは決定的に異なると言える。

そして、そのような読みを「ピュロン主義的」と形容するならば、ウィトゲンシュタインは相反する立場のいずれも支持しておらず、どちらが正しいのかという判断を保留し、私的言語に関するいかなる断言も避けようとしていると解釈すべきである。そのような態度はまた、別の節における記述からもうかがえる。『探究』において「私的言語」という語はほとんど登場しないものの^x、269 節では、「他人は誰も理解しないが、しかし、私が『理解していると思われる』音声を、『私的言語』と呼べるかもしれない」(269 節、強調原文イタリック) とある^{xi}。この節で注目すべきは、私的言語についての見解を断言していないことにある。「呼べるかもしれない」という曖昧な言い回しであり、議論の余地を残しているように見受けられる。

節番号のみが振られた『探究』本文の記述方法は、書物全体の目的はおろか、各節ごとの関係や方向性さえも曖昧に留めておくのに有効である。だからこそ、書かれたことをどう受け取るかは読み手に委ねられている。つい我々が、語り手の声を書き手であるウィトゲンシュタインの声として直ちに同定しようとするのは、平叙文と主張文とを混同するからである。けれども、仮に 243 節をウィトゲンシュタインの自問自答として再構成できるならば、いずれの立場もウィトゲンシュタインと同定すべきではない。つまり、いずれの文中にもウィトゲンシュタインの主張は現れていないと解釈すべきだ。その主張は、肯定文と否定文とを超えたところにある^{xii}。したがって、ウィトゲンシュタインは私的言語が可能であるとも言っておらず、私的言語が不可能であるとも言っておらず、そこでは何も語ってはいないのだ。

(4) 今後の展望

以上より、243 節に関する二つの読みとは、私的言語論の要点である私的言語の不可能性という結論に影響を及ぼすような違いでない限りにおいて、本質的とは言えない。だが、243 節をウィトゲンシュタインの一つの声として読み解くならば、ウィトゲンシュタインは私的言語の不可能性を主張していないと解釈できるので、両者の読みの違いが重要となる。独特な形式で書かれた『探究』において、ウィトゲンシュタインの立場をどこに見出すべきかについて深い見解の相違がある。しかしながら、ウィトゲンシュタインの主張が文中

で語られていることを前提として、平叙文を主張文と即座に判断するのは得策ではない。本稿の読みに従うならば、肯定文と否定文とを重ねた先に、ウィトゲンシュタインの主張が消極的にあらわれているはずだ。このような読みは、『探究』のいくつかの節でのウィトゲンシュタインの位置づけにひとつの視座を与えるかもしれない。

しかしながら、『探究』のいくつかの節は、かつてのウィトゲンシュタイン自身を含んだ論敵が明らかに想定されているため、そうした節の記述を主張文だと認めざるをえない。とはいえ、ある記述をウィトゲンシュタインの主張とみなし、別の記述を論敵の主張とみなすためには、そこで想定されている論敵の主張に一致するであろう内容を、その論敵の著作から援用できなければならない。つまり、『探究』の仮想敵が実際にそう述べているということを、『探究』以外の著作によって裏付けなければならない。そのためには、『探究』での論敵の主張と、『探究』以外での論敵の主張との類似点や一致を明確に示す必要がある。それにもかかわらず、ロックやデカルト、フレーゲやラッセルといった他の哲学者たちが、原著に言及されることもなしに、『探究』の仮想敵として扱われることは少なくない。私的言語論も例外ではないⁱⁱⁱ。無論、そうした解釈が誤りであると断定したいわけではない。そうではなく、いくつかの節において、ウィトゲンシュタインが他の哲学者を念頭に置いて批判していると解釈するためには、その哲学者の著作もまた参照すべきであるということだ。けれども、本稿では、ウィトゲンシュタインと他の哲学者とを比較する解釈について言及するに至らなかった。したがって、こうした点を今後の課題としたい。

注

ⁱ この引用は、『探究』243節による。とはいえ、改行と補足は、引用者が後述の内容を分かりやすくするために、あえて行っている。また本稿では、節番号のみが記載されている場合、『探究』からの引用であることとする。訳は引用者によるものである。

ⁱⁱ 『探究』の何節かは、一方をウィトゲンシュタインとして、他方を対話者として対話篇のように読み取られることがある。たとえば、(飯田 2005, pp. 248-9) はその一例である。

ⁱⁱⁱ 言語の概念と人々の一致とが内的関係にあると主張されるのは、言語が成立するためには複数人による使用が必須であると解釈されるからだ。この解釈では、規則論の帰結がすでに私的言語の不可能性を含意していると考え、(Kripke 1982) がその代表例である。同様の指摘は、(野矢 2022) にある。

^{iv} ウィトゲンシュタインのいう「文法」とは、いわゆる日本語の文法とは異なる。詳しくは、(野矢 2022, pp. 139-142, p. 188) などを参照されたい。

^v (Stern 2011, p. 333) に、具体的な論者が示されている。

^{vi} これまで258節などの特定の節に対する解釈が集中してきたという指摘は、(Stern 2011, p. 338) にある。

^{vii} 同様の指摘は、(Mulhall 2008, p. 20) にある。

^{viii} スターンも、「ウィトゲンシュタインと対話者」という見方を批判している。しかしながら、スターンは二つよりも多くの声が存在すると言う (Stern 2011, p. 338)。

^{ix} 大谷によれば、ウィトゲンシュタインが目指しているのは、「私的言語は想像不可能である」という決定的論証を与えるのではなく、「私的言語」という像を明確化し、吟味することにある (大谷 2020, pp. 162-6)。スターンもまた、(Stern 2011, p. 339) などで同様の主張をしているように思われる。そうとはいえ、スターンの主張は曖昧である。なぜなら、「ピュロン主義的な読み」が前提 (2) も拒絶するとは明言していないからだ。さらには、ピュロン主義の伝統を最初に説明している箇所は、前提 (4) に関する論述であり、一見すると前提 (2) とは無関係であるように思われるからだ (Stern 2011, pp. 338)。

- ^x 「私的言語」という語が登場するのは、259 節、269 節、275 節のみである。
- ^{xi} 中村は、269 節で私的言語と呼ばれるものに具体例を与えたうえで、その語が言語の場に登場しないと指摘している（中村 2021, pp. 150-3）。
- ^{xii} スターンは、「ひとはそこで、『私に正しいと思われるものは何でも正しい』と言いたくなるかもしれない。そして、それは、ここでは『正しい』ということについて語れない、ということにほかならない」（258 節）という箇所に対して、「ウィトゲンシュタインの締め言葉は、私がそのような基準をもてるかもしれないという考えそのものへの強固な拒絶として読むべきだ」（Stern 2011, p. 350）と言っており、その二文をウィトゲンシュタインの主張とみなさないどころか、258 節のうちにウィトゲンシュタインを見出していないように思われる。
- ^{xiii} たとえば、ウィトゲンシュタインは私的言語の不可能性を示すことによって、デカルトが前提としていた心の像を否定していると簡潔に語られることもある。このようにデカルト批判として私的言語論を説明するのは、一般的な解釈として受け入れられてきた。詳細に関しては、（Stern 2011, p. 336）や（伊藤 1995, pp. 167-8）を参照されたい。しかしながら、ウィトゲンシュタインとデカルトをそもそも同じ土俵で比較できるとは限らない。たとえば、伊藤は『第二省察』などを援用することで、私的言語論における内的体験はデカルトの観念説における観念や思考とは異なると指摘する（伊藤 1995, pp. 169-70）。つまり、そもそも心の捉え方がウィトゲンシュタインとデカルトでは異なるため、ウィトゲンシュタインを単純な反デカルト主義とみなせないと述べている。

(5) 参考文献

- Hacker, P.M.S. 2019, *Wittgenstein : meaning and mind*, Wiley Blackwell.
- Kripke, S. 1982, *Wittgenstein on Rules and Private Language*, Harvard U. P. (クリプキ, A. 黒崎宏訳, 2022, 『ウィトゲンシュタインのパラドックス：規則・私的言語・他人の心』, 筑摩書房.)
- Mulhall, S. 2008, “Wittgenstein’s Monologuists (§243)”, *Wittgenstein’s Private Language : Grammar, Nonsense, and Imagination in Philosophical Investigations*, §§ 243-315, Oxford University Press, pp.16-22.
- Stern, D. 2011, “Private Language”, Kuusela, O. & McGinn, M. ed. *The Oxford Handbook of Wittgenstein*, Oxford University Press, pp. 333-50.
- Wittgenstein, L. 2009, *Philosophical Investigations*, revised 4th ed. Hacker, P.M.S. & Schulte, J. eds. Wiley-Blackwell. (ウィトゲンシュタイン, L. 鬼界彰夫訳, 2020, 『哲学探究』, 講談社.)
- 飯田隆, 2005, 『ウィトゲンシュタイン：言語の限界』, 講談社.
- 伊藤邦武, 1995, 「デカルト批判・私的言語の議論」, 飯田隆編, 『ウィトゲンシュタイン読本』, 法政大学出版局, pp. 167-180.
- 大谷弘, 2020, 『ウィトゲンシュタイン 明確化の哲学』, 青土社.
- 中村昇, 2021, 『続・ウィトゲンシュタイン『哲学探究』入門』, 教育評論社.
- 野矢茂樹, 2022, 『ウィトゲンシュタイン『哲学探究』という戦い』, 岩波書店.
(中央大学)

普遍的指令が「適切に」機能するためには何が必要か

What are Necessary for 'Happy' Functioning of Universal Prescriptions?

駒田珠希

Abstract

In this paper I propose the things which are necessary for performing universal prescription 'happily' within Prescriptivism on the basis of Austin's theory of speech act. According to Hare's Prescriptivism, moral statements have two meanings; *descriptive meaning* and *evaluative meaning*. The latter constitutes a prescription, which is a form of illocutionary act; i.e. in saying that 'x is good' one prescribes x. There are three distinctive features of universal prescriptions that distinguish them from mere prescriptions: *universalizability*, *motivational internalism*, and *moral principles*.

Austin proposed the things which are necessary for the 'happy' functioning of performatives. While Austin's original proposal explains mere prescriptions, it does not explain universal prescriptions which exhibit the three distinct features above. This paper solves this problem by proposing the things which are necessary for 'happy' functioning of universal prescriptions.

(1) 研究テーマ

本研究では、Austin(1962)による言語行為論の議論を援用することで、道徳的発話を普遍的指令であると主張する指令主義において、普遍的指令が適切に遂行されるために必要なこと (the things which are necessary for 'happy' functioning of universal prescriptions) がどのようなものであるか検討する。

(2) 研究の背景・先行研究

指令主義とは、道徳文は記述的要素の他に指令的要素を持つとするメタ倫理学上の立場である。この立場によれば、道徳文は「記述的意味 *descriptive meaning*」と「評価的意味 *evaluative meaning*」の二つの意味を持つ。前者は、事実に関する記述的な情報を伝え、後者は、記述的意味によって記述された対象や性質についての指令である。指令主義を提唱した Hare の初期著作 *The Language of Morals* (1961)においては、彼は評価的意味を言語行為

として明示的に述べてはいない。しかし、彼は後に、評価的意味である指令が言語行為であることを認め、さらに、指令主義の先駆けとなった Stevenson の情動主義理論における情動的意味が発語媒介行為であるのに対し、自身の理論における評価的意味は発語内行為であると述べる¹。

言語行為論の提唱者である Austin(1962)によれば、真偽はある言明を評価する際の尺度の一つに過ぎず、特権的なものではない。彼は、言明を正確に評価するためには真偽のみを重視するのではなく、全体的な発話状況をも踏まえた「適切性」の尺度も必要であると述べている²。従って、道徳的発話を指令という言語行為として捉える指令主義において道徳的言明を評価するためには、少なくともその文の発話において遂行される指令という言語行為が適切に遂行されることが必要であるといえる。

Austin は、言語行為が「円滑に、あるいは『適切に』機能するために必要なことから(Austin 1962, p.14=2021, p.35)」として、以下に挙げる(A.1)から(Γ.2)までの六つを提示している。これらに違反した場合、その言語行為は違反したことがらに応じて異なる仕方で不適切なものとなる。ただし、すべての言語行為がすべての形式の不適切さに陥るわけではない³。

言語行為を適切に遂行するために必要な条件：

(A.1)一定の慣習的効果をもつ、受け入れられた慣習の手順が存在していなければならない、その手順は一定の状況において一定の人物による一定の言葉の発話を含んでいる必要があり、そしてさらに、

(A.2)そのケースにおける特定の人物や状況が、呼び起こされた特定の手順の発動にとって適切なものでなければならない。

(B.1)手順は、すべての参加者によって、正確に、かつ

(B.2)完全に実行されなければならない。

(Γ.1)よくあるように、その手順が一定の思考や感情をもつ人物によって行われるようになっている場合、また参加者の誰かの側で一定の後続的なふるまいが開始されるようになっている場合、参加して手順を発動する人物は実際にそうした思考や感情をもっていなければならない、また〔他の〕参加者たちも自身そのようにふるまうことを意図しなければならない、さらに

(Γ.2)引き続き実際にそのようにふるまわなければならない。

(Austin 1962, pp.14-5=2021,pp.35-6)

(A.1)への違反の事例として、次のようなケースが挙げられる。すなわち、キリスト教国に住む両者ともキリスト教徒の夫妻において、夫から妻に向けられた「私は君を離縁にする」のケースである。この場合、離婚を認めないというキリスト教の規範に鑑みると、キリスト教徒による離縁の慣習的手順は存在しえないため、この夫によって発せられた「君を離縁する」の発話が実際に離縁を完遂することはない。

(A.2)への違反の事例として、次のようなケースが挙げられる。すなわち、命名権がない人物が、船首にぶら下がる瓶をたたき割って「私は本船を『ミスタ・スターリン号』と命名する」と宣言するケースである。この場合、船の命名手順は存在し、遂行されてはいるが、行為者に命名権がないため、話者が自称したようには適用できない。

(B.1)への違反の事例として、次のようなものがある。すなわち、家を二つ持っている話者によって言われた「私の家」という語を含む文は、「私の家」という語がどの対象（どちらの家）を指示しているのかが明確でないため、漠然としていて不確定な指示の使用、つまり正確さに欠ける場合となる。

(B.2)への違反の事例として、次のようなケースが挙げられる。すなわち、話者が「私はあなたと6ペンスの賭けをする」と言っても、相手が「よし、乗った」と応じなければ「賭ける」という行為は実行されない場合である。

