広義の自己言及のパラドクスの一般構造と解決方法

山森真衣子

0　はじめに

パラドクスについての研究は古今東西を問わず行われてきた。では、なぜ人々はパラドクスについて研究をしてきたのだろうか。もちろん、パラドクスを研究するのが面白いからでもあるが、それ以上に、パラドクスの存在が、我々の思考や理論に対して重大な問題提起を行っているからである。パラドクスは「正しそうな前提」と「妥当そうな推論」から「受け入れがたそうな結論」[*1]が生じるという困難だ[*2]。一般に我々は、正しい前提と妥当な推論からは、必ず正しい結論が導出されると考える。前提が正しく推論が妥当であるにもかかわらず間違った結論が導出されるとすると、判明している事実から新しい事実を導出することが難しくなる[*3]。そのため、パラドクスの存在は、そのパラドクスが生じる理論やその理論を受け入れている人の思考において、「判明している事実から新しい事実を導出する」という知的営みが困難である可能性を示唆する。それゆえ、パラドクスを孕む理論においても上述の知的営みを望むのであれば、そのパラドクスに対する探求を避けて通ることは許されないのだ（もちろん、探求の結果、その思想や理論自体が合理的でないという結論が導かれる可能性も多大にあるが）。

さて、本稿は、種々のパラドクスのうち、「自己言及のパラドクス（自己言及性のパラドクス、self-referential paradox, paradox of self-reference）」と呼ばれるものを扱う。自己言及のパラドクスは、「自分自身に言及することで生じるパラドクス」と特徴付けられることが多い。だが、第 1 節で述べるように、自己言及のパラドクスとみなされている（みな

---

*1 この「受け入れがたい結論」として矛盾のみを念頭に置いている説明もある（cf. [17, p.1987]、[7, p.21]、[14, p.15]）が、「受け入れがたい結論」は矛盾のみではない。矛盾以外の「受け入れがたい結論」には、「自明化」（第 1 章で紹介する）や「否定的自明化」（第 3 章で紹介する）がある。

*2 なお、広い意味、ないし元々の意味では、パラドクスとは「一般に受け入れられている見解に反する命題」（cf. [17, p.1987]）のことである。だが、本稿では、このような広い意味におけるパラドクスではなく、狭い意味——論理学で用いられる意味——におけるパラドクス、つまり「正しそうな前提と妥当そうな推論から受け入れがたい結論が生じる」という困難を扱う。本稿で断りなく「パラドクス」という語を用いる際は、この狭い意味でのパラドクスを指す。

*3 例えば、正しい前提と妥当な推論からは必ず正しい結論が導出される場合、A「今日が月曜日ならば、明日は火曜日である」と B「今日は月曜日である」が正しい場合、この二つから導出される結論である C「明日は火曜日である」は必ず正しい。A と B という判明している事実から、C という新しい事実を導出することができるのだ。だが、正しい前提と妥当な推論から間違った結論が導出される可能性がある場合、C という導出された結論が正しいのか否かわからず、C が事実かどうかわからなくなるのである。

されることもある）パラドクスの中には、「自分自身『のみ』に言及する」わけではないものもある。本稿では、このようなパラドクスに対する我々の直観を救うために、自己言及を二種類に分ける。つまり、自分自身のみに言及する「狭義の自己言及」（既存の「自己言及」）と、自分自身を含むグループに言及する「広義の自己言及」があると考える。また、それに応じて「狭義の自己言及のパラドクス」と「広義の自己言及のパラドクス」があると考える。

本稿の目的を述べる前に、筆者の研究の目的について軽く触れておきたい。それは、狭義・広義両方の自己言及のパラドクスに共通する、一般的な解決方法を探ることだ。もちろん、「自己言及のパラドクスの『一般的な』解決方法を探る」という研究に対して懐疑的な人も少なくないだろう。一般に、「共通の解決方法」を探るためには、「共通のパラドクス発生条件」がなければならない。しかし、広義であれ狭義であれ、自己言及のパラドクスに「共通の発生条件」というものは存在するのか、という疑問が湧くのは自然なことだ。広く知られるように、Ramsey[10, p.171–172] は自己言及のパラドクスを二つのグループに――論理的、数学的語や概念そのものから生じるパラドクスと、経験的な語や概念とも関わるパラドクスに――分類した。つまり、Ramsey や彼の分類を受け入れる人にとって、自己言及のパラドクスは「論理的語のみを原因とするもの」「論理的語以外も原因とするもの」に区別されるため、「自己言及のパラドクスに共通の発生条件がある」という考えは受け入れ難いかもしれない。だが、Priest[8, p.142] が指摘するように、Ramsey の区別はあくまで表層的なものにすぎない。Ramsey は「どの概念、どの語がパラドクスに現れているか」に着目し、それによって自己言及のパラドクスを分類しようとする。だが、「自己言及のパラドクスに共通の発生条件がある」と主張するために要請される前提は、「自己言及のパラドクスは同じ（ないし同種の）概念や語を用いている」ということではない。「自己言及のパラドクスは同じ構造を持ち、かつ、その構造がパラドクスの発生に本質的である」ということだ*4。そしてその「自己言及のパラドクスに共通の本質的な構造」は、第 2 節で見るように、これまで事実発見されてきている。そのため、「自己言及のパラドクスの一般的な解決方法の探求」には全く見込みがないとするのは早計であろう。

とはいえ、その「これまで発見されてきた、自己言及のパラドクスに共通の構造」はいずれも、狭義の自己言及のパラドクスの構造でしかなかった。一般に、狭義の自己言及のパラドクスについてはこれまでも豊富な議論があるものの、広義の自己言及のパラドク

*4 さらには、Ramsey の「論理的・数学的」と「非論理的・非数学的」という二つの概念の境界線は、時代に応じて変わる流動的なものでしかないとも Priest[8] は指摘する。

スに関しては十分な議論があるとは必ずしも言えない状況である。そのため、本稿では、「狭義・広義両方の自己言及のパラドクスに共通の一般的構造、および共通の一般的解決方法」を考察するのではなく、そのために必要な準備作業としての「広義の自己言及のパラドクスの共通構造の提示」を行う。また、その構造を分析することで広義の自己言及のパラドクスの解決方法の候補がわかるのだが、その解決方法の一つを扱い、「パラドクスのある解決方法を採用するための哲学的正当化はいかにして可能なのか」「その解決方法に問題があるかどうかはどのようにチェックされるのか」の一つの例を提示する[*5]。

本稿は以下のように構成される。まず第 1 節で、議論の本筋に入る前に、いくつかの自己言及のパラドクスについて確認する。第 2 節では、自己言及のパラドクスの一般的な解決方法を探る先行研究を紹介する。これらの研究たちは、「自己言及のパラドクスの発生に本質的な、自己言及のパラドクスの共通構造」を提示し、そこから「一般的な解決方法」を分析している。第 3 節では、まず、第 2 節で確認した先行研究は、「狭義の自己言及のパラドクス」に対するものであったため、広義の自己言及のパラドクスをカバーしきれていないことを確認する。そして、広義の自己言及のパラドクスの一般構造を提示し、その構造を分析して、どのような「一般的な解決方法」が可能か、またその解決方法に問題はないかを考察する。

1　自己言及のパラドクスの例たち

ではまず、いくつかの自己言及のパラドクスについて確認しよう。

■嘘つきのパラドクス　まずは、最も有名な自己言及のパラドクスである「嘘つきのパラドクス（the Liar Paradox）」について見ていこう。これは、嘘つき文 φ「この文は偽である」から生じるパラドクスだ。

このパラドクスについて説明をする前に、「真である（**Tr**）」という述語に関して簡単に説明しておこう。この述語がどのように振る舞うか（我々がこの述語をどのように使っているか）に着目し形式化した代表的な人物として、Tarski が挙げられる。彼は「真である」という述語の振る舞いについて、以下のように形式化した。任意の文（論理式）P に対して、「文 P が真であるのは、P であるとき、またその時に限る」、すなわち

$$\mathbf{Tr}(\ulcorner P \urcorner) \equiv P$$

[*5] 本稿では広義の自己言及のパラドクスの解決方法の候補の「一つ」を扱っているが、筆者は [16, 第 4 章、付録] において、複数の解決方法について考察をしている。

である（ただし $\ulcorner P \urcorner$ は文 P の名前である[*6]）。これは T-図式（T-Schema）と呼ばれる（以下、証明図で T-図式が用いられている際、「TS」と記載する）。嘘つきのパラドクス（および後で見るカリーのパラドクス）には、この T-図式が用いられている。

では、嘘つき文 φ「この文は偽である」からどのようにパラドクスが生じるのだろうか。一般に文は真か偽かのどちらかである。もし φ が真なら、「この文は偽である」が真であるのだから、φ は偽であるということになる。よって、φ が真と仮定すると、「φ は真であり、かつ、偽である」という矛盾した結論が導かれる。一方で、φ が偽の場合、「この文は偽である」が偽なので、φ は真ということになる。そのため、φ が偽と仮定しても、「φ は真であり、かつ、偽である」という矛盾が導かれる。

嘘つきのパラドクスを形式的に確認しよう。嘘つき文は、形式的に、

$$\varphi := \neg\mathbf{Tr}\ulcorner\varphi\urcorner$$

と表すことができる。ここから矛盾は以下のように導出される[*7*8]。

$$\dfrac{\varphi \vee \neg\varphi \qquad \dfrac{\dfrac{[\varphi]_{\vee_1}}{\mathbf{Tr}\ulcorner\varphi\urcorner}\,\mathrm{TS} \qquad \dfrac{[\varphi]_{\vee_1}}{\neg\mathbf{Tr}\ulcorner\varphi\urcorner}\,\mathrm{def.}}{\mathbf{Tr}\ulcorner\varphi\urcorner \wedge \neg\mathbf{Tr}\ulcorner\varphi\urcorner}\,\wedge I \qquad \dfrac{\dfrac{\dfrac{[\neg\varphi]_{\vee_2}}{\neg\neg\mathbf{Tr}\ulcorner\varphi\urcorner}\,\mathrm{def.}}{\mathbf{Tr}\ulcorner\varphi\urcorner}\,\mathrm{DNE} \qquad \dfrac{[\neg\varphi]_{\vee_2}}{\neg\mathbf{Tr}\ulcorner\varphi\urcorner}\,\mathrm{TS}}{\mathbf{Tr}\ulcorner\varphi\urcorner \wedge \neg\mathbf{Tr}\ulcorner\varphi\urcorner}\,\wedge I}{\mathbf{Tr}\ulcorner\varphi\urcorner \wedge \neg\mathbf{Tr}\ulcorner\varphi\urcorner}\,\vee E$$

■ラッセルのパラドクス　次に紹介する「ラッセルのパラドクス（Russell's paradox）」は、素朴集合論（naive set theory）におけるパラドクスである。

「自分自身を要素として含まない集合」の集合、ラッセル集合 $\omega := \{x : x \notin x\}$ を考えよう。パラドクスは、「ラッセル集合はラッセル集合に含まれるか？」という問題を考えた時に発生する。ラッセル集合がラッセル集合に含まれる（$\omega \in \omega$）と仮定すると、ラッセ

[*6] 本稿は真理述語について詳述することを目的としていないため、詳しくは述べないが、文 P と文の名前 $\ulcorner P \urcorner$ とに一対一対応をつけることが可能であるため、文から文の名前への変換をする際に、「文 A を変換したら文 B の名前になってしまった」という事態は生じない。（もちろん、文の名前から文への変換の際も同様である。）