(Γ.1)への違反の事例として、次のようなケースが挙げられる。すなわち、「約束する」と言いながら約束を果たす意図がない場合である。

(Γ.2)への違反の事例として、次のようなケースが挙げられる。すなわち、「歓迎いたします」と言った後に相手をぞんざいに扱う場合が挙げられる。

Austinは、以上の(A.1)から(Γ.2)がすべての言語行為の適切さを規定するものではないと述べる一方、「非常に多くの倫理的行為は、全体として、あるいは部分的に、慣習的もしくは儀式的な性格を一般にもっており、それゆえとりわけ不適切さへの危険にさらされているのである」と述べる⁴。それでは、道徳的発話を普遍的指令という言語行為であるとする指令主義において、普遍的指令を適切に遂行するためにはどのようなことが必要だろうか。指令主義における指令を発語内行為として扱うことは一般的である⁵ものの、それが適切に遂行されるために何が必要であるかについては明確に示されていない。本稿では、発語内行為としての普遍的指令が適切に遂行されるために必要なことを、Austinの言語行為論の議論を援用しつつ再構成することを目標とする。

(3) 筆者の主張

道徳的発話の評価的意味が発語内行為としての指令を適切に遂行するためには、道徳的発話もまた、一定の基準を満たす必要があることになる。単なる指令であれば、Austin の挙げている条件で十分かもしれない。しかし、指令主義における道徳的発話は、単なる指令と区別される重要な特徴として、次の三つの特徴を持っている。一つ目は普遍化可能性である。これは、ある記述的性質 p を持つもの x に対して「 x はよい」と判断を下した場合、その発話後は、 x と同様の記述的性質 p を持つ全てのものに対して「よい」と判断するように拘束される、という原理である⁶。二つ目は、動機付けの内在主義である⁷。これは、道徳判断と行為の動機付けの間に必然的な関係がある、つまり道徳判断を下したならば、当該の行為を導くように動機づけられると考える立場である。三つ目は、道徳判断は「道徳原理 moral principles」を参照して下される⁸、というものである。この道徳原理とは、人々によって設立されたものではあるが、修正と継承を繰り返すことで歴史的・文化的に確固たる地盤をもつという意味での客観性を有する原理である⁹。

以上の特徴をもつ普遍的指令を適切に遂行するために必要なことを検討するにあたり、Austin の議論を援用する。以降では、普遍的指令を適切に遂行するために必要であると考えられることとそれぞれに違反した際の不適切な事例を挙げる。

普遍的指令を適切に遂行するために必要な条件：

(A.1)*：発話された判断は一般に受け入れられている道徳原理に含まれる必要がある。

(A.2)*：下された判断は、首尾一貫した判断者によって下される必要がある。

(B.1)*：道徳言明の記述的意味が正確である必要がある。

(B.2)*：なし

(Γ.1)*：一定の思考や感情を持った判断者によって、真剣に下された判断である必要がある。

(Γ.2)*：発話された判断は、発話者の後続行為を拘束する必要がある。

(A.1) は手順の存在・非存在に関する条件である。Austin はこの条件における不適切な例としてキリスト教徒による「離縁する」という発話を挙げている。これは普遍的指令の遂行においては、次のように言うことができる。すなわち、受け入れられた道徳原理には「盗みをすべきだ」という普遍的指令を与える手順は存在しない。そのため、「盗みをすべきだ」という発話は普遍的指令になりえない。

(A.2)は、手順が自称されたように適用できるか否かに関する条件である。Austinはこの条件に違反する例として、命名権のない人物による命名行為を挙げていた。指令主義においてこれと対応する不適切さは次のようなものとなる。すなわち、首尾一貫しない人物や移り気な判断を下す人物は、Austinの例における命名するのに適当でない人物と同様で、普遍的指令を与える人物として適当でない。指令主義において、ある状況に置かれている人物Aに対して「嘘をつくべきでない」と言い、それと同様の状況に置かれている人物Bに対しては「嘘をつくべきだ」と言う人物は、「べき」という語の意味を理解していないと見なされる¹⁰。そのため、このような人物による普遍的指令の手順は、話者が自称したようには適用できない。普遍的指令を与えるには、「よい」や「べき」の語の意味を理解しており、首尾一貫した判断を下せる状態にある人物による発話である必要がある。

(B.1)は手順の正確さに関する条件である。Austinの例では、複数の家を所有する人物による「私の家」という語の発話は、その語の指示対象が曖昧であるため不正確であるというものが挙げられていた。普遍的指令の場合、これは次のようになる。すなわち、道徳言明の持つ記述的意味が曖昧である場合、その道徳的発話は適切に普遍的指令を遂行できないというものである。募金をしていることを理由にある人物を「彼はよい人だ」と言う際、この発話は記述的意味として「募金をすること」をもつため、募金を指令していることが明らかである。しかし、何を理由に「彼はよい人だ」と言っているのか明らかでないとき、つまり、その言明の記述的意味が不明確であるとき、その言明の発話は何を指令しているのか明らかではないため、指令として正確ではない¹¹。

(B.2)は手順の完全さに関する条件である。この条件におけるAustinの例は、ある人が「私はあなたと6ペンスの賭けをする」と言ったとしても聞き手が「よし、乗った」と応答しなければ行為は完遂されないというような、発話者の発話とは別の要因が障害となって行為が遂行されない例であった。普遍的指令の場合においては、この条件は必要にならない。というのも、普遍的指令は相手の応答とは関係なしに遂行可能であるためである。相手が指令どおりに振舞うかどうかは発語内行為としての指令とは独立の問題である。Hareによると、Stevensonの情動主義における道徳言明が相手を説得して特定の振る舞いをさせるもの(getting him to do)であるのに対して、指令主義の主張はあくまで「何がよいか」について教えるものである(telling someone to do)¹²。そのため、相手の応答とは独立に普遍的指令を遂行することは可能である。以上のことから、普遍的指令の適切性条件として、Austin

の挙げる(B.2)に対応する条件は不要であるといえる。

(Γ.1)は誠実さに関する条件である。本当はよいと思っていないのに「募金をするのはよい」と言う場合、この発話は、Austinの言葉を借りれば「うわべだけの」、Hareの言葉を借りるならば「引用符付き」あるいは「慣用表現」のものとなる¹³。この場合、その発話は聞き手に社会的事実としての情報、例えば、募金をすることは、「人びとが一般に認めている基準に従うために要求されている¹⁴」という社会的事実の言明を与えることはあるが、発話者自身はその判断に与しておらず、その対象を本気で指令しているわけではない。そのため、この発話は評価的意味としての指令を欠いたものとなる。

(Γ.2)は、後続行為についての条件である。Austinによるこの条件への違反として、「歓迎する」と言いながらその発話の後に続く行為が聞き手をぞんざいに扱うものであるケースを挙げていた。この条件を普遍的指令の場合に当てはめた場合、後続行為は次の二種類、①判断と②振る舞いに分けて考えられる。

①は、普遍化可能性に関する。もし記述的性質 p を持つために「 x はよい」と判断したにもかかわらず、記述的性質 p を持つ y に対しては「 y はよくない」と言うのであれば、その人は「よい」という語の意味を知らないか、誠実に判断したわけではないと見なされるため、その発話が普遍的指令を遂行することはない。

②は、動機づけの内在主義に関する。つまり、「募金をすべきだ」という判断をした場合にはその発話の後続行為として実際に募金をするよう動機づけられなければならないというものである。もしも「募金をすべきだ」と述べながら実際には募金をするよう動機づけられないのであれば、①への違反と同様、発話者は「よい」や「べき」という語を理解していないか、誠実に発話したわけではないとみなされる。(Γ.2)*をまとめると、ある道德判断を発話した場合、①それに後続する判断は最初の判断と矛盾したものであってはならず、②それに後続する振る舞いは下された判断に沿って動機づけられたものでなければならない、というものになる。

以上に、Austinの議論を基にして普遍的指令を適切に遂行するために必要なことがらを提示した。Austinの提示した条件では、単なる指令の適切性は説明可能であるかもしれないが、指令主義における普遍的指令の適切性を説明するには不十分であった。本稿で挙げた(A.1)*から(Γ.2)*は、指令主義における道德的発話の三つの特徴を押さえている（普遍化可能性は(Γ.2)*の①、動機づけの内在主義は(Γ.2)*の②、道德原理の参照については(A.1)*で説明できる）点で、普遍的指令が適切に遂行されるための尺度を提供できる。

(4) 今後の展望

本稿では、道徳的発話の評価的意味にあたる普遍的指令を適切に遂行するために必要なことがらを提示した。Hare (1997)は、一方で、道徳的言明の真理適合性を認めており、記述的意味がその真理条件であると述べている¹⁵。彼によれば、記述的意味によって記述されている事実が存在する場合には、その道徳的言明は真、そうでない場合には偽となる。彼の挙げている例は次の通りである。すなわち、優しく、寛容で、トランプでずるをしない人物 A に対する「A はよい人だ」という発話の記述的意味は、「優しく、寛容で、トランプでずるをしない」ことである。この発話が真となるのは、この発話の記述的意味（すなわち真理条件）を満たす場合となるので、実際に A が優しく、寛容で、トランプでずるをしないとき、かつそのときに限る、というものである¹⁶。

この記述を踏まえると、指令主義における道徳的発話の評価するには、一方で評価的意味における適切性を、他方で記述的意味に関係する真理を検討する必要がある。それでは、これら二つの評価尺度はそれぞれ完全に独立したものなのか。普遍的指令を適切に遂行するには真理は必要ではないのか。今後は、指令主義における適切性と真理の関係を検討する必要がある。

注

1. Hare 1997 pp.15-6
2. Austin 1962 p.3=2021 p.17
3. Ibid pp.18-9=2021 p.40
4. Ibid pp.18-20=2021 pp.40-1
5. 児玉 2022 pp.184-5
6. Hare 1963 p.15=1982 p.31
7. Hare 1961 p.162=2003 pp.214-5
8. Ibid p.78=2003 p.104
9. Ibid pp.146-7=2003 pp.193-4
10. Ibid pp.80-1=2003 p.108
11. Hare によれば、記述的意味はよさの判定基準や情報を与えるという、「よい」という語の仕事の半分を担っている（この語のもう半分の仕事は指令である）。しかし、この記述的意味が不明確な場合、その語は、その記述的意味が明らかになるまでは無意味になる (ibid pp.116-7 =2003 pp.153-4)。

12. Ibid p.13=2003 pp.17-8
13. Austin 1962 p.16=2021 p.37, Hare 1961 pp.124-5=2003 pp.163-6
14. Hare 1961 p.167=2003 p.221
15. Hare 1997 p.52
16. Ibid pp.52-3

(5) 参考文献

- 児玉聡(2022)「第9章 反實在論・非認知主義」:『入門・倫理学』、赤林朗・児玉聡編、勁草書房、pp.177-193
- Austin, J.L. (1962) *How to do things with words*, Oxford University Press
(飯野勝己訳、『言語と行為 いかにして言葉でものごとを行うか』、講談社、2021)
- Hare, R.M. (1961) *The Language of Morals*, Oxford University Press (小泉仰・大久保正健訳、『道徳の言語』、勁草書房、2003)
- (1963) *Freedom and Reason*, Oxford At The Clarendon Press
(山内友三郎訳、『自由と理性』、理想社、1982)
- (1970) “Meaning and Speech Acts”, *The Philosophical Review*, Vol.79, No.1, Duke University Press, pp.3-24
- (1997) *Sorting Out Ethics*, Clarendon Press・Oxford
- Stevenson, C.L. (1944) *Ethics and Language*, Yale University Press (島田四郎訳、『倫理と言語』、内田老鶴圃、1990)

(北海道大学)

外界懐疑論に対するカヴェルの応答
Cavell's Response to External World Skepticism

森内康太

Abstract

Stanley Cavell deals with skepticism of the existence of the world mainly in Part II of his book *The Claim of Reason*. As far as I can see, he distinguishes between the kinds of skepticism he is dealing with, but he does not explicitly indicate the distinction. This study investigates what kinds of skepticism Cavell responds to and how he responds to them. In conclusion, this study shows that he is responding to each of the two types of skepticism: high-standards skepticism and radical skepticism about the criteria of justification.