[*7] なお、

$$\frac{A}{B}\ r$$

は、「論理規則 r を用いることで、A から B が導出される」を意味する。

[*8] 証明中の DNE は二重否定除去則（double negation elimination）

$$\frac{\neg\neg A}{A}$$

を意味する。

ル集合は「自分自身を要素として含まない集合」の要素であるので、ラッセル集合は自分自身を要素として含まない（$\omega \notin \omega$）ということになる。これは矛盾である。一方で、ラッセル集合がラッセル集合に含まれない（$\omega \notin \omega$）と仮定すると、ラッセル集合は「自分自身を要素として含まない集合」というラッセル集合の要素としての条件を満たすため、ラッセル集合はラッセル集合の要素ということになる（$\omega \in \omega$）。これも矛盾だ。つまり、ラッセル集合がラッセル集合に含まれると仮定しても含まれないと仮定しても、矛盾が生じるのである。

■カリーのパラドクス　嘘つきのパラドクスとラッセルのパラドクスは、その「受け入れがたい結論」が矛盾であるパラドクスであった。次に確認するパラドクス、「カリーのパラドクス（Curry’s Paradox）」では、それらとは異なり、「自明化（triviality）」と呼ばれる「任意の文が導出される」という結論が生じる。自明化が生じる、つまり任意の文が導出されるということは、（正しい文だけでなく）間違った文も導出されるということであるため、この結論は明らかに受け入れがたいものである[*9]。

カリーのパラドクスを生じさせる文の一例は、文 φ「この文が真なら、月はチーズでできている」である。φ が真であると仮定しよう。φ が真であるということは φ が成り立つということであり[*10]、それはつまり「φ が真なら、月はチーズでできている」が成り立つということである。いま、「φ は真である」かつ「φ が真なら、月はチーズでできている」なので、モーダスポネンスにより、「月はチーズでできている」が導かれる。さて一般に、P を仮定して Q が導かれたとき、「P ならば Q（$P \rightarrow Q$）」が成り立つと言える。いま「φ が真である」を仮定して「月がチーズでできている」が導かれたため、「φ が真であるならば、月がチーズでできている」が成り立つと言える。つまり、「φ が真であるならば、月がチーズでできている」が真であるということだ[*11]。これはまさに、φ が真であるということに他ならない。「φ が真なら、月はチーズでできている」かつ「φ は真である」なので、モーダスポネンスにより、「月はチーズでできている」が導かれる。

ここでは「この文が真なら、月はチーズでできている」から「月はチーズでできている」を導出したが、この「月はチーズでできている」の部分を別の文に変えれば、その別の文が——どんな文であっても——導出されることは明らかであろう。それゆえこのパラドクスは自明化を引き起こすと言えるのだ。

[*9] ちなみに、自明化を受け入れる立場も存在する。その立場、自明主義（trivialism）は、あらゆる文は真であると主張する。この立場を支持する者もいるが（cf.[3]）、批判も非常に多い立場である（cf. [1]、[9, sec.3]）。筆者もこの立場には懐疑的なので、本稿では自明化を「受け入れがたいもの」として扱う。

[*10] 「T-図式を受け入れる場合」という留保がつくが。

[*11] こちらも、「T-図式を受け入れる場合」という留保がつくが。

カリーのパラドクスを形式的に確認しよう。カリー文 φ は、

$$\varphi := \mathbf{Tr}\ulcorner\varphi\urcorner \to \psi$$

と定義される（ψ は任意である）。このとき、ψ は——-つまり任意の文は——以下のようにして導出される[*12]。

$$\dfrac{\dfrac{\dfrac{\dfrac{\varphi \to \varphi}{\varphi \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\,\text{def.}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\,\text{TS}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi}\,\text{Contraction} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\varphi \to \varphi}{\varphi \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\,\text{def.}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\,\text{TS}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi}\,\text{Contraction}}{\mathbf{Tr}\ulcorner\varphi\urcorner}\,\text{def.}}{\psi}\,\to E$$

ところで、カリー文における ψ にはどの式を入れてもよいのだが、ここに $\bot$ を入れると

$$\varphi := \mathbf{Tr}\ulcorner\varphi\urcorner \to \bot$$

となる。これは嘘つき文に他ならない。そのため、カリーのパラドクスは嘘つきのパラドクスを一般化したものであり、嘘つきのパラドクスはカリーのパラドクスの特殊例であると言うことができる。

■語り得ぬもののパラドクス　語り得ぬもののパラドクス（Ineffability Paradox, paradox of ineffability）は、嘘つきのパラドクスやラッセルのパラドクス、カリーのパラドクスとは異なり、これまで宗教哲学での研究が主であり、論理的観点からの研究があまりなされてこなかったパラドクスである。「語り得ぬもの」という概念の特徴付けについては論者によって様々であるが、最も制限のない一般的な特徴付けをするならば、語り得ぬということは、「それについて描写したり表現したりすることが不可能である」ということだと言えるだろう。

多くの哲学者や宗教家が「x は語り得ぬものである」と言ってきた。例えば「神は語り得ぬものである」（神秘主義的キリスト教）、「道は語り得ぬものである」（道教、特に老子）、「クオリアは語り得ぬものである」（心の哲学）、等々である。だが、広く知られるように、この概念はパラドクスを引き起こす。「x は語り得ぬものである」と言われるとき、

[*12] 証明図に現れる「Contraction」は縮約規則（contraction rule）を意味する。この規則は、

$$\alpha \to (\alpha \to \beta) \vdash \alpha \to \beta$$

という規則である。「前提 α を『二回』使用して結論 β を導出できる場合、前提 α の使用が『一回』であっても結論 β を導出できる」と読める。

x はまさに「語り得ぬものである」という言葉によって描写され、表現されている。これはすなわち x は語り得ぬものではないと言うことだ。そのため、「x は語り得ぬものであり、かつ、語り得ぬものではない」という矛盾した結論が生じる。