(1) 研究テーマ

本研究では、スタンリー・カヴェルの外界懐疑論への応答を扱う。カヴェルは主にその著書 *The Claim of Reason* の第II部において外界懐疑論の問題を扱う。筆者の見る限り、カヴェルは自身が相手とする懐疑論の種類について区別を行ってはいないが、その区別について明示的に論じてはいない。そこで、本研究は、カヴェルの外界懐疑論に対する応答を再構成し、それがどのような種類の外界懐疑論に対して、どのような仕方で応答を行なっているものであるかを明らかにすることを目的とする。

本稿の構成は次の通りである。まず(2)研究の背景・先行研究では二つのことを論じる。2-1節では外界に対する懐疑論が大きく二つに分類可能であることを確認する。2-2節では、カヴェルの記述とそれに関する先行研究の調査をもとに、カヴェルの外界懐疑論に対する応答を再構成する。次に(3)筆者の主張では、再構成したカヴェルの議論が、先に確認した二種類の外界懐疑論のそれぞれへの応答になっていることを論じる。最後に(4)今後の展望では、本研究が今後取り組む問題を示す。

(2) 研究の背景・先行研究

2-1.二種類の懐疑論

ここでは、Williams(2001)の記述に基づいて、外界懐疑論を、外界に関する我々の知識をどのように疑うかという観点から二つに分類する。

一つ目は、「高い基準の懐疑論[high-standards skepticism]」である。高い基準の懐疑論は一般に、「我々の信念の多くがプラスの認識論的地位[positive

epistemic status]—高いプラスの認識論的地位さえも—を持つことは認めるが、それが適切にいわゆる知識に相当するほど十分に高い地位を有していることを否定する」(Williams,2001:6). この懐疑論が知識として認めるのは、「否定不可能なほど[indefeasibly]に正当化された真である信念」と、「絶対的な確実性」をもって正当化された真である信念(Williams,2001:6)である。

この種の懐疑論を外界に適用した場合の主張は次のようになる。すなわち、我々の知覚は外界についての信念をもつことをある程度は正当化しうが、その信念を確実性をもって正当化することはできない。そのため、我々は外界についての知識をもたない。

二つ目は、「徹底的懐疑論[radical skepticism]」である。徹底的懐疑論は、「正当化された信念をもつレベルに達する我々の能力すらをも否定する」(Williams,2001:6)懐疑論である。この懐疑論のもとでは、我々は「認識論的地位に関して根拠のある区別をつける」(Williams,2001:6)ことさえもできないということになる。

高い基準の懐疑論では、知覚に基づいて形成された信念は、知覚に基づかずに形成された信念よりも正当化されているということは認められていた。例えば、目の前にリンゴがある（レモンはない）ような知覚経験を得た際には、「目の前にリンゴがある」という信念は、確実性をもった正当化はなされていないとしても、「目の前にレモンがある」という信念よりは正当化されているということは認められていた。

しかし徹底的懐疑論ではそれさえも否定される。たとえリンゴの知覚経験を得ていたとしても、「目の前にリンゴがある」という信念と「目の前にレモンがある」という信念は、どちらも同じようにまったく正当化されていないということになる。したがって、徹底的懐疑論のもとでは、外界懐疑論の主張は次のようになる。すなわち、我々の知覚は外界についての信念をもつことをまったく正当化しない。そのため、我々は外界についての知識をもたない。

2-2.カヴェルの外界懐疑論に対する応答

ここでは、カヴェルの著作 *The Claim of Reason*(Cavell,1979, 以下 CR と表記)第II部、特に外界懐疑の問題について詳しく論じている章である‘The Quest of Traditional Epistemology: Opening’と‘The Quest of Traditional Epistemology: Closing’の記述と、それについて論じた先行研究として Pritchard(2022)の記述をもとに、カヴェルの外界懐疑論に対する応答を確認する。

カヴェルは、自身が問題とする懐疑論者は「我々誰もが最良の状態の知識を例証していると認識する[外界についての]知識の事例に疑問を投げかける」(CR:135)懐疑論者であると述べる。この懐疑論者は「この事例において私は知らない、あるいは知ることができないのだから、私はいかなる事例においても知ることができない」(CR:133)という形で、その主張を一般化することを目的とする。

カヴェルによれば、このような懐疑論者はまず「この事例において私は知らない」ということを帰結させるために、「ある主張の例が与えられ；その根拠が明示され；それに答えられなければ当の主張が否定されるような疑いの根拠が提起される」(CR:132)という形式の論証を用いる。そしてこの論証によって、「その主張が「明らかに誤りである」ということではなく、主張者はやそれが真であると主張することができないということ」(CR:132)を帰結させる。

例としてカヴェルは「私はここにテーブルがあると知っている」という主張についての懐疑を次のように構成する。

根拠の要求： 例えば、私はどうしてここにテーブルがあると知っているのか？

根拠： なぜならば、私はそれを見ているからだ。あるいは：感覚によって。

疑いの根拠： a)しかし、私は本当は何を見ているのか？私はもしかしたら夢を見ていたり、幻覚を見ていたりするかもしれないのではないか？

b)しかし、それでは十分ではない。それはデコイかもしれない。

c)しかし私はそのすべてを見てはいない。私が見ているほとんどは...

結論： それゆえ、私は知らない。(CR:144)

先に述べた通り、ここでの懐疑論者はその懐疑を外界についてのあらゆる知識に一般化させようとするが、カヴェルによれば、そのためには論証の出発点となる主張が、特定の対象[specific object]（例えばそのゴシキヒワ[the goldfinch]、ルイ 14 世の書物机）についてのものではなく、一般的な対象[generic object]（例えば、テーブル一般）についてのものでなくてはならな

い(CR:138). ここで一般的な対象と言われているのは、それらについて問題が生じるとしたら、「識別、特定、描写の問題」ではなく、「それらが存在しているかどうか、本物であるかどうか、実際にそこにあるかどうかを私たちが知ることができるか」(CR:52)という問題であるような対象である。

カヴェルによれば、特定の対象について私が知識を持たないことが明らかになったとしても、「そのことは、いったい何が知られうるかについての、あるいは全体として、プロジェクトとしての知識についての含意を与えず、あなたの訓練不足や判断における軽率さ、あるいはこの事例における機会や物理的条件の相対的な貧しさについての含意しか与えない」(CR:133). それに対して、一般的な対象についての事例は、「個人的な訓練や判断の慎重さといった要素や、機会や条件といった文脈上の問題が無関係な事例」(CR:133-4)であるために、知識全体についての帰結をもたらすことができる。

カヴェルは以上のようにして懐疑論者の論証を構成した上で、そもそも懐疑論者が出発点として想定する（想定せざるをえない）一般的な対象についての言明は主張として成立しえないものであり、従って懐疑論者はジレンマに陥らざるを得ないことを次のように指摘する。

伝統的な知識の探求が巻き込まれるジレンマは、このように定式化することができる：もしその手続きが首尾一貫したものであるべきならば、それは具体的な主張の探求でなくてはならない；もしその結論が一般化されるべきならば、それは具体的な主張の探求ではあり得ない。その首尾一貫性がなければ、それはそれが持つように思われる明白さを持っていない；その一般化がなければ、その結論は懐疑的なものではあり得ない。(CR:220)

ではなぜ、一般的な対象についての言明は首尾一貫した有意味な主張ではあり得ないのか。カヴェルによれば、それは「あなたが言うことが理解可能であるべきならば、文法上、あなたが言うことの原因やあなたが何かを言うことのポイントがなくてはならない」(CR:208)が、一般的な対象についての発言にはそれらがないためである。つまり一般的な対象は誰もがそれをそれと識別できるような対象であり、また識別を妨げる要因もないような環境に置かれているものであるから、我々はそれについて「私はここにあると知っている」という言明がなされる理由やポイントを見出すことができない。したがってそれを主張として理解することができないのである。

またカヴェルによれば、懐疑論者は、疑いの根拠として「私には対象の一部しか見えていない」という事実を提示するときにもジレンマに陥っている。もしこの主張が一般的な対象に向けられたものであるとするならば、ここで懐疑論者は我々と対象との関係を、我々の現実におけるモデルとは異なるモデル、すなわちまるで我々の位置が対象に対して固定されており、常に対象の半分しか見えていないようなモデルで考えている。したがって懐疑論者の主張は我々の知識に対しては何の影響も及ぼさない。またこの主張が我々の知識に対して影響を及ぼすためには、それは具体的な主張でなくてはならず、そのとき懐疑論者は我々の知識一般に疑いを向けることができない (CR:202-203)。

したがって、懐疑論者の論証は、疑いの対象として我々に理解不可能な主張を置いており、また疑いの根拠を示すために我々の現実とは異なるモデルを設定しているということになる¹。

ただしカヴェルによれば、懐疑論者が構成する論証は合理的なものではないが、懐疑論者が世界についての知識全体に対する問いを立てること自体は自然である。というのも、このような問いは、「言語を所有するに足りるほど複雑な生き物、あるいは、重荷を負わされている生き物の自然な経験を表現する反応」(CR:140)であるためである。

では我々はどのようにしてそのような問いに至るのか。カヴェルは一例として、「それ[懐疑論の問い]はあなたが「完全に確信している」あるいは「確実であると思っている」事柄について、あなたが明らかな仕方で間違っていたことに対する反応として生じる」(CR:140)ことがあるとする。つまり我々は、自分が確実なものであると考えていた事柄が実は間違っていたという経験から、自分が現在知っていると思っているすべての事柄について、我々はそう思い込んでいるだけで、本当は知らないのではないか、という一般化された疑いに向かってしまうことがある。

しかしながら、カヴェルによれば、このような経験から帰結するのは、「我々人間は可謬的である」という道徳(CR:143)であって、先に示したような懐疑論ではない。このような経験にさらに別の種類の経験、すなわち「自分の感覚は自分がそれについてのものだと思っている世界についてのものではまったくないかもしれない、あるいは、自分が知ることができるのは、対象がどのようにあるように(我々に)見えるかだけであって、対象それ自体がどのようなものであるかは決して知ることができない」と悟るという経験が加わった時に、我々は「徹底的なあるいは形而上学的な懐疑論」(CR:143)に至っ

てしまう。

しかしこのような経験に至ることが自然であると認めるとしても、カヴェルは「私は世界が存在していることを知らない」という主張が正しいということをも認めるわけではない。それは、そもそも世界全体に対する我々の関係は「知ること」(CR:241)ではなく、「世界は受容される[*accepted*]べき」(Cavell,2002:298)ものであるためである。

Pritchard(2022)は、ここでのカヴェルの指摘は「[ワイトゲンシュタインの]『確実性について』でさらに明確に示されている」(Pritchard,2022:269)とする。ここでは詳しく論じることはできないが、ワイトゲンシュタインはそこで「我々の最も日常的な確信は[...]実際には完全に理論の及ばない仕方[*arationally*]抱かされている」ものであり、「合理的評価が可能であるために適当なところになくなくてはならない根本的な蝶番のような確信」(Pritchard,2022:269)であるとしている。Pritchardの解釈を踏まえるならば、カヴェルは、「世界が存在する」という確信は、我々が普通知識と考えるもの、つまり何らかの正当化のプロセスを経て知識となったものではなく、我々が日常的に様々な合理的な探求を行うためにその前提として受け入れており、それを疑うということが意味をなさないようなものであると主張していると考えられる。

(3) 筆者の主張

ここでは、以上で確認したカヴェルの議論が、どのような懐疑論に対するどのような応答となっているのかを論じる。

まず、人は自分が確実だと思っていた事柄が実は間違っていたということから、他の事柄についても自分は間違っているかもしれないため、本当は自分は何も知らないのだという懐疑論に陥ってしまうことがある。これは知識の可謬性から生じる懐疑論であるから、高い基準の懐疑論であると言える。カヴェルはこの種の懐疑論に至ったとしても、それは不可謬主義を取ろうとする立場からの帰結であって、我々の知識が可謬的であることを道徳として認めれば問題ない考える。

しかしここで懐疑論から抜け出すことができず、さらに別の種類の経験が加わったときに、懐疑論者は、我々は夢を見ているかもしれない、あるいは我々は対象の一部しか見ていないといったことから、結果として世界の存在に関する知識を疑ってしまうことになる。これは我々の感覚による信念の正当化自体を疑う懐疑論であるから、徹底的懐疑論であると言える。カヴェル

はこのような懐疑論者に対して、二つの観点から応答する。

一つ目は、我々が世界の存在を知り得ないという結論に至るために懐疑論者が構成する論証は、ジレンマに陥らざるを得ないというものである。しかしこれが指摘するのは、あくまでも懐疑論者の論証がうまくいっていないということであって、世界が存在することが証明可能であるということではない。したがって、我々は世界の存在を知り得ないという主張自体は反駁されおらず、また我々は世界の存在を知っているという主張がなされているわけでもない。

そこでカヴェルが提示するのが、「受容[acceptance]」(Cavell,2002:298)という概念である。カヴェルは、世界の存在に関する信念は、我々が他の知識と同じように何らかの正当化プロセスを経てたどり着いたものではなく、我々が生活の前提としてあらかじめ受け入れているものであるということを示す。そしてそれによって、世界の存在に関する信念の正当化を疑おうとする懐疑論者に対して、世界の存在についてはそもそも正当化が問題にならないということを示している。

以上より、カヴェルの議論は、高い基準の懐疑論と徹底的懐疑論の、それぞれへの応答となっていることが明らかとなった。

(4) 今後の展望

筆者の見る限り、カヴェルの外界懐疑論に対する応答においては、懐疑論者が提示する主張が主張として成立していないという点が重要性を持っている。本稿ではその点について十分な検討をすることはできていない。しかし、懐疑論者が提示する主張が主張たり得ていないということは明らかではない。そのため、カヴェルの外界懐疑論に対する応答を説得力のある形で再構成する為には、この箇所についてのカヴェルの言語観を踏まえた検討が必要となる。

それから、本稿ではカヴェルが用いる「受容」という言葉は、「世界は存在する」ということを我々が正当化プロセスなしに受け入れているということを示していると解釈している。これは先に述べたとおり、Pritchard(2022)の解釈によったものであるが、この解釈が妥当であるかという点も含め、本研究では「受容」概念を十分に検討することができていない。したがって、カヴェルの「受容」概念については、筆者の解釈を根拠づけるために更なる議論が必要となる。

本研究では、引き続き、以上の点を詳細に検討することによって、カヴェルの外界懐疑論に対する応答をより詳細に再構成し、その説得力を評価する

ことを目的とする。

注

1. カヴェルによれば、懐疑論者が提示する、根拠についての主張（「私はその見ている」や「感覚によって」）もまた、それが「特定の主張の特定の根拠」としてではなく、世界について何かを知る際の一般的な方法として提示しようとする限りにおいて、ジレンマが生じている（CR:160 参照）。

(5) 参考文献

- Cavell, Stanley. *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*. New York: Oxford University Press. 1979.
- “The avoidance of love: A reading of King Lear.” *Must We Mean What We Say?: a book of essays* Updated edition. Cambridge University Press. 2002.
- Pritchard, Duncan. Putnam on radical skepticism: Wittgenstein, Cavell, and occasion-sensitive semantics. In J. Conant, & S. Chakraborty (Eds.), *Engaging Putnam* (pp. 263-288). (Berlin Studies in Knowledge Research; Vol. 17). De Gruyter. Advance online publication. 2022.
<https://doi.org/10.1515/9783110769210-012>(2023/10/27 にアクセス)
- Williams, Michael. Contextualism, externalism and epistemic standards. *Philosophical Studies* 103 (1):1 - 23. 2001.

(千葉大学)

根拠づけと帰結の関係の小考
A Brief Consideration of The Relation between Grounding and
Consequence

中川和彦

Abstract

This paper aims to briefly consider the relation between grounding and consequence. In this paper the concept of consequence is characterized in the following way. B is a consequence of A, if and only if, necessarily, if A then B. Although many grounding theorists agree that consequence is a necessary condition for grounding, they don't agree the inverse for some reasons. I propose that a constrained form of consequence becomes a sufficient condition for grounding. Finally I will show two things the proposal suggests and the shortcoming it has.