ところで、このパラドクスはパラドクスではないと考える人もいるだろう。語り得ぬもののパラドクスは、これまで挙げた他のパラドクスと異なり、「x は語り得ぬものである」を否定さえすれば矛盾は生じない。そのため、「x は語り得ぬものである」から矛盾が生じることは、x が語り得ぬものでないことの論証であり、パラドクスではないと考えることもできる。だが、語り得ぬもののパラドクスは「x は語り得ぬものである」からだけでなく「語り得ぬものが存在する」からも生じる。語り得ぬもののパラドクスを「前提が偽であることの論証」と捉えるならば、「語り得ぬものが存在する」が偽であるということになるが、「語り得ぬものが存在する」が偽であるということ――あらゆるものは語りうるということ――は受け入れがたいだろう[*13]。

また、語り得ぬもののパラドクスがパラドクスであることは受け入れても、「自己言及の」パラドクスとみなすことに異論のある人もいるだろう。嘘つきのパラドクスの場合、嘘つき文「この文は偽である」における「この文」は嘘つき文自体を指示していた。カリーのパラドクスの場合も、カリー文「この文が真ならば、月はチーズでできている」における「この文」はカリー文自体を指示していた。語り得ぬもののパラドクスの場合、「この文」といったものは登場しないため、自己言及性と関わらないパラドクスではないか、という疑念はもっともである。

とはいえ、語り得ぬもののパラドクスが「自己言及のパラドクスらしく見える」ことに異論のある者は少ないだろう。語り得ぬもののパラドクスにおいて、「x は語り得ぬものである」という文は「x は語り得ぬものである、わけではない（『語り得ぬものである』によって x が語られているから）」を導く。「x は語り得ぬものである」を P とすると、P が $\neg P$ を導く。あるいはより強い言い方をするならば、P は $\neg P$ を意味するとも言える。これは自己言及のパラドクスらしく見える。

このような「X らしく見えるが、一般的な（ないし既存の）基準からは X ではない」という事象に対して、一般に、二つの態度がとれる。第一の態度は、一般的な基準にのっとって「X ではない」と判断することだ。第二の態度は、「X らしく見える」という直感を救えるような新たな基準を作り出すことだ。この二つの態度は、一方のみが良い態度でもう一方は悪いというわけではないだろう。前者は、既存の X についての議論を厳密に

[*13] 少なくとも、「語り得ぬものが存在する」ことをサポートする議論はいくつも存在する（cf.[4]）が、それを否定する議論はほとんど行われていない。

進めていくためには重要である。後者は、X（ないし X と関連するもの）についての議論に新しい視座を与えてくれるため、これまでとは全く別の地点の発見が期待される。目的の違いが態度の違いなのである。

語り得ぬもののののパラドクスの場合、第一の態度は、このパラドクスに対して、「これには（既存の基準における）自己言及性がないから、『自己言及のパラドクスらしく見える』だけであって自己言及のパラドクスではない」と断じるということだ。そして第二の態度は、「既存の基準では自己言及のパラドクスではないが、『自己言及のパラドクスらしく見える』ため、これを『自己言及のパラドクス』と捉えられるような新たな意味の『自己言及』を考えよう」というものだ。本稿は後者の態度を採用し、「新たな意味の『自己言及』」を提案する。

その、我々の直観を救ってくれる「新たな意味の『自己言及』」は、冒頭でも述べたように、「自分自身を含むグループに対する言及」である。この新たな意味の自己言及を「広義の自己言及」、既存の基準の自己言及（自分自身のみに対する言及）を「狭義の自己言及」と呼ぼう。語り得ぬもののパラドクスには、この「広義の自己言及」が確認できる。上述のように、語り得ぬということは、「それについて表現することが不可能」ということである。すなわち、「x は語り得ぬものである」とは「x は、それについて表現することが不可能なものである」すなわち「あらゆる表現は、x に帰属させられない」ということである。この「あらゆる表現」には、「語り得ぬもの」という表現も含まれる。つまり「あらゆる表現」は「語り得ぬもの」という表現を指示しうる。ここに語り得ぬもののパラドクスの自己言及がある。そして「あらゆる表現」は「語り得ぬもの」だけでなく、「赤い」「善である」「全知である」等の他の表現も指示しうるため、この自己言及は「自分自身のみに対する自己言及」ではなく、「自分自身を含むグループに対する自己言及」なのだ。

長くなったが、語り得ぬもののパラドクスに関しても形式的に矛盾の発生を確認しよう。そのためにまず、「語り得ぬもの」という概念を形式的に定義してみよう。「x は語り得ぬものである（$\neg\mathbf{Ef}(x)$）」とは、「x は、それについて表現することが不可能なものである」すなわち「あらゆる表現は、x に帰属させられない」ということであった。そして「x は語り得ぬものではない（語りうるものである）（$\mathbf{Ef}(x)$）」とは、「ある表現を x に帰属させることができる」である。$\mathbf{Ef}(x)$ および $\neg\mathbf{Ef}(x)$ の形式的定義は様々なものが考えられるだろうが、本稿では最もシンプルと思われるもの、すなわち

$$\mathbf{Ef}(x) := \exists P(P(x))$$
$$\neg\mathbf{Ef}(x) := \forall P(\neg P(x))$$

という定義を採用しよう[*14]。「x は語りうるものである（$\mathbf{Ef}(x)$）」は「ある表現 P について、x にそれを帰属させたものが成り立つ（『$P(x)$』が成り立つような表現 P が存在する）」と、「x は語り得ぬものである（$\neg\mathbf{Ef}(x)$）」は「いかなる表現 P についても、それを x に帰属させたものは成り立たない」と読める。

このように定義すると、「x は語り得ぬものである」という言明が矛盾を生じさせることは、以下のように確認できる。

$$\cfrac{\cfrac{\cfrac{\cfrac{\neg\mathbf{Ef}(x)}{\forall P(\neg P(x))}\ \text{def.}}{\neg\neg\mathbf{Ef}(x)}\ \forall E}{\mathbf{Ef}(x)}\ DNE \qquad \neg\mathbf{Ef}(x)}{\mathbf{Ef}(x)\wedge\neg\mathbf{Ef}(x)}\ \wedge I$$