1 研究テーマ

本稿では、近年分析形而上学分野において注目を浴びている根拠づけ（グラウンディング）関係（grounding relation）と帰結関係の関係について簡単に（大まかで非形式的な仕方ではあるが）考察する。ここで帰結（consequence）関係と呼ぶものは演繹的推論関係あるいは必然化関係とも呼ばれうるもので、以下の実例において成立しているものとする。

1. 雪が白いということからは雪が白いということが真であるということが帰結する。
2. ソクラテスが存在するということから $\{\text{ソクラテス}\}$ が存在するということが帰結する。
3. ソクラテスが存在するということとアリストテレスが存在するということからは、ソクラテスが存在しかつアリストテレスが存在するということが帰結する。
4. 冬よりも夏の方が温度が高いということからは、冬よりも夏の方が正常に機能する温度計の数値が高いということが帰結する¹。

以降では差し当たりこの関係を、標準的な仕方で²次のように特徴づけておく。まず帰結関係を命題間の関係とする。この関係を記号「 \Rightarrow 」によって表し、命題 A_1, \dots, A_n, B について、 $A_1, \dots, A_n \Rightarrow B$ である（ A_1, \dots, A_n から B が帰結する）のは A_1, \dots, A_n が真であるならば B が真である、ということが必然的である時かつその時に限る、とする。

根拠づけ関係は本稿では事実間関係とする。それは以下のようなものを
実例とする、因果関係とは異なる世界におけるある種の先行関係、順序関係
である。

1. 雪が白いという事実は雪が白ということが真であるという事実を根拠
づける。
2. ソクラテスが存在するという事実が { ソクラテス } が存在する という事
実を根拠づける。
3. ソクラテスが存在するという事実とアリストテレスが存在するという事
実が、ソクラテスが存在しかつアリストテレスが存在するという事実を
根拠づける。
4. 冬よりも夏の方が温度が高いという事実が、冬よりも夏の方が正常に機
能する温度計の数値が高いという事実を根拠づける。

根拠づけ関係は全面的 (full) 根拠づけと部分的 (partial) 根拠づけに二分さ
れる。上記は全面的根拠づけの例である。全面的根拠づけを記号「 \prec 」によっ
て表し、諸真命題 Γ と真命題 A にそれぞれ対応する事実 $[\Gamma], [A]$ に対し、

$$[\Gamma] \prec [A]$$

によって「 $[\Gamma]$ は $[A]$ を全面的に根拠づける」ということを意味する。部分的
根拠づけを記号「 \prec 」を用いて全面的根拠づけから次のように定める。

$[A] \prec [B]$ であるのは、その内に $[A]$ を含むような諸事実 $[\Gamma]$ があり、 $[\Gamma] \prec [B]$
である時かつその時に限る。

また本稿では全面的、部分的根拠づけ関係を共に非反射的、非対称的、推
移的な関係とする³。

2 研究の背景・先行研究

根拠づけ関係と帰結関係について、以下のことは多数の論者が認めている
ところである⁴。

$$[\Gamma] \prec [\varphi] \text{ ならば } \Gamma \Rightarrow \varphi.$$

ここで Γ は少なくとも一つの命題を表す。これはすなわち帰結関係が (全面
的) 根拠づけ関係の必要条件になっているという主張である。

他方、逆の十分条件の主張は同意が得られていない。それには主に三つの理由がある。

第一に根拠づけ関係は事実間関係であり、(通例) 事実は真命題のみに対応し、偽命題に対応する事実はない。他方帰結関係は真命題のみならず偽命題間関係でもある。例えば太陽系の惑星の数は9であるという偽命題から太陽系の惑星の数は奇数であるという偽命題が帰結するが、両者に対応する事実はなくそれゆえ根拠づけ関係は成り立たない。

第二に擬似必然化の問題 (the problem of spurious necessitation) と呼ばれる⁵ ものの存在である。これは帰結関係は前件が必然的に偽あるいは後件が必然的に真であれば成立するが、必ずしもその前件と後件間に根拠づけが成り立つとは限らないという問題である。例えば、 $2 + 3 = 5$ は必然的真理であるので、任意の真命題 A を前件とし $2 + 3 = 5$ を後件とする帰結関係が成立するが、必ずしも $[A]$ と $[2 + 3 = 5]$ とで根拠づけ関係が成り立っている訳ではない。

第三に非対称的根拠づけ間の対称的必然化の問題 (the problem of symmetric necessitation between asymmetrically grounded entities) や非対称性問題と呼ばれる⁶ ものの存在である。これは(擬似的でない) 相互帰結関係が認められるのに、根拠づけ関係は相互に成り立つことはないという問題である。例えば、ソクラテスが存在すると $\{\text{ソクラテス}\}$ が存在するとは相互帰結関係が認められている。すなわち、ソクラテスが存在するならば $\{\text{ソクラテス}\}$ が存在する、ということが必然的であると同時に、 $\{\text{ソクラテス}\}$ が存在するならばソクラテスが存在する、ということが必然的である。他方で $[\text{ソクラテスが存在する}]$ と $[\{\text{ソクラテス}\} \text{ が存在する}]$ とには片方向の根拠づけ関係すなわち $[\text{ソクラテスが存在する}] < [\{\text{ソクラテス}\} \text{ が存在する}]$ のみしか認められない。

3 筆者の主張

以上のようにいくつかの問題があり帰結関係が根拠づけ関係の十分条件であることの同意は得られていない。他方、しかしながら、上の第三の非対称性問題から示唆されると思われるが、「真命題 A, B に関して、 B が帰結するために A が必要不可欠な仕方存在するのであれば、両者に対応する事実の間でどちらか片方向の根拠づけが成立している」といった直観は得られうる。

そこで筆者が提案したいのは、一定の制限を加えた帰結関係が部分的根拠づけ関係に十分である、ということである。その提案は具体的には次のようになる。

A, B を真命題、 Γ を諸真命題とし、それぞれ個別にあるいは合わさって

矛盾することがないとする。以下の三つの条件を満たす場合、 $[A] \prec [B]$ か $[B] \prec [A]$ のいずれかが成り立つ。

- (1) $A, \Gamma \Rightarrow B$
- (2) $(\Gamma \not\Rightarrow B)$ あるいは $(\Gamma \Rightarrow A$ かつ $\Gamma \Rightarrow \Delta$ かつ $\Delta \not\Rightarrow B)$
- (3) $[A] \neq [B]$

(ただしここで $\Gamma \Rightarrow \Delta$ は諸命題 Δ におけるどの命題も Γ から帰結することの略記とし、 Δ の全ての命題から別々にあるいは合わさって A が帰結しないこととする)。

大まかに言えば、第一の条件で A と Γ から B が帰結すること、第二の条件でその帰結関係に A が必要であること、第三にしかし A と B について、それらに対応する事実としては異なることを述べている (そしてこの条件により根拠づけ関係の、自身が自身を根拠づけることはないという非反射性が保たれる)。

例えば、 $A =$ ソクラテスが存在する、 $B = \{ \text{ソクラテス} \}$ が存在する、 $\Gamma =$ アリストテレスが存在する とすると、上の (1) から (3) の条件を満たすので、 $[A] \prec [B]$ か $[B] \prec [A]$ のいずれかが成り立つことが導かれる。そして実際 $[A] \prec [B]$ は正しい。

「 $[A] \prec [B]$ か $[B] \prec [A]$ のいずれかが成り立つ」というように選言とする理由は、上の三つの条件が、 $A =$ ソクラテスが存在する、 $B = \{ \text{ソクラテス} \}$ が存在する、とした場合と、 $A = \{ \text{ソクラテス} \}$ が存在する、 $B =$ ソクラテスが存在する、とした場合のいずれの場合も満たされるからである。本稿の提案のもともとの動機は本節の冒頭に記したように「真命題 A, B に関して、 B が帰結するために A が必要不可欠な仕方 で存在するのであれば、両者に対応する事実の間でどちらか片方向の根拠づけが成立している」という直観を言い表すことである。ソクラテスと $\{ \text{ソクラテス} \}$ の場合のように、相互帰結関係にある場合に主張できることは、少なくともどちらか片方向には (部分的) 根拠づけが成立しているということである。すなわち異なる真命題 A と B について、 A から B が帰結するときに、 $[A]$ が $[B]$ を根拠づけていると言い切ることは不可能であるが、少なくとも $[A]$ が $[B]$ を根拠づけているかあるいは $[B]$ が $[A]$ を根拠づけていると言い切ることはできる。そしてこの選言によって第三の非対称性問題が回避されている。

上の条件の (2) の「あるいは」の右側の条件が必要な理由は次のような場合の根拠づけを認めるためである。

例えば、 $A = \exists f; (f = [X]), B = (X \wedge Y) \wedge \exists f; (f = [X]), \Gamma = X \wedge Y$ とすると、 $[A] \prec [B]$ は成り立つ。しかし $X \wedge Y \Rightarrow X$ であり、かつ $X \Rightarrow$

$\exists f; (f = [X])$ である⁷ので、 $\Gamma \Rightarrow B$ が成り立つ。他方、 $\Gamma \Rightarrow X, \Gamma \Rightarrow Y$ かつ $Y \neq B$ であり、このことが示すのは「あるいは」の右側の条件を満たしていることである。従って、この根拠づけを導くために「あるいは」左側のみならず右側の条件が必要となる。

以上のような条件づけによって、根拠づけの十分条件を考えた場合に生じる前節の問題を回避していることは明らかである。第一の問題が回避されているのは、真命題に範囲を絞ることにしたという点に尽きる。第二の擬似必然化の問題は(2)によって回避されている。第三の問題が回避されるのは上の通りである。

注意すべきは(3)に関わる事実の同一性の問題である。上の根拠づけの十分条件の主張が成り立つためには、以下のような二つの事実を同一とみなす必要がある。それは $A = a$ は b よりも大きい、 $B = b$ は a よりも小さい、である場合の $[A], [B]$ である。この時(1),(2)は満たすように思われるので、(3)の同一性を認めるのであれば条件三つ全て満たすが、実際には $[A] \prec [B]$ でも $[B] \prec [A]$ でもないように思われる。従って事実の同一性条件を $[A] = [B]$ を認める仕方では与えなければならない⁸。

最後にここで本稿と類似した提案である McDaniel (2022) との差異について言及しておく。彼は根拠づけ関係の特徴づけのために最少必然化 (minimal necessitation) 関係を用いている。この提案と本稿との大きな差異は、前者がその必然化関係の特徴づけるのに帰結関係とは異なるであろう複雑性 (complexity) 概念を用いている点である⁹ (根拠づけ関係が単純なものから複雑なものへの関係であるという考えが背景にあるのであろう)。本稿の提案においては帰結関係に付される条件としてそのような、根拠づけをそれへと還元することがもしかしたら可能でありそうであつた帰結関係とは異なる別概念に訴えてはいない。その上、このことに関連して、少なくともこの提案では単純、複雑という関係にないもの間で言われているところの根拠づけの主張をなすことができない。例えば、上述の「冬よりも夏の方が温度が高いという事実が、冬よりも夏の方が正常に機能する温度計の数値が高いという事実を根拠づける」という主張をなすことはできないであろう¹⁰。

4 今後の展望

前節の、一定の制限を加えた帰結関係が部分的根拠づけ関係に十分であるという主張が正しいとみなされた場合に考えられることは、第一に根拠づけ関係の内実理解に帰結関係が重要な役割を占めている可能性があるということである¹¹。第二にこの十分性の主張が、(帰結関係をもとに) ある事実とある事実が根拠づけ関係にあることを示すための基準として用いることが出来

るということである。

しかしながらこの十分性の主張は、根拠づけ関係の必要条件にまでは至ることが出来ない。というのも、 B が必然的命題である場合に、任意の (諸) 命題 Γ に対し、 $\Gamma \Rightarrow B$ が成り立つので (2) の条件が満たされないからである。確かに擬似必然化の問題を回避しているが、 B が必然的命題である時に許容されるものとそうでないものの区別が出来ていない。この一因は帰結関係を厳密含意的なものとして捉えてしまうことであろう。帰結関係、あるいは演繹的推論関係あるいは必然化関係をそれとは別に捉えた場合¹² の上の条件を満たした根拠づけの十分性の主張がどのようなものであるか検討することが今後の課題である¹³。

注

¹ この例は Tatzel(2002)[10], p.8 や Poggiolesi(2020)[7], p.221 に倣った。

² Schnieder(2018)[9], s1335 を参照。

³ その他根拠づけの前提については、Correia(2012)[2] のイントロダクション、Fine(2012)[4] に従う。根拠づけ全般については Bliss(2021)[1] を参照のこと。

⁴ 例えば Fine(2012)[4], p.38 など。

⁵ MacDaniel(2022)[5], p.155 を参照。ただしこの呼び名の「必然化」は本稿で定めた帰結関係としての必然化とは意味が異なる。

⁶ MacDaniel(2022)[5], p.156 及び秋葉 (2016)[11], p.88 を参照。

⁷ Fine(2010)[3], p.100 を参照。

⁸ ここで問題になるのは事実のきめ細かさ (granularity) である。これについては Correia(2012)[2], p.14 などを参照。

⁹ MacDaniel (2022)[5], p.160。加えて、そもそも本稿と根拠づけ関係の関係項が異なる。彼はその関係項に事実以外の存在者を入れることをも許容している (MacDaniel (2022)[5], p.155-156, 159)。

¹⁰ 詳細は省くが、具体的にはそれら事実が (恐らく事実の複雑性をどのように考えたとしても) MacDaniel の最少必然化の諸条件の内 PC (proportionality constraint) に反しているであろう (MacDaniel (2022)[5], p.161)。また、少なくとも彼の提案から言えないであろう根拠づけのその他の例としては、「神が存在するという事実が宇宙が存在するという事実を根拠づける」(Rabin and

Rabern(2016)[8], p.7) などがある。

¹¹ 秋葉 (2016)[11] では、根拠づけ関係が、本稿で言うところの帰結関係の最善の体系における一種の依存関係である、という根拠づけの統合説が提唱されている。

¹² 本稿では扱いきれなかったが、関連含意 (relevant implication) と根拠づけ関係を結びつけて論じる文献がある。(Poggiolosi (2020) [6] を参照。)

¹³ あるいはこの帰結関係を根拠づけ関係を用いて明らかにするという方向性も考えられる。

文献

- [1] R. Bliss and K. Trogon. Metaphysical grounding. In *Stanford Encyclopedia of Philosophy*. 2021.
- [2] F. Correia and B. Schnieder eds. *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press, Cambridge, 2012.
- [3] K. Fine. Some puzzles of ground. *Notre Dame Journal of Formal Logic*, 51(1):97–118, 2010.
- [4] K. Fine. Guide to ground. In F. Correia and B. Schnieder eds., *Metaphysical Grounding*, pp. 37–80. Cambridge University Press, 2012.
- [5] B. McDaniel. Grounding as minimal necessitation. *Inquiry*, 65(2):154–175, 2022.
- [6] F. Poggiolosi. Grounding principles for (relevant) implication. *Synthese*, 198(8):7351–7376, 2020.
- [7] F. Poggiolosi. logics. In M. J. Raven ed., *The Routledge Handbook of Metaphysical Grounding*, pp. 213–227. Routledge, 2020.
- [8] G. O. Rabin and B. Rabern. Well founding grounding grounding. *Journal of Philosophical Logic*, 45(4):349–379, 2016.
- [9] B. Schnieder. On ground and consequence. *Synthese*, 198(Suppl 6):1335–1363, 2018.
- [10] A. Tatzel. Bolzano’s theory of ground and consequence. *Notre Dame Journal of Formal Logic*, 43(1):1–25, 2002.
- [11] 秋葉剛史. 根拠づけ (grounding) に関する統合説. 科学哲学, 49(1):85–90, 2016.