なお、紙幅の都合上、本稿では広義の自己言及のパラドクスとして語り得ぬもののパラドクスのみを記しているが、[16] ではあと二つの広義の自己言及のパラドクスについても述べている（ベリーのパラドクス、無性質のパラドクス）。

2　先行研究：（狭義の）自己言及のパラドクスの一般構造

冒頭でも述べたように、本稿は広義の自己言及のパラドクスの共通の解決方法を探るが、ただ闇雲に探すのではなく、まずは広義の自己言及のパラドクスに共通の（そしてそのパラドクスの発生に本質的な）構造を探り、それを分析することで共通の解決方法を見つける、という方法をとる。

この手法で自己言及のパラドクスに対する共通の解決方法を考察した研究として、Russell[12]、Priest[8]、Pleitz[6] が挙げられよう（彼らにおける「自己言及のパラドクス」は、当然、「狭義の自己言及のパラドクス」であったが）。Russell の研究は、（狭義の）自己言及のパラドクスに対して一般構造を考察するという手法の初めてのものであったと考えられる。この構造（以下、Priest[8] に倣って「Russell's Schema」と呼ぶ）を元に、Russell は自己言及のパラドクス一般に対する解決方法を提案した。Priest と Pleitz の研究は、それを引き継いだものだ。まず Priest が Russell の提出した一般構造をさらに一般化させた構造「Inclosure Schema」を考案した。そして、Inclosure Schema に適合するパラドクスは全て、哲学的には「真矛盾主義的（dialetheism）」を採用することで、また論理学的には

[*14] 上述したように、語り得ぬもののパラドクスについて形式的に考察した研究はあまり多くない。そのため、議論をわかりやすくするという点からも、叩き台にしたときに応用しやすいという点からも、シンプルな形式化から始めるのがよいと筆者は考える。

「矛盾許容型論理（paraconsistent logic）」を採用することで解決されるという診断を下した（「真矛盾主義」「矛盾許容型論理」に関しては後述する）。Pleitz も狭義の自己言及のパラドクスの一般構造を提出したのだが、彼の構造「Curry’s Schema」は Inclosure Schema をさらに一般化させたものである。この構造から導き出された Pleitz の診断は、自己言及のパラドクスの共通の解決方法は「縮約規則なしの論理（contraction-free logic）」だ、というものである（「縮約規則なしの論理」に関しても後述する）*15。

それぞれの研究について概観する前に、包括的アプローチに関して一点注意することがある。狭義の自己言及のパラドクスに対して包括的アプローチを採用する、すなわち「共通の解決方法を探る」という方法を採るという意味では、ここで挙げた三人以外にも論者は存在する。しかし、上述したように、共通の解決方法を見つけるためには、共通構造をまず見つけることが肝要であると筆者は考える。この「共通構造の構築」をしている論者は、筆者の知る限り、Russell、Priest、Pleitz の三人である。

なお、本節の内容については [13] および [16, 第二章] で詳しく述べているため、簡単に説明するに止める。

2.1 Russell: Russell’s Schema

Russell[12] はカントールのパラドクス、ラッセルのパラドクス、ブラリ=フォルティのパラドクスという狭義の自己言及のパラドクスについて扱い、これらのパラドクスの本質、つまりこれらのパラドクスの共通点を [12, p.35] に示した。この共通点を Priest[8, p.129] が形式的に表したものを、より形式的に表すと、

性質 φ と函数 δ が与えられたとき、

1. $\exists\Omega(\Omega = \{y : \varphi(y)\})$
2. $\forall x(x \subseteq \Omega \rightarrow \delta(x) \notin x)$
3. $\forall x(x \subseteq \Omega \rightarrow \delta(x) \in \Omega)$

と表される。以降、これを「Russell’s Schema」と呼ぶ。Russell’s Schema の三条件は、それぞれ以下のように説明できる。第一条件は、ある性質 φ を満たす対象の全体（総体、totality）Ω の存在を宣言するものである。第二条件と第三条件に現れる函数 δ は、簡単に

*15 Pleitz[6, p.12] がもともと提案した「（狭義の）自己言及のパラドクスの共通の解決方法」は、「縮約規則なしの論理（contraction-free logic）」ではなく「縮約規則なしの矛盾許容型論理（contraction-free paraconsistent logic）」であった。しかし、後述するように、矛盾許容型である必要はない。筆者が 2017 年に Pleitz 自身とこの点について議論をした際、縮約規則の制限で十分だと彼も同意したため、「縮約規則なしの矛盾許容型論理」ではなく「縮約規則なしの論理」を彼の診断として記述する。

言うと、「集合 x を入力すると、x に含まれないある要素を出力する」というものが念頭に置かれている。このような δ を考えたとき、第二条件が成り立つことは自然である。$\delta(x)$ とは「x に含まれないような要素」である——$\delta(x) \notin x$ である——からだ。Ω のどの部分集合 x に関しても、δ が上のような函数である限り、$\delta(x) \notin x$ となることは首肯できるだろう。そして第三条件は、「そのような『x に含まれないような要素』($\delta(x)$) であっても、対象の全体 Ω には含まれる」というものだ。

さて、この三つの条件が満たされるときに矛盾が導出されることを確認しよう。条件 2 と 3 より、

$$\Omega \subseteq \Omega \to \delta(\Omega) \notin \Omega$$

と

$$\Omega \subseteq \Omega \to \delta(\Omega) \in \Omega$$

が成り立つ。Ω が Ω 自身の部分集合であることは明らかなので、$\Omega \subseteq \Omega$ が成り立つ。よって、モーダスポネンスより、

$$\delta(\Omega) \notin \Omega$$

と

$$\delta(\Omega) \in \Omega$$

が成り立つ。そのため、

$$\delta(\Omega) \notin \Omega \wedge \delta(\Omega) \in \Omega$$

という矛盾が導出される。

Russell は、このようなパラドクスに対する解決方法として、タイプ理論を提案した*16。紙面の都合上、タイプ理論については詳しく説明しないが、様々な論者から指摘されるように、タイプ理論は我々の日常の言語使用と乖離してしまっているという欠点があることを述べておこう。そのため、数学におけるパラドクスを扱う分にはよい解決方法となるか

*16 正確には、[12] において Russell が紹介したのはタイプ理論ではなく、ジグザグ理論、サイズの制限の理論、無クラス理論という三つの理論であった。とはいえ、『数学の原理（The Principle of Mathematics）』第二版で追加された序論において Russell は、「[パラドクスに対する] 可能な解決方法に関して言われたものは、タイプ理論について書かれている付録 B を除いて、無視して構わない」[11, p.xiii] と述べているため、本稿では「Russell's Schema によってカバーされるパラドクスの、Russell による解決方法」を（分岐）タイプ理論とみなす。（とはいえ、分岐タイプ理論は Russell's Schema によってカバーされるパラドクスだけでなく、それによってカバーされないパラドクス（ベリーのパラドクスなど）の解決方法としても考案されているのではあるが。）

ちなみに、[12] で挙げられている解決方法（特に無クラス理論）と分岐タイプ理論との間の関係については、[18] に詳しい。

もしれないが、語り得ぬもののパラドクスをはじめとする「そうではないパラドクス」の解決方法としてはあまり妙手とは言えない*17。

2.2 Priest: Inclosure Schema

Russell's Schema は、彼の挙げた三つのパラドクスをカバーしている。しかしながら、Russell's Schema がカバーしきれていない狭義の自己言及のパラドクスがあることも事実である*18。Priest はそのようなパラドクスをカバーするために、Russell's Schema をベースに、それをより一般化させた構造を考案した。この構造、「Inclosure Schema」*19も三つの条件を持っている。Priest の記述 [8, p.134] をより形式的に表すと、

性質 φ、ψ と函数 δ が与えられたとき、

1. $\exists\Omega(\Omega = \{y : \varphi(y)\} \land \psi(\Omega))$
2. $\forall x(x \subseteq \Omega \land \psi(x) \to \delta(x) \notin x)$
3. $\forall x(x \subseteq \Omega \land \psi(x) \to \delta(x) \in \Omega)$

と書ける。Inclosure Schema の三条件は Russell's Schema と似ているが、「ψ」という新たな性質が付け加えられている。そのため、上述の三条件は、それぞれ以下のようになる。第一条件は、ある性質 φ を満たす対象の全体 Ω の存在を宣言するだけでなく、その Ω が ψ という性質を満たすことも述べる。第二条件、第三条件に現れる函数 δ は、Russell's Schema と同じく、「集合 x を入力すると、x に含まれないある要素を出力する」ようなものだ。それゆえ第二条件は、「Ω のどの部分集合も、それが性質 ψ を満たすなら、$\delta(x)$ を要素として持つことはない（$\delta(x) \notin x$ である）」と読める。2.1 で確認したように、函数 δ の性質上、これが成り立つのは自然である。第三条件の読み方は Russell's Schema と同じく「そのような、x に含まれないような $\delta(x)$ であっても、対象の全体 Ω には含まれる」となる。ここで、ψ として当然成り立つような性質を選んだとき、ψ に関する記述は空虚なものとなり、その Inclosure Schema は Russell's Schema と同一視できる。そのため Russell's Schema は Inclosure Schema の特殊例と考えられる。

Inclosure Schema の三つの条件が満たされると、Russell's Schema の場合と同じく、矛盾した結論 $\delta(\Omega) \notin \Omega \land \delta(\Omega) \in \Omega$ が導出される（詳しい証明は省くが、この矛盾が導出さ

*17 また、本稿では割愛するが、[16, p.86-88] で述べたように、タイプ理論は「我々の日常の言語使用と乖離している」だけでなく、原理的に広義の自己言及のパラドクスを解決できない。

*18 [8, p.133-135] で Priest がそれを示しているため、参照されたい。

*19 「inclosure」は「enclosure（囲い込み）」と「in-closure（close の否定）」の掛け言葉となっている。

れることは明らかであろう)。Priest は Inclosure Schema の三条件を満たす（そしてそれゆえ矛盾を導く）パラドクスを、「inclosure paradox」と呼ぶ。

さて、inclosure paradox に対して、Priest が提案した解決方法は、前述のように、「真矛盾主義的」解決方法、論理学的には「矛盾許容型論理の採用」だ。真矛盾主義は、「真なる矛盾が存在する」と主張する哲学的立場である。すなわち、真矛盾主義者は、ある文 S について、S と S の否定（¬S）の両方が同時に真となると主張する。すなわち、Priest の提案する「真矛盾主義的解決方法」は、「矛盾した結論（$\delta(\Omega) \notin \Omega \wedge \delta(\Omega) \in \Omega$）を受け入れよう」ということだ。そして、ほとんどの真矛盾主義者が採用するのが矛盾許容型論理、文字通り「矛盾を許容する」論理体系だ。矛盾許容型論理が矛盾を許容するのは、この論理体系が爆発律（the Law of Explosion）

$$\alpha \wedge \neg\alpha \vdash \beta$$

を持たないからだ。この法則は、「矛盾があれば、そこから任意の文が導出される（矛盾があれば、自明化が起こる)」と言い換えられる。自明化はほとんどの場合で避けるべきものとして扱われる[*20]。爆発律が成り立つ論理体系においては――古典論理や直観主義論理といった代表的な論理体系のほとんどがそれに該当する――矛盾が一つでもあれば自明化が生じてしまう。それゆえ、そのような論理体系においては、矛盾も避けるべきものとして扱われるし、矛盾した結論は受け入れがたいものとなる。一方で矛盾許容型論理を採用した場合、矛盾即自明化とはならないため、矛盾は「矛盾である」というだけで避けられるべきものというわけではなくなる。それゆえ矛盾した結論も、「矛盾している」というだけで受け入れがたい結論となるわけではなくなる。そのため、inclosure paradox という、その結論が矛盾であるパラドクスたちは、矛盾許容型論理を採用すると、その結論が「受け入れがたい」ものではなくなるため、解消されるのである。