(東京大学)

文脈主義による無欠陥な不一致への応答とその射程
Contextualism's reply to faultless disagreement and its scope

檜岡寛己

Abstract

In this essay, I investigate how contextualism about predicates of personal taste replies to the criticism in terms of faultless disagreement, and whether that reply is successful. The criticism is that although contextualism can explain faultlessness, it cannot explain disagreement. But, it does provide some ideas to explain disagreement. After outlining these ideas, I consider whether contextualism successfully explains faultless disagreement, and I suggest that it does explain faultless disagreement in one interpretation but not in another.

(1) 研究テーマ

小論では、個人的趣味述語(predicates of personal taste;以下、ppt)の意味論の文脈主義(contextualism)が無欠陥な不一致(faultless disagreement)へどのような応答をもたらし、その応答が上手くいっているのかを探究する。

本論に入る前に ppt の外延をいくつか示したい。多くの論者は、「楽しい(fun)」、「美味しい(delicious, tasty)」のような主観的な述語を ppt の典型例だと見なす。本論では、「リンゴは美味しい」という文を、ppt を含む文のモデルケースに採用し、以降の議論を進めていくことにする。

2節では、ppt の文脈主義、無欠陥な不一致、無欠陥な不一致によって文脈主義が抱える問題、文脈主義によるこの問題への応答を紹介する。3節では、文脈主義による無欠陥な不一致への応答は上手くいっているのかを考察し、その応答はある意味では上手くいっているが別な意味では上手くいっていないことを示唆する。4節では、今後の研究の展望を簡単に述べる。

(2) 研究の背景・先行研究

2-1.ppt の文脈主義

典型的な ppt の文脈主義は、ppt を含む文には趣味に相対化された隠れた指標的な構成要素が含まれていると考える(Cappelen&Hawthorne 2009)。話し手 A が「リンゴは美味しい」と発話したとしよう。この場合、その趣味は典型的には話し手のものであり、文脈主義によれば、話し手 A によるこの発

話が表す命題は《リンゴは A にとって美味しい》というものである。以上が ppt の文脈主義的な意味論である。2-2 では、無欠陥な不一致を見ていく。

2-2. 無欠陥な不一致

Kölbel(2004)によれば、無欠陥な不一致とは、思考者 A、思考者 B、命題（判断の内容） p が存在する以下のような状況である。

(a) A は p と信じ（判断し）、B は $\neg p$ と信じる（判断する）。

(b) A も B も間違いを犯していない（欠陥があるのではない）。

(Kölbel 2004: 53-54)

この定式では、(a)は不一致を表していて、(b)は無欠陥を表している。また、Zeman(2019)によれば、この定式には二つの特徴がある。一つ目は、(a)において、不一致は命題のみならず、信念ないし判断に言及する仕方で定式化されている。二つ目は、(b)において、欠陥の意味について説明がなされていない。一つ目の特徴に関して言えば、不一致に関する説明では、信念ないし判断という命題的態度は省略されて、話し手それぞれの発話が表す命題が両立しないというような仕方で説明されることが多い。二つ目の特徴に関して言えば、Zakkou(2019)によれば、二つの意味がある。一つ目は、ppt を含む発話によって表現される命題が真であると考えることが正当化されるために、話し手は必要なことをすべて行っているという意味での無欠陥である。二つ目は、話し手による ppt を含む発話が表す命題が真であるという意味での無欠陥である。

2-3. 文脈主義が抱える不一致問題

しかし、無欠陥な不一致に対して、文脈主義は話し手それぞれが無欠陥であることを説明できる一方で、話し手同士の間には生じる不一致を説明できないと言われてきた。以下の談話(1)を見てほしい。

(1)A と B は、以下のそれぞれの発話によって表現される命題が真であると考えることが正当化されるために必要なことをすべて行った上で、真剣かつ誠実に以下のようなやりとりをする。

A: 「リンゴは美味しい」

B: 「いや。リンゴは美味しくない」

文脈主義によれば、(1)における A と B のそれぞれの発話が表す命題は(2)のそれぞれとなる。

(2)A: 《リンゴは A にとって美味しい》

B: 《リンゴは B にとって美味しくない》

文脈主義によれば、(1)の談話では、A の命題と B の命題はどちらも真であると言えるだろう。そのため、文脈主義は、両者の発話が表す命題が真であるという意味で、A と B が無欠陥であることを説明できる。一方で、A の命題と B の命題は矛盾をきたすものではなく、両立する。そのため、文脈主義は不一致を説明できないと指摘されてきた(以下、このことを不一致問題と述べる)。

しかし、López de sa(2022)によれば、不一致には排除的内容が必要だとする考え方は明らかに不当な制限である。通常的不一致の概念はもっと柔軟であり、排除的内容を含む必要はないのだ。実際に、多くの文脈主義者が不一致をより柔軟に捉え、不一致問題に応答しようとしてきた。2-4 では、文脈主義による不一致問題への応答の一つとして語用論的内容に訴える説明を見ていく 1。

2-4.文脈主義による不一致問題への応答

2-4-1.優越性の前提

Zakkou(2019)は、ppt を含む文が優越性(superiority)の前提を語用論的に伝えると主張し、不一致問題に応答する。Zakkou(2019)によれば、優越性の前提とは、話し手がそれぞれ自分の趣味を相手の趣味よりも優れていてベストであると考えた前提である。そして、(1)のような談話において A も B も、(2)のような命題を意味論的に表しているだけでなく、優越性の前提を語用論的に伝える。結果として、(1)のような談話において、A と B によって伝達される語用論的内容は以下の(3)となる。

(3)A: 「リンゴは A にとって美味しいし、リンゴに関しては A の趣味の基準がベストだ。」

B: 「リンゴは B にとって美味しくないし、リンゴに関しては B の趣味の

基準がベストだ。」

優越性の前提の説明が正しければ、文脈主義は、(1)の談話において A と B はそれぞれ(2)の命題を表し、(3)の内容を語用論的に伝達していると説明する。そして、この説明では、A と B は、命題が真であるという意味で、共に無欠陥であり、語用論的内容に関して A と B の間に不一致が生じているということになる。

2-4-2.メタ言語的交渉

文脈主義は、不一致問題を説明するための別な仕方としてメタ言語的交渉に訴える仕方も提案してきた。Sundell(2011)、Plunkett&Sundell(2013, 2021)によれば、メタ言語的交渉とは、二人の話し手が、ある言語表現を用いて、その言語表現がある文脈ではどのように使用されるべきかをめぐる規範的な論争である。そして、メタ言語的交渉を用いて、文脈主義は、(1)のようなやりとりが(2)のような命題を表しながら、(4)のようなメタ言語的交渉もしていると考ええる。

(4)A:「A と B の文脈では、美味しいという語はリンゴに当てはまるような基準で用いられるべきである」

B:「A と B の文脈では、美味しいという語はリンゴに当てはまるような基準で用いられるべきではない」

メタ言語的交渉の説明が正しければ、文脈主義は、(1)の談話において A と B はそれぞれ(2)の命題を表しつつ、それぞれ(4)のようなメタ言語的な伝達をしているのであり、この伝達された語用論的内容で不一致が生じていると説明することができる。そして、この説明では、A と B は、命題が真であるという意味で、共に無欠陥であり、語用論的内容に関して A と B の間に不一致が生じているということになる。

(3) 筆者の主張

これまで、文脈主義による無欠陥な不一致への応答を見てきた。では、文脈主義のこれまでの説明によって、無欠陥な不一致は説明されたことになるのだろうか。私は、ある意味では無欠陥な不一致が説明されていて、また別な意味では無欠陥な不一致が説明されていないと考えている。では、それはどういうことなのだろうか。その説明のために、まずは Lowe(2006)による 4

カテゴリー存在論を簡単に導入したい。

3-1.4 カテゴリー存在論

Lowe(2006)は存在論の基本的なカテゴリーとして、対象(Objects)、様態(Modes)、種(Kinds)、属性(Attributes)の四つを導入する 3。対象は個別の実体であり、そのカテゴリーには、この人間というような個体が属する。様態は個別的性質ないし個別関係であり、そのカテゴリーには、この壁の白さというような個別的性質、AとBの間の特定の友情関係というような個別関係が属する。種は実体的普遍者(substantial universals)であり、そのカテゴリーには、人間のような生物種が属する。属性は非実体的普遍者(non-substantial universals)であり、そのカテゴリーには、白いという普遍的性質、友情関係のような普遍的関係が属する。

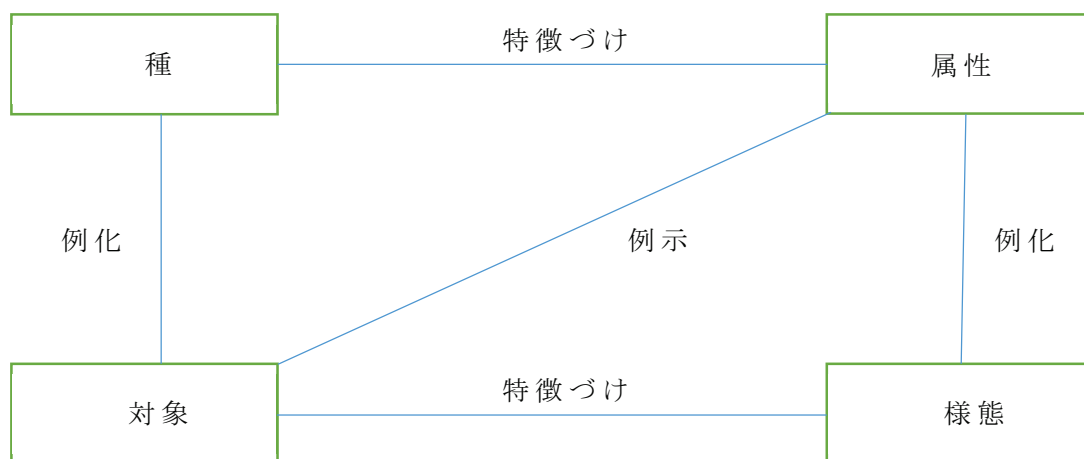


図 1. 存在論的スクエア

ここで示した四つのカテゴリーの間には例化(instantiation)関係、特徴づけ(characterization)関係、例示(exemplification)関係という三つの形式的関係が成り立つ 4。例化関係は、種カテゴリーに属する存在者と対象カテゴリーに属する存在者の間に、また、属性カテゴリーに属する存在者と様態カテゴリーに属する存在者の間に成り立つ。例えば、人間という種は、メッシというような個別の実体に例化される。同様に、より背が高いという関係は、ロナウドはメッシより背が高いというような様態に例化される。特徴づけ関係は、種カテゴリーに属する存在者と属性カテゴリーに属する存在者の間に、また、対象カテゴリーに属する存在者と様態カテゴリーに属する存在者の間に成り立つ。例えば、後者なら、ロナウドとメッシはより背が高いという様態によって特徴づけられる。例示関係は、対象と属性の間に成り立つ。そし

て、例示関係は例化関係と特徴づけ関係によって定義づけられ、対象は様態を経由して属性へと例示関係に立つ 5。例えば、ロナウドとメッシはより背が高いという属性を例示する場合、ロナウドとメッシは実際により背が高いという様態によって特徴づけられ、より背が高いという様態はより背が高いという属性を例化するということになる。以上が、Lowe(2006)による 4 カテゴリー存在論の概要説明である。

3-2.無欠陥な不一致から得られる二つの命題への二つの解釈

Kölbel(2004)の無欠陥な不一致の定式は、A と B の判断の間に不一致がある(①)こと、A と B が共に無欠陥である(②)ことを表している。これまで、不一致は A と B の命題間に生じる不一致である(③)として捉えられてきた。無欠陥に関しては、両者の発話によって表現される命題が真であるという意味で A と B は無欠陥である(④)。また、不一致は語用論的内容での不一致である(⑤)とする文脈主義によるアイデアも考慮に入れて議論をする必要がある。

文脈主義では説明ができなかったところの無欠陥な不一致は、(①)、(②)、(③)、(④)によって説明され、次の(5)の 1 の命題が真であることにコミットすることになるだろう。一方、文脈主義が提示する無欠陥な不一致は、(①)、(②)、(④)、(⑤)によって説明され、次の(5)の 2 の命題が真であることにコミットすることになるだろう。

- (5)1.《A と B の発話によって表現される命題はそれぞれ真であり、これらの命題の間には不一致が生じている》
- 2.《A と B の発話によって表現される命題はそれぞれ真であり、両者の発話によって伝達される語用論的内容の間には不一致が生じている》。

ここで、3-1 の 4 カテゴリー存在論で示した対象と属性の間に成り立つ例示という形式的関係を考慮すると、無欠陥な不一致から導くことができる(5)の 1、2 にはそれぞれ、不一致関係と真であるという性質に関して成り立つ規則について、以下の 2 つの解釈の余地があるように思われる。

- 解釈 1.命題が不一致関係を例示し命題が真であるという性質を例示することが求められる。あるいは、語用論的内容が不一致関係を例示し語用論的内容が真であるという性質を例示することが求められる。
- 解釈 2.命題が不一致関係を例示し命題が真であるという性質を例示することが求められる、というわけではない。あるいは、語用論的内容