2.3 Pleitz: Curry Schema

本章で最後に紹介するのは Pletiz の研究である。だが彼の構造を紹介する前に、Inclosure Schema ではカバーできないパラドクスについて述べたい。

Priest が診断したように、真矛盾主義的解決方法（矛盾許容型論理の採用）は、多くの自己言及のパラドクスに対して有効であるように思われる。パラドクスが引き起こす問題である「矛盾」を無害化してくれるからだ。しかし問題は、パラドクスが引き起こすものが「矛盾」ではない場合だ。第 1 節で見たカリーのパラドクスは、まさにそのようなパラ

[*20] 上述した自明主義においては、自明化は避けるべきものではないため、「ほとんどの場合で」と留保した。

ドクスに他ならない。カリーのパラドクスのような自明化を引き起こすパラドクスに関して、Priest 自身は「Inclosure Schema にカバーされるとも言えるし、されないとも言える」[8, p.169] と答えるが、自明化を引き起こすパラドクスを inclosure paradox として解釈することは実際難しい（紙面の都合上、この点の説明は割愛するが、[13] で述べているので参照されたい）。

本章で紹介する最後の人物、Martin Pleitz は、ラッセルのパラドクスや嘘つきのパラドクスといった inclosure paradox はカリーのパラドクスの特殊例でしかないと考え、Inclosure Schema よりもさらに一般的な構造——カリーのパラドクスをカバーできるような構造——を考察した人物である。その構造、Curry Schema は、Russell's Schema や Inclosure Schema と同じく三つの条件から構成される [6, p.9]。

> φ と ψ を述語、δ を函数とする。このとき以下の三つの条件が成立する。
> 1. $\exists\Omega(\Omega = \{x|\varphi(y)\} \wedge \psi(\Omega))$
> 2. $\forall x(x \subseteq \Omega \wedge \psi(x) \rightarrow (\delta(x) \in x \rightarrow p))$
> 3. $\forall x(x \subseteq \Omega \wedge \psi(x) \rightarrow \delta(x) \in \Omega)$

（第二の条件における「p」は任意の文を指示する。）Russell's Schema や Inclosure Schema と異なり、ここでの函数 δ は「集合 x を入力すると、x に含まれないある要素を出力する」というものではなく、「集合 x を入力すると、『x に含まれているとき、文 p を出力するような要素』を出力する」ものである。δ がこのような函数であるため、第二条件が成り立つことは明らかである。（第一条件、第三条件については、Inclosure Schema と同じである。）

さて、Curry Schema の第二条件における p には、任意の式を入れることが可能である。すなわち、p に $\perp$ を入れることも可能だ。p に $\perp$ を入れたものは、Inclosure Schema の第二条件に他ならない[*21]。それゆえ、Inclosure Schema は Curry Schema の特殊例であるということができる。

さて、上述したように、Russell's Schema や Inclosure Schema の場合、それらの条件が満たされた時に導出されるのは矛盾であった。一方、Curry Schema の場合、上の三つの条件が満たされた時に導出されるものは自明化である。Curry Schema の (2)、(3) より、

$$\Omega \subseteq \Omega \wedge \psi(\Omega) \rightarrow (\delta(\Omega) \in \Omega \rightarrow p)$$

と

$$\Omega \subseteq \Omega \wedge \psi(\Omega) \rightarrow \delta(\Omega) \in \Omega$$

[*21] もっとも、否定を $\neg\alpha := \alpha \rightarrow \perp$ と定義するときに限るが。

が成り立つ。$\Omega \subseteq \Omega$ が成り立ち、かつ、(1) より $\psi(\Omega)$ も成り立つため、$\Omega \subseteq \Omega \wedge \psi(\Omega)$ が成り立つ。よって、

$$\delta(\Omega) \in \Omega \to p$$

と

$$\delta(\Omega) \in \Omega$$

が成り立つ。これにモーダスポネンスを適用することで、p が導出される。この p は任意であったため、この結論は自明化を意味することとなる。

では、Curry Schema によってカバーされるパラドクスはどのように解決されるだろうか。Pleitz の提案は、縮約規則のない論理体系（contraction-free logic, non-contractive logic）を採用することである[*22][*23]。これは、名前の通り、「縮約規則（contraction rule）」

$$\alpha \to (\alpha \to \beta) \vdash \alpha \to \beta$$

を持たない論理体系だ。縮約規則は、「前提 α を『二回』使って結論 β を導出できるなら、前提 α の使用が『一回』であっても結論 β を導出できる」という規則である。カリーのパラドクスはこの縮約規則を制限することで解決される。

$$\dfrac{\dfrac{\dfrac{\dfrac{\varphi \to \varphi}{\varphi \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\ \text{def.}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\ \text{TS}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi}\ \text{Contraction} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\varphi \to \varphi}{\varphi \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\ \text{def.}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)}\ \text{TS}}{\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi}\ \text{Contraction}}{\mathbf{Tr}\ulcorner\varphi\urcorner}\ \text{def.}}{\psi}\ \to E$$

結論 ψ（＝自明化）の導出の過程で、「$\mathbf{Tr}\ulcorner\varphi\urcorner \to (\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi)$」から「$\mathbf{Tr}\ulcorner\varphi\urcorner \to \psi$」を導出する箇所が二箇所ある。この移行が縮約規則に則った移行であることは明らかである。そして、この移行を制限すれば、結論 ψ が導出されなくなり、パラドクスは解決される。

3　広義の自己言及のパラドクスの一般構造

前節で見た自己言及のパラドクスについての研究は、狭義の自己言及のパラドクスについての研究であった。本節ではまず、それらの研究が広義の自己言及のパラドクスをカバーできていないことを確認する。そして、広義の自己言及のパラドクスの一般構造を提