が不一致関係を例示し語用論的内容が真であるという性質を例示することが求められる、というわけではない。

3-3. 文脈主義は無欠陥な不一致を説明できるのか

3-2 で得られた二つの解釈を 2 節での文脈主義の説明に照らし合わせてみよう。また、Zakkou(2019)による無欠陥への説明を適切に拡張することで、両者の発話によって表現される命題が真であるという意味での無欠陥を、両者の発話によって表現される命題/語用論的内容が真であるという意味での無欠陥として言い換えられると仮定してみよう。すると、以下の①~④の説明が得られるだろう。

- ①.真であるという性質も不一致関係も二つの特定の命題に例示される(解釈 1 に適合する)6。
- ②.真であるという性質も不一致関係も二つの特定の語用論的内容に例示される(解釈 1 に適合する)。
- ③.真であるという性質が二つの特定の命題に例示され、不一致関係が二つの特定の語用論的内容に例示される(解釈 2 に適合する)。
- ④.真であるという性質が二つの特定の語用論的内容に例示され、不一致関係が二つの特定の命題に例示される(解釈 2 に適合する)。

2 節における文脈主義の説明を考慮すると、不一致関係は A による発話が伝達する語用論的内容と B による発話が伝達する語用論的内容に例示され、真であるという性質は A による発話が表す命題と B による発話が表す命題に例示されると言えるだろう。これは③にあたり、解釈 2 に適合する。したがって、文脈主義による説明では、真であるという性質を例示している対象と不一致関係を例示している対象が別である。そのため、文脈主義の説明では、解釈 1 を説明することができず、さらに解釈 1 はもともとの無欠陥な不一致から導かれる解釈であるはずなので、無欠陥な不一致が適切に説明されているとは必ずしも言えないだろう。

2 節での文脈主義の説明では、無欠陥な不一致の解釈 2 を問題なく説明できるようには思われるが、解釈 1 を説明できないように思われる。というのも、文脈主義の説明によれば、不一致関係と真であるという性質を例示する対象が同じではなく、別だからだ。文脈主義は不一致を意味論的に説明するのが難しいというこれまでの反論を考慮すると、①と④はあてにはできないので、本論が示すところによれば、無欠陥な不一致から導かれる解釈 1 を説

明するために文脈主義が取り得る戦略候補は②となるだろう。

(4) 今後の展望

註6でも示したように、相対主義は解釈1をうまく説明できるように思われる。そのため、相対主義が解釈2を説明できるのかを考察したい。その上で、もし相対主義が解釈2を説明できないのだとしたら、解釈1と解釈2を説明し尽くす理論を構築するのか、それとも相対主義と文脈主義を無欠陥な不一致に関する解釈によって使い分けるのかを探究したい7。また、これまでの議論を踏まえると、相対主義と文脈主義は無欠陥な不一致に関して異なる解釈に適う理論を組み立てたことになり、無欠陥な不一致に関して両者に論争はないと評価することができるかもしれない。さらに、文脈主義と相対主義は無欠陥な不一致というそもそもの現象をそれぞれ異なる現象に還元していて、どちらがもっともらしく還元しているのかが論争的になると考えることもできるかもしれない。これらは重要な論点であるように思われるので、これらについても探究したい。

註

1. 語用論的な応答には López de Sa によるアイデアをはじめとして多くあるが、紙幅の都合により、本論では 2-4 の二つのみをあげる。

2. メタ言語的交渉を意味論的に説明する余地もある。詳しくは Zakkou(2019)、Plunkett&Sundell(2021)、小田(2021)を参照してほしい。しかし、Plunkett & Sundell (2013)がメタ言語的交渉は語用論で説明されると考えていること、Plunkett & Sundell(2021)でもメタ言語的交渉は語用論で説明されると考えることを否定していないことの二つを理由に、本論では、メタ言語的交渉は語用論で説明されるものとして仮定している。

3. Lowe(2006)の解釈やその説明の仕方は倉田(2017)を参考にした。

4. 訳語は倉田(2017)を参考にした。

5. 本論では触れないが、対象が種を経由して属性へと例示関係に立つ仕方もある。

6. 文脈主義が対立する相対主義(Relativism)はまさにこのような説明をすることができるように思われる。相対主義による意味論には三平(2019)が詳しい。

7. この点は飯川遥氏に負う。

(5) 参考文献

- Cappelen, H & Hawthorne, J, 2009, *Relativism and monadic truth*, Oxford University Press.
- Kölbel, M, 2004, "Faultless Disagreement.", *Proceedings of the Aristotelian Society* 104, 53-73.
- López de Sa, D, 2022, "Disagreements and Disputes About Matters of Taste.", in *Perspectives on Taste Aesthetics, Language, Metaphysics, and Experimental Philosophy*, edited by Jeremy Wyatt, Julia Zakkou and Dan Zeman, Routledge.
- Lowe, E. J, 2006, *The four-category ontology: a metaphysical foundation for natural science*, Oxford Clarendon Press.
- Plunkett, D & Sundell, T, 2013, "Disagreement and the Semantics of Normative and Evaluative Terms." *Philosopher's Imprint* 13, 1-37.
- , 2021, "Metalinguistic Negotiation and Speaker Error.", *Inquiry* 64, 142-167.
- Sundell, T, 2011, "Disagreements About Taste.", *Philosophical Studies* 155, 267-288.
- Zakkou, J, 2019, *Faultless Disagreement: A Defense of Contextualism in the Realm of Personal Taste*, Klostermann.
- Zeman, D, 2019, "Faultless disagreement.", in *The Routledge Handbook of Philosophy of Relativism*, edited by Martin Kusch, Routledge.
- 小田拓弥, 2022, 「メタ言語的交渉をめぐる研究について」, *科学哲学* 54 (2), pp. 93-111。
- 三平正明, 2019, 「相対主義意味論の可能性」, *日本大学文理学部人文科学研究 研究所研究紀要* 97, pp. 1-30。
- 倉田剛, 2017, 『現代存在論講義 I』, 新曜社。

(無所属)

問答的理由説にもとづく規範理由の区別へ向かって

Towards a Distinction of Normative Reasons Based on the Erotetic View
of Reasons

池田 開

Abstract

Many philosophers have attempted to provide reductive definitions of normative reasons, with prevalent views being the Reasoning view and the Explanation view of reasons. Amidst this controversy, Artūrs Logins develops and defends a new theory: the Erotetic view of reasons, according to which normative reasons are appropriate answers to normative ‘Why?’ questions (Why should I do this?). This paper not only affirms the capacity to distinguish normative reasons within the framework of the Erotetic view but also critically examines the shortcomings of Logins’ argument and identifies avenues for future work.

(1) 研究テーマ

理由とは何か？という問いは論争的である。本稿では、特に規範理由とは何かについて扱う。なかでも、本稿で論じたいテーマは、規範理由に関する還元的定義と規範理由の分類である。

(2) 研究の背景・先行研究

理由の還元的な定義を与える方法のひとつに、その役割を考えること、すなわち、理由の機能的分析がある。理由には、少なくとも2つの役割があるとされている(McHugh and Way 2022; Logins 2022)。それは推論的な役割と説明的な役割である。

第一に理由は、実際に私たちが何をすべきか、どのような態度を取るべきかを考え出すのに役立つ。ある行為をすべきかどうか考えるとき、私たちは理由を参照してその行為をすべきだと判断するだろう。このことから理由は、熟慮や推論を作り出す材料、すなわち推論の前提とも言えるだろう。私たちは推論を踏まえてある行為をしたり、ある態度を取ったりする。このことから理由には、行為者が実際に行為をするよう導く役割があると言えるだろう。これが理由の推論的役割である。

第二に理由は、私たちが何をなすべきか、どのような態度を取るべきかを

理解可能にするという役割がある。理由は、なぜ私たちがある行為をすべきなのか、ある態度を取るべきなのかといった規範的地位を決定し、説明する。

例えば、もうすぐ 23 時だということは、サッカープレミアリーグの試合が始まろうとしていると考えてもよいということであり、しかし明日は朝 6 時から仕事があるということは、サッカーの試合をリアルタイムで観戦することを諦めるべきだということになる。明日は朝 6 時から仕事があるということは、実際に行為を導くかどうかはさておき、サッカーの試合をリアルタイムで観戦することを諦めるべきだということを説明している。これが理由の説明的役割である。機能的分析によって理由に還元的定義を与える試みは、理由のこの 2 つの役割を掬えるものでなければならないとされてきた (McHugh and Way 2022, 13–16)。

理由の推論的役割に着目した議論は、Williams (1979) や Raz (1999) をはじめとして、最近では McHugh and Way (2022) によって展開されている。この立場にはさまざまなバリエーションがあるものの、共通するのは次のような見解である。

考慮事項 R は、R が F すること（すなわち帰結態度）に向かう主体 S の可能な良い/健全な推論における前提態度の内容である場合にのみ、S が F するための規範理由となる。(Logins 2022, 39–40)

この立場は、規範理由と動機づけ理由とのかかわりを重視している。理由の内在本義のように、規範理由は動機づけ理由にもなると主張する立場もあるため、規範理由は動機づけ理由の部分集合であると、この立場では前提されている。そのため、この見解は動機づけ理由のどの部分集合が規範理由であるのかを特定する提案として、推論的見解を提示している。この見解では、規範理由は、行為者が F することを引き受けることができる推論の良いパターンの適切な前提に対応する潜在的または実際の動機づけ理由である (Logins 2022, ch.2)。この立場では推論の健全性、つまり、良い（パターンの）推論とはどのような推論かということについて、各論者は価値適合性や欲求基底性を用いて論じている (McHugh and Way 2022; Smith 1994, 161–64=218–221)。

他方、理由の説明的役割を重視する立場では、規範理由を「べし (ought)」と説明によって定義している。Broome は次のように規範理由を定義する。

〔行為者〕N にとって F することの *pro toto* な理由とは、なぜ N が F すべきかの説明である。(Broome 2013, 50)

pro toto な理由とは、他の理由によって凌駕されない理由のことであり、この Broome の定義に従うならば、F すべきことの説明項 (explanans) になっているならば、それは pro toto な理由である。つまり、ここでの「べし」とは規範的要請であり、「説明」は説明することを表しているのではなく説明項となる事実のことである。このように、Broome は、べしと説明項によって規範理由を定義するのである (Broome 2013, ch.4)。

それぞれの立場には、それぞれ反論がある。推論的な見解に対しては、良い推論とは何かという反論と、そもそも理由によって支持される全ての行為や態度が、常に良い推論パターンによって導かれることに対応しているわけではないという反論がある。さらには、良い推論を考えたときに、理由ではなく、行為を遂行するための単なる背景条件が前提になることもある。これらは、理由の説明的役割を推論的な見解が掬えていないことに起因する反論だろう。なぜなら、推論は行為者の行為を導く心の働きなのであって、行為者の行為を説明するためのものではないからだ。他方、説明的な見解には、私たちの実際の行為を理由がどのように導くのが明らかではないという反論がある。これは理由の推論的役割を説明的な見解が掬えていないことに起因する反論である。このようにして推論的な見解に対しては説明的役割に起因する反論 (Logins 2022, secs.2.4-2.6) が、説明的な見解に対しては推論的役割に起因する反論が提出されている (Logins 2022, ch.3)。

このような規範理由の還元的説明にかんする論争の中で Logins は理由の間答的な見解 (Erotetic view) を提案する (Logins 2022)。Logins によれば規範理由とは、規範的な問いに対する適切な回答である。Logins はこの見解を詳細に明示化した定式として、以下を提示している。

理由の間答的な見解 (明示的): p が S が F する理由だということは、その p が (a) 説明を必要とするように読める、(S に向けられた)「なぜ F するのか/なぜ S は F すべきなのか」という問いに対する答え (のパターン) を提供する適切な説明の内容 (の一部) であるか、(b) 論証を必要とするように読める、(S に向けられた)「なぜ F するのか/なぜ S は F すべきなのか」という問いに対する答え (のパターン) を提供する適切な論証や推論の前提の内容である、ということである。(Logins 2022, 170)

この Logins の見解は「なぜ F するのか?」という問いに着目する。非形式論理では、「なぜ?」の問いには、回答として説明を求める読みと、論証を求める読みとの 2 種類の問いがあるとされている (Logins 2022, 166-67; cf. McKeon 2013)。ここでは、この非形式論理の区別を用いることで、理由の推

論的役割と説明的役割のそれぞれを擲うことができる見解を示している。

「なぜ F するのか？」という問いは、一方では、F すべきだという事実の説明を求めている場合がある。このとき、「なぜ F するのか？」という問いでは、F すべきであるということは前提されている。そのため、この読みでは、質問の発話者は、どのような根拠に基づいて F すべきなのか、ということをお問うている。他方、別の回答が必要とされる場合もある。それは、S が F すべきだという主張をサポートするような適切な論証・推論の前提を提示する必要がある場合だ。このとき質問の発話者は、S が F すべきだとは考えていない。質問の発話者は、S が F するということに異議を唱えているか、異議を唱える立場にあると言える。つまり、ここで求められている回答は、F すべきだということを裏付ける論証であり、理由は論証の前提となるのである (Logins 2022, 170)。まとめると、「なぜ？」質問の発話者が、「F すべし」という前提を持っている場合には、説明が求められ、その前提を持っていない場合には論証が求められている。

そのため、「なぜ F するのか？」という問いが説明を求めている場合には、その回答となる理由は、説明的役割に基づいたものである必要がある。つまりこのときは、Broome のような説明的見解が理由の定義として求められていることになる。他方、「なぜ F するのか？」という問いが論証を求めている場合には、理由は推論的役割に基づいたものである必要がある。このときは、推論的見解が定義として求められていると言える。このように規範理由には 2 つの側面があり、「なぜ F するのか？」という問いが何を求めているのかによって、どちらの側面に着目した理由となるかが決まると Logins は主張する (Logins 2022, ch.5)。つまり、Logins によれば理由とは「なぜ F するのか？」に対する適切な回答に当たるものなのだ。

(3) 筆者の主張

この Logins の理由の問答的見解は、理由の還元的定義の 2 つの立場をうまく扱える立場として概ね妥当であると考えられる。さらに、これは規範理由の 2 つの役割を捉えたものとなっている。

繰り返しになるが、Logins によれば、「なぜ F するのか？」という規範的な問いの答えが理由である。この「なぜ F するのか？」は、回答に説明を必要とする読みと、回答に論証を必要とする読みの 2 つの読みがある。例えば、道で倒れている人がいたので、あなたは道で倒れているその人を助けたとしよう。「なぜ道で倒れている人を助けたのか？」と問われたら、少なくとも 2 通りの仕方で答えることができる。まず、その状況からあなたが推論したことを（それが適合的であるなら）そのまま論証として伝える回答を与えるこ

とができる。「道で倒れている人がいた。倒れている人を助けなければ、その人は辛い思いをするだろう。しかし、誰かが助ければその辛さは緩和されるだろう。今、その人を助けることができるのは自分だけだ。だから助けたのだ。」と。他方で、別の仕方で答えることもできる。「聖書に書かれていたから」「義務だから」「功利原理に従って」といったように。前者は論証を、後者は説明を与えている。このことから、「なぜ？」を分析すれば理由には二面性があると言える。

前述した2つの立場に即して捉えるならば次のようになる。この問答的見解は、理由が「義務的な事実（例えば、SはFすべきである）の説明（の一部）であるという点で説明的見解に同意する」(Logins 2022, 171)。しかし、説明理由的な規範理由のみが規範理由であるとは措定していないという点で説明的見解とは異なる(Logins 2022, 171)。他方、理由は「ある意味で、推論の健全なパターンにおける前提の内容になる」という点で、推論的見解に同意している(Logins 2022, 171)。しかし、行為者の推論に関わる規範理由のみが規範理由だとは考えていない点で、推論的見解とは異なっている(Logins 2022, 171)。よって、この見解は規範理由の2つの役割を掬える。これがLoginsの主張であった。

つまり、Loginsは、冒頭で提示した理由の2つの役割はそもそも別の観念についてのものであり、一方を他方に還元した統合的な理論は不可能であると述べていることになる。すなわち、そもそも推論的役割と説明的役割のどちらか一方によった分析はできないとLoginsは述べていることになる。Loginsの貢献は、規範理由を（推論の前提に用いられる）動機づけ理由的な規範理由と説明理由的な規範理由とに区別したことにある。

だが、Loginsの説明には不十分だと思われるところもある。Loginsは、説明と論証には明らかに共通点があるにもかかわらず、その共通点についてはほとんど述べていない。例えば、「村人が困っていたことは、フリーレンが村人を助けるべき理由だ」という場合、説明を与えているとも論証を与えているとも、どちらともとれるだろう。というのは、「なぜフリーレンは村人を助けるのか？」という質問によって、発話者が「フリーレンは村人を助けるべきではない」と前提していても前提していなくても、いずれにしても「村人が困っているから」という答えはその回答として成り立つからだ。Loginsは事例ベースではこの曖昧性を認めているが、その含意を明らかにしていない。ここで私が問題にしたいことは、このような説明と論証の曖昧性を認めた場合には理由はどちらの働きをするのか不明だということだ。説明として提示される理由が、論証の前提として論証的理由として扱われる場合は大い

に考えられるしその逆も考えられる。このような場合の理由は、説明と論証の両方の役割を果たしているといって良いのだろうか。もしそうって良いのなら、理由の2つの役割はそもそも別の観念についてのものであり、一方を他方に還元した統合的な理論は不可能だと Logins が主張しているのだとすると、その主張と矛盾するようにさえ思われる。加えて、論証を求める問いの回答であっても「功利原理に従って」のように答えることはあり得るだろう。そのため、Logins の主張を擁護するならば、説明と論証のより詳細な区別は必要だろう。

さらに、「なぜ？」質問において、どうして説明を求める読みと論証を求める読みが可能で重要なのかということに Logins は説明を与えていない。Logins は非形式論理で区別される説明と論証の区別を「なぜ？」質問に応用させているが、それがどうして可能で重要なのかは説明が必要だろう。例えば、Dancy が規範理由と動機づけ理由の区別として提示したように、「なぜ？」質問に着目すると、規範的な問いと動機づけの問いを分けることもできる (Dancy 2000, 2–3)。Logins は規範的な「なぜ？」質問に焦点を絞っているが、その応答がどうして適切な規範理由となるのかは不明である。というのも、「なぜ？」質問によって、適切な規範理由と動機づけ理由や説明理由をどのように区別するのかは不明だからである。さらに、説明と論証の区別が Logins の議論にとっては重要であるのだから、Logins が提示する質問の発話者が「すべき」だと前提しているかどうか以上の区別が必要だろう。例えば、「フリーレンが魔族から村人を助けるのは、魔族が憎いからだ」という文を考えよう。このとき「なぜフリーレンは村人を助けるのか」と問われた時に、フリーレンの視点から「魔族が憎いからだ」とフリーレンが答えることは想像がつくし、説明にもなっている。しかし、この理由は、私たちが考えている規範理由とは別の理由、むしろフリーレンの動機づけ理由であると言えるのではないだろうか。だが、Logins の枠組みでは規範理由として扱われることになるだろう。これは、Logins の議論の欠陥というよりむしろ「規範理由」というときの「規範的」という語の意味が多義的であることの問題だ。

加えて、Logins の説明と論証の区別という主張でも捉えられていない規範理由の区別についての論点がある。それは、規範性、特に「べし」における当為と評価の区別を用いた理由の区別だ (cf. 鴻 2021)。「べし」には当為的な読みと、評価的な読みができることは、近年の義務様相についての意味論の議論の中で多く指摘されている (cf. Castañeda 1970; Schroeder 2011; Chrisman 2016; Väyrynen 2019)。大雑把に述べるならば、「べし」には事態に帰属する「べし」と、行為に帰属する「べし」がある。事態に帰属する「べ

し」は評価的な「べし」であり、行為に帰属する「べし」は当為的な「べし」である(鴻 2021, ch.8)。「べし」についての評価と当為の区別を受け入れるならば、理由についても同様の当為と評価の区別を認めることができるだろう。「村人が困っていることは、フリーレンが村人を救うべき理由である」という文を考えてみる。このとき当為的な読みでは、この文はフリーレンが村人を助けることを要請していると読むことができる。他方、評価的な読みでは、この文の話者が、村人が困っているという事実によって、フリーレンが村人を助けるという事態に対する賛成的態度を表出をしていると読むことができる(鴻 2021, 208)。

このような「べし」に関する意味論的区別を受け入れるならば、Loginsの説明による区別だけで、規範理由の働き全てを捉えられるわけではない。Loginsの理論では、たしかに、本稿冒頭で提示した理由の2つの役割にもとづく区別は行われている。しかし、この2つの役割だけではなく規範理由には、事態に対する評価を行う役割と、行為に対する要請を表す役割とがある。この役割はどのようにして区別されるのか Loginsの枠組みでは想定されていない。

以上のことから、Loginsの問答的理由説にもとづく規範理由の区別の説明において不十分、または今後取り組むべき三つの点が示唆された。ひとつ目は、説明と論証の区別の曖昧性である。前述したように説明と論証には共通点があるように思われるし、その区別は曖昧である。ふたつ目は、どうして「なぜ？」質問の二つの読み方が適切な規範理由を特定するのか不明瞭であること。さらに「なぜ？」の問いに答えることが適切な理由の提示であると言えるのかは不明瞭である。みつつ目は、推論的役割と説明的役割の区別のほかにも、規範理由には当為的理由と評価的理由の区別ができそうであり、それらとLoginsの理論との関係の明確化である。

(4) 今後の展望

私は、Loginsの理論に基づいて規範理由を区別することは良い路線であると考えている。そのために前述した三つの点を洗練させる必要がある。その中でも、論証と説明の区別と「べし」に関する区別の二つは重要である。論証と説明の区別は、曖昧で良いと考えており、この区別はプラグマティックに区別されるだろう。この曖昧性によって、規範理由が説明と論証の両方の役割を果たすことを表していると今後論じたい。また、規範理由というときの「規範的」という語の意味の多様さによって議論が交錯していることも考えられるので、その整理も必要である。

さらに、ここまでの議論で不十分な点は規範的な「べし」に関する議論とそれにもとづく当為的理由と評価的理由の区別だ。すでに Castañeda や Schroder、Chrisman が議論しているような、義務様相意味論を用いた「べし」に関する意味論のさらなる精緻化が必要である(cf. Castañeda 1970; Schroeder 2011; Chrisman 2016)。そして、具体的にこの区別を理由にどう応用させるのが今後の課題である。

また、Logins が「なぜ？」質問にもとづく区別を行ったように、動機づけ理由に関してもより詳細な区別を行うことが可能であると考えられる。動機づけ理由に関する心理主義と反心理主義との論争も、本稿で見た説明的見解と推論的見解のように適切な区別を用いれば調停可能な立場であると今後論じたい。

(5) 参考文献

- Broome, John. 2013. *Rationality through Reasoning*. John Wiley & Sons.
- Castañeda, Hector-Neri. 1970. "On the Semantics of the Ought-to-Do." *Synthese* 21 (3-4): 449-68.
- Chrisman, Matthew. 2016. *The Meaning of "Ought": Beyond Descriptivism and Expressivism in Metaethics*. Oxford University Press.
- Dancy, Jonathan. 2000. *Practical Reality*. Oxford, GB: Oxford University Press.
- Logins, Artūrs. 2022. *Normative Reasons: Between Reasoning and Explanation*. Cambridge University Press.
- McHugh, Conor, and Jonathan Way. 2022. *Getting Things Right: Fittingness, Value, and Reasons*. Oxford University Press.
- McKeon, Matthew W. 2013. "On the Rationale for Distinguishing Arguments from Explanations." *Argumentation* 27 (3): 283-303.
- Raz, Joseph. 1999. *Engaging Reason: On the Theory of Value and Action*. Oxford University Press.
- Schroeder, Mark. 2011. "Ought, Agents, and Actions." *Philosophical Review* 120 (1): 1-41.
- Smith, Michael. 1994. *The Moral Problem*. Blackwell. (樞則章訳.2006『道徳の中心問題』ナカニシヤ出版.)
- Väyrynen, Pekka. 2019. "Normative Commitments in Metanormative Theory." In *Methodology and Moral Philosophy*, edited by Jussi Suikkanen and Antti Kauppinen, 193-213. Routledge.

Williams, Bernard. 1979. "Internal and External Reasons." In *Rational Action*, edited by Ross Harrison, 101–13. Cambridge University Press.

鴻浩介. 2021. 「行為の理由と規範性のジレンマ.」 博士論文.

(一橋大学)

様々な定式化による算術の無矛盾性証明と Gentzen の当初の目標
Consistency proofs for first-order arithmetic in various formulations and
Gentzen's original goal

安東 里沙子

Abstract

This paper compares and analyzes the consistency proofs for first-order arithmetic in various formulations, in light of Gentzen's original goal. First, Natural deduction and Sequent calculus systems are compared, and proofs of Normalization theorem and Cut-elimination theorem are outlined. Next, Gentzen's consistency proofs for first-order arithmetic, subsequent works on the consistency proofs for first-order arithmetic (by normalization) after Gentzen, and works on infinite proof systems are introduced. Then, focusing on the compatibility of "naturalness" and "technical advantage" of proof systems, this paper focuses on the consistency proof in Troelstra and Schwichtenberg [16].

1 研究テーマ

本稿では、特に自然演繹に着目した上で、様々な定式化による算術の無矛盾性証明を比較し、関連を探求する。Gentzen により提案された自然演繹体系は、証明体系の自然な定式化が実現されている代わりに、特にその構造の分析に際して、ある種の技術的困難が伴っているといえる。そこで、主にその技術的困難をどう克服するかという観点から、無矛盾性証明で用いられる様々な手法を概観し、算術の無矛盾性証明に関する諸研究の位置付けについて整理する。また、それを通じて、無矛盾性証明に関する Gentzen のオリジナルの着想について検討する。最後に、形式的体系の「自然さ」と、その正確な無矛盾性証明に必要な「技術的利点」は両立するかといった問いについて、両立する体系が可能であることを示唆する。

2 研究の背景・先行研究

証明論の代表的な証明体系として、**自然演繹 (NJ, NK)** と **シークエント計算 (LJ, LK)** がある。自然演繹において、仮定を閉じる操作が与えられ、さらに単一の論理式のみが結論に現れること、シークエント計算において、複数の論理式の集合が結論に現れてもよいことは、両者の体系を比較した際に、各々を特徴づける性質の一部として挙げられるだろう。自然演繹における導入規則 (除去規則) はシークエント計算における R 規則 (L 規則) にそれぞれ対応しており、例えば、条件法の導入規則は以下の左の推論規則であり、条件法の R 規則¹ は以下の右の推論規則である。

$$\frac{\begin{array}{c} [A] \\ \vdots \\ B \end{array}}{A \rightarrow B} \rightarrow I \qquad \frac{A, \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \rightarrow B} \rightarrow R$$

自然演繹は、数学の証明を典型とするような、我々が実際に行なっている推論とできるだけ近い形で推論規則を定式化した体系である (Gentzen [5])。その一方で、自然演繹は、古典論理においてはその定式化に成功しているとはいえないだろう。後で詳しく述べるが Gentzen [5] には、正規化定理は、直観主義的自然演繹 (NJ) では直ちに証明できたが、古典的自然演繹 (NK) では証明できなかったとある。Gentzen は正規化定理の帰結を古典的体系に拡張するために、古典的シーケント計算 (LK) を考案し、それに対応する定理であるカット消去定理を証明した。たしかに、シーケント計算は (特に古典論理における) その技術的利点においては優れた定式化であるといえる。しかし、その定式化において「自然さ」が失われている点は軽視できない²。例えば、Steinberger [15] は、論理的推論主義者 (証明論的意味論の立場) は、複数の結論をもつ論理体系を拒否すべきであると主張する。証明論的意味論の観点からいえば、演繹体系は我々の推論実践と適切に結びついていなければならないが、複数の結論をもつ論理体系は我々の通常の推論様式を自然に表現できていないからである³。このように、自然演繹は、推論の自然な定式化に成功している代わりに技術的な困難を伴い、シーケント計算は、定式化において自然さを失っている代わりに技術的な利点を備えている。すなわち、両者の体系はそれらの体系がもつ自然さと技術的利点において、各々相反する優位性を備えているといえる。

Gentzen の純粹論理に関する研究は、算術と解析学の無矛盾性証明のための予備的なものであった。Gentzen は当初、自然演繹の正規化定理を直観主義算術の体系に拡張することで、算術の無矛盾性を証明することを目指していた (Gentzen [5])。

正規化定理とは、ある前提 Γ から結論 A への証明 d が与えられたとき、 d に簡約手続きを有限回適用することで、 Γ から A への正規形の証明を得ることができる、というものである。この定理の帰結として、自然演繹体系の無矛盾性が示される。正規形の証明とは、遠回りを含まない証明、すなわち、導入規則が適用された直後に除去規則が適用されているような論理式 (極大論理式) を含まない証明のことであり、以下は、簡約手続きの一部である。

$$\frac{\frac{[A] \quad \vdots d_2}{B} \quad \frac{\vdots d_1}{A} \quad \rightarrow I}{A \rightarrow B} \rightarrow E \quad \hookrightarrow \quad \frac{\vdots d_1}{A} \quad \frac{\vdots d_2}{B} \quad \vdots d_3}{A} \rightarrow E$$

長らく、Gentzen は正規化定理に関しては、証明の背後にある直観的な考えを説明するにとどまっておき、完全な証明は Prawitz [10] で与えられたとされていたが、von Plato [17] によれば、2005 年 2 月に発見された Gentzen の博士論文の初期の手書き原稿には、直観主義的自然演繹の正規化に関する詳細な証明が含まれていたのである。それにもかかわらず、Gentzen はカット消去定理を応用することで、算術の無矛盾性証明を得た。**カット消去定理**とは、シーケント計算において、ある前提 Γ から結論 Δ への証明が与えられたとき、有限回の操作により、 Γ から Δ へのカット規則の適用を含まない証明（遠回りを含まない証明）を得ることができる、というものであり、ちょうど自然演繹における正規化定理に相当するものである⁴。

Gentzen が無矛盾性証明のために自然演繹ではなくシーケント計算による定式化を選択したのは、ひとつには、自然演繹における、直観主義的体系に追加されねばならない**二重否定除去則** ($\neg\neg A \rightarrow A$) の特別な立ち位置による。二重否定除去則は、否定の導入則から正当化できないという点から、簡約手続きを与えることができないのである。これはシーケント計算を採用する際には全く問題とならない。もうひとつには、自然演繹ではおおよそ保たれるインフォーマルな証明における文の自然な連続が、シーケント計算では、ある目的のもとつくられた人工的な列に置き換わり、それが無矛盾性証明において技術的に有利であることによる。

最小論理に算術の公理と爆発則 ($\perp \rightarrow A$) を追加することで、**直観主義算術** (**ハイティング算術**、以下 **HA**) の体系が得られ、さらに二重否定除去則を追加することで、**古典算術** (**ペアノ算術**、以下 **PA**) の体系が得られる。論理体系の無矛盾性を算術の理論に拡張する際の問題点は、主に、算術の場合は複雑な論理式（量子子を含む論理式）が論理規則の適用のみならず、**数学的帰納法**の使用によって導出されることもある点である。つまり、数学的帰納法の使用を構成的に正当化することに困難が生じる。

Gentzen は生涯に数回にわたって算術の無矛盾性証明に取り組んでいる。1934 年に発表された最初の無矛盾性証明は、ファン定理を暗黙のうちに使用しているという Bernays の批判により取り下げられたが、この批判は Bernays 自身により撤回され、Gentzen [7] として Gentzen の死後に出版されている。

また、Gentzen [5] は算術の無矛盾性に関するもうひとつの重要な帰結として、PA、特にその無矛盾性の問題が、HA に還元されることを示した。1936 年に発表された第二の無矛盾性証明では、証明図への順序数記法の割り当てが定義され、 ϵ_0 までの超限帰納法の原理の構成的証明に依拠したものとなっている。1938 年に発表された第三の無矛盾性証明は、第二の無矛盾性証明を純粹なシーケント計算で行った改訂版である。Akiyoshi and Takahashi [1] の分析によれば、第一の証明が内容的（意味論的）正しさに関する証明であり、第三の証明が形式的（証明論的）正しさに関する証明である。また、Akiyoshi and Takahashi [1] によれば、第二の証明は内容に関する証明でも形式に関する証明でもあり、二つの側面が互いにどのように関連しているかに関する議論もなされている。第三の無矛盾性証明までとは対照的に、1943 年に発表された第四の無矛盾性証明は、非導出可能性に基づいている。また、Gentzen は 1939 年から 1943 年にかけてさらなる無矛盾性証明に取り組んでいた。1938 年の第三の無矛盾性証明を直観主義的シーケント計算によって再構成し、HA の無矛盾性を証明するというものである。

Gentzen によって与えられた標準的な無矛盾性証明はシーケント計算によるものであり、後続の研究の多くはそれに基づいているが、自然演繹による無矛盾性証明に関する研究も存在する。Kanckos [9] は、自然演繹で形式化された HA の無矛盾性証明を与えた。この証明は、Howard [8] による証明図へのベクトル記法の割り当てに基づく、正規化の手法がベースとなっている。この手法を用いた証明は、Gentzen のオリジナルの証明で必要となる等高線 (Höhenlinie) の概念の使用を避けられている。また、この HA の無矛盾性証明での議論を一般化し、Siders [14] は自然演繹で形式化された PA の無矛盾性証明を与えている。一方で、Prawitz [11] は、Gentzen の 1936 年の無矛盾性証明での手法を自然演繹に応用することで、PA の無矛盾性証明を与えている。

また、無矛盾性証明に関する議論を順序数解析というより数学的なトピックに移行するために、 ω 規則という新たな道具立てが考案された。 ω 規則とは、 $A(x)$ (x は任意の数) に具体的な数項を代入して得られた無限個の論理式 $A(i)$ ($i < \omega$) から、 $\forall x A(x)$ を導出する推論規則であり、以下の形をしている。

$$\frac{\begin{array}{c} \vdots \\ \dots A(i) \dots \end{array}}{\forall x A} \omega$$

ω 規則の使用は、超限的命題の有意味性を実無限の概念に訴えることなく正当化することを目的とする Gentzen による無矛盾性証明の手法とはかけ離れてい

るように見えるが、実は有限的に正当化される。Buchholz [3] によれば、 ω 規則に基づくシーケント計算体系におけるカット消去の手続きから、Gentzen によるカット消去の手続きを復元することができるからである。 ω 規則に基づく体系（無限的体系）による証明の分析は、現代の証明論研究においては標準的である。Buchholz [3] は Gentzen の 1938 年の無矛盾性証明を、無限的シーケント計算体系におけるカット消去法による現代的な形で再提示している。一方で Troelstra and Schwichtenberg [16] は、無限的自然演繹体系に基づく算術の正規化定理を証明しており、無矛盾性証明に相当する帰結が得られている。

3 筆者の主張

ここまで、自然演繹とシーケント計算の体系の比較から始まり、正規化定理とカット消去定理、Gentzen による算術の無矛盾性証明、Gentzen 以降の（正規化による）算術の無矛盾性証明に関する研究、無限的証明体系に基づく研究についてごく簡単に紹介してきた。これらの研究の中でも、特に Prawitz [11] と Troelstra and Schwichtenberg [16] による証明は、正規化定理を算術の体系に拡張することで算術の無矛盾性を証明するという仕事を、Gentzen による 1936 年の無矛盾性証明の手法に基づき達成しているものであるといえよう。Prawitz [11] ではかなり詳細な証明が示されている。Troelstra and Schwichtenberg [16] の証明はカーリーワード対応に基づいて、 λ 記法によって記述されている⁵。

ここで、自然さと技術的利点を両立させるような体系は存在するかという問いが重要である。筆者は、Troelstra and Schwichtenberg [16] の証明で採用された無限的自然演繹体系はその一つに値するのではないかと考える。無限的体系は、無限個の前提を持つ推論規則という、我々の推論の実践からかけ離れているように見える概念を含んでいるという点で、自然さを備えていないと思われるかもしれない。しかし、標準的な証明体系に現れる**固有変項 (eigenvariable)** の解釈も、果たして自然な思考の対象であるだろうか。標準的な証明体系と無限的な証明体系の間に、無限の解釈に関する実質的な違いが存在するようには思えないのである。

固有変項というある種の超越的な概念に関して、それが何であるかについては様々な立場があり、たとえば、道具主義的説明、量化的説明、そして指示的説明などが与えられている (Breckenridge and Magidor [2])。道具主義的説明によれば、固有変項は意味をもたず、意味のある前提から意味のある結論に到達するための便利な道具としての役割にすぎない。量化的説明によれば、固有変項は、量子子によって暗黙のうちに束縛される変数である（あるいは、

それ自体が形を変えた量子子である)。最後に、固有変項とはある種の対象を指示する名前であるという、指示的説明がある。指示的説明の代表的論者の一人である Fine [4] によれば、固有変項とは任意の対象 (arbitrary object) であり、また、Breckenridge and Magidor [2] によれば、なんらかの通常の対象である。ただし、どの特定の対象を指示しているのかは知りえないという。

ω 規則の従来の位置付けは、シュッテ学派により単なる技術的・数学的な工夫のため考案された道具立てである、というものである (Schütte [13])。それを使用することの利点は、証明図に現れる固有変項を消去できることにある。例えば、数学的帰納法の適用を含む証明は、無限的自然演繹体系に以下のよう埋めこまれる⁶。

$$\frac{\begin{array}{c} [A(x)] \\ \vdots \\ A(0) \quad A(S(x)) \end{array}}{\forall x A(x)} \text{Ind} \quad \hookrightarrow \quad \frac{\begin{array}{c} [A(0)] \quad [A(1)] \\ \vdots \\ A(0) \quad A(1) \quad A(2) \quad \dots \end{array}}{\forall x A(x)} \omega$$

このように埋め込まれた証明は固有変項を含んでいないため、簡約手続きを問題なく適用することができる。ここで、固有変項を消去するということは、固有変項の現れを ω 規則の適用に置き換えているということでもあるが、そのような意味で、 ω 規則は固有変項、すなわち任意の対象を、証明論的意味論の立場から意味づけているといえないだろうか。

4 今後の展望

算術が無矛盾であるという帰結自体は、Gentzen によりほとんど完全な形で得られている。しかし、どのような定式化によってそれが与えられるかについては今後も探求していく必要がある。特に、自然演繹で定式化された算術の無矛盾性証明、すなわち算術に対する正規化定理に関する研究は、これまで十分に注目されてこなかった。しかし、自然演繹は、推論規則の適用によって論理式の意味を捉えようとする証明論的意味論の立場において特に重視されており、この点において、自然演繹は証明論研究において重要な意義をもつはずである。さらに、人間の証明活動を自然に定式化することに成功し、証明支援系の研究をはじめとする様々な研究分野での応用可能性をもつ自然演繹について、より詳細な分析が与えられることは不可欠である。

注

¹ 規則に含まれる Γ 及び Δ はそれぞれ 0 個以上の論理式の集合である。

² 自然演繹は 1980 年代以降、証明支援系の研究の基礎にある体系でもあり、人間にとって使いやすい証明支援系を設計する上でも、証明の「自然さ」の探究は重要性をもつ。

³ 反対に、Restall [12] は、主張と否認の概念があれば、論理的帰結を複数の前提と結論をもつ論証として分析してよいと主張する。しかしこの主張は、ある種の意味論的観点からの分析において、複数結論をもつ論理体系の優位性を示したに過ぎないように思われる。

⁴ ただし、正規化の手続きとカット消去の手続きの間に対応関係が成り立つか否かについては議論の余地がある (Zucker [18])。

⁵ 安東 [19] では、この証明を自然演繹に変換し、より詳細な証明を与えた。

⁶ 実際には、埋めこみ後の証明図は無限個の前提をもつ。

文献

- [1] Ryota Akiyoshi and Yuta Takahashi. Contentual and formal aspects of Gentzen's consistency proofs. In *Philosophical Logic: Current Trends in Asia: Proceedings of AWPL-TPLC 2016*, pages 95–137. Springer, 2017.
- [2] Wylie Breckenridge and Ofra Magidor. Arbitrary reference. *Philosophical Studies*, 158:377–400, 2012.
- [3] Wilfried Buchholz. Explaining Gentzen's consistency proof within infinitary proof theory. In *Kurt Gödel Colloquium on Computational Logic and Proof Theory*, pages 4–17. Springer, 1997.
- [4] Kit Fine. A defence of arbitrary objects. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 57:55–89, 1983.
- [5] Gerhard Gentzen. Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1935. English translation in [6].
- [6] Gerhard Gentzen. *The Collected Papers of Gerhard Gentzen*. North-Holland, 1969.

- [7] Gerhard Gentzen. Der erste Widerspruchsfreiheitsbeweis für die klassische Zahlentheorie. *Archive for Mathematical Logic*, 16(3-4):97–118, 1974.
- [8] William Alvin Howard. Assignment of ordinals to terms for primitive recursive functionals of finite type. In *Studies in Logic and the Foundations of Mathematics*, volume 60, pages 443–458. Elsevier, 1970.
- [9] Annika Kanckos. Consistency of Heyting arithmetic in natural deduction. *Mathematical Logic Quarterly*, 56(6):611–624, 2010.
- [10] Dag Prawitz. *Natural Deduction: A Proof-theoretical Study*. Courier Dover Publications, 2006.
- [11] Dag Prawitz. A note on how to extend Gentzen’s second consistency proof to a proof of normalization for first order arithmetic. In Reinhard Kahle and Michael Rathjen, editors, *Gentzen’s Centenary: The Quest for Consistency*, pages 131–176. Springer, 2015.
- [12] Greg Restall. Multiple conclusions. In *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, pages 189–205. Kings College Publications London, 2005.
- [13] Kurt Schütte. Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie. *Mathematische Annalen*, 122:369–389, 1950/51.
- [14] Annika Siders. Normalization proof for Peano arithmetic. *Archive for Mathematical Logic*, 54:921–940, 2015.
- [15] Florian Steinberger. Why conclusions should remain single. *Journal of Philosophical Logic*, 40:333–355, 2011.
- [16] Anne Sjerp Troelstra and Helmut Schwichtenberg. *Basic Proof Theory*. Cambridge University Press, 2000.
- [17] Jan von Plato. Gentzen’s logic. In Dov Gabbay, editor, *The Handbook of the History of Logic*, pages 667–721. Elsevier, 2009.
- [18] Jeffrey Zucker. The correspondence between cut-elimination and normalization. *Annals of mathematical logic*, 7(1):1–112, 1974.
- [19] 安東 里沙子. ω 規則に基づく自然演繹体系による算術の無矛盾性証明. 慶應義塾大学文学研究科修士論文, 2024.

(慶應義塾大学)