[*22] Priest がパラドクスの解決方法を論理的観点のみならず哲学的観点からも提案した（矛盾許容型論理だけでなく、真矛盾主義的解釈を提案した）のと異なり、Pleitz の提案は論理的観点のみなされている。

[*23] 前述したよう、[6] に記載されている Pleitz 自身の提案は「縮約規則のない矛盾許容型論理」であったが。

示し、そこから分析される広義の自己言及のパラドクスの解決方法の一つについて考察をする。

3.1 （狭義の）自己言及のパラドクスの先行研究の限界

一見、語り得ぬもののパラドクスは矛盾許容型論理によって容易に解決されるように思われる。というのも、第 1 節で見たように、このパラドクスの結論は「x は語り得ぬものであり、かつ、語り得ぬものではない」という矛盾だったからである。結論が矛盾であるパラドクスは、矛盾を無害化してくれる論理体系である矛盾許容型論理によって解消される。矛盾許容型論理で解決されるということは、Inclosure Schmea によってカバーされるということであるため、「語り得ぬもののパラドクスは既存の研究でカバーされる」と言いたくなるかもしれない。

これはある意味においては正しい。一般に「語り得ぬもののパラドクス」と言ったとき、その「受け入れがたい結論」は、「x は語り得ぬものであり、かつ、語り得ぬものではない」という矛盾としてのみ解釈されるからだ。しかし、語り得ぬもののパラドクスは別の「受け入れがたい結論」を持っている。この「受け入れがたい結論」は、自明化の一種だ。自明化とは、第 1 節で見たように、任意の文が導出できてしまうという問題であった。ここで我々が確認する自明化（のようなもの）は、「任意の文について、それが成り立たないということが導出される」ということである。この問題を「否定的自明化」と呼ぼう[*24]。

否定的自明化は以下のようにして導出される。「x は語り得ぬものである（$\neg\mathbf{Ef}(x)$）」を考えよう。これはすなわち、「あらゆる表現は x に帰属させられない（$\forall P(\neg P(x))$）」すなわち、「あらゆる表現 P について、それを x に帰属させたもの（$P(x)$）は成り立たない」ということだ。さてこの「あらゆる表現」には、「語り得ぬものである」という当の表現も含まれるが、例えば「x は白色である、または、今日は月曜日である」という表現も含まれる。この「x は白色である、または、今日は月曜日である」が x に帰属させられない——この表現を x に帰属させたものは成り立たない——ということは、「x は白色である、または、今日は月曜日である」が成り立たないということになる。一般に、「A または B」が成り立たないとき、「A」と「B」の両方が成り立たないと解される。そのため、「x は白色である、または、今日は月曜日である」が成り立たないということから、「今日は月曜日である」が成り立たないということが導かれる。同様の方法で、我々は、「今日は火曜日である」「柴犬はかわいい」等々の任意の文について、それらが成り立たないという結論

[*24] 「否定的自明化」という名称を考えるにあたり、Graham Priest 氏にアドバイスをもらった。

を——否定的自明化を——導出することができる。
　否定的自明化の導出を形式的に確認しよう。

$$\cfrac{\cfrac{\cfrac{\cfrac{\neg\mathbf{Ef}(x)}{\forall P(\neg P(x))}\ \text{def.}}{\neg(\varphi(x)\vee\psi)}\ \forall E}{\neg\varphi(x)\wedge\neg\psi}\ \text{DeMorgan}}{\neg\psi}\ \wedge E$$

（ここでの φ と ψ は任意である。）この証明の結論は、矛盾ではなく否定的自明化だ。そのため爆発律を拒否しても——矛盾許容型論理を採用しても——助けにはならない。また、縮約規則も用いられていないため、縮約規則のない論理を採用してもどうにもならない。語り得ぬもののパラドクスが矛盾許容論理や縮約規則のない論理によって解消されないということは、このパラドクスが Inclosure Schema や Curry’s Schema によってカバーされていないということを（そしてもちろんそれらの構造よりも一般性の低い Russell’s Schema によってもカバーされていないことも）意味する[*25]。
　ところで、「あらゆる表現 P に関して、それを x に帰属させたものは成り立たない：$\forall P(\neg P(x))$」から「『x は白色である、または、今日は月曜日である』が成り立たない：$\neg(\varphi(x)\vee\psi)$」を導出することに対して違和感を抱く人もいるだろう。というのも、「今日は月曜日である」は、（それが「または」で結ばれていたとしても、）x に帰属される表現というより、独立した表現（文）だからだ。それゆえ、この議論は成り立たないと感じる人もいるだろう。そのような人には別の議論を提示しよう。「あらゆる表現 P に関して、それを x に帰属させたものは成り立たない」から「『x が全能ならば、世界から悪はなくなる』は成り立たない」を導出することに違和感を抱く人はいないだろう。つまり、「$\forall P(\neg P(x))$」から「$\neg(\varphi(x)\rightarrow\psi)$」を導出することは許容できるということだ。さて、$\varphi$ として「x と同一である」を採用しよう。このときもちろん、$\varphi(x)$ は「$x = x$」である。ここまできたら否定的自明化まで一直線だ。「あらゆる表現 P に関して、それを x に帰属させたものは成り立たない：$\forall P(\neg P(x))$」から「『$x = x$ ならば、今日は月曜日である』は成り立たない：$\neg(x = x\rightarrow\psi)$」が得られる。いかなる対象であれ自己同一性があることは成り立

[*25] ちなみに、「$\forall P(\neg P(x))$」から「$\neg(\varphi(x)\vee\psi)$」ではなく「$\neg(\varphi(x)\vee\neg\psi)$」を導出すれば、結論として「$\neg\neg\psi$」すなわち「$\psi$」が導出されるため、否定的自明化だけでなく、（否定的自明化よりも強い結論である）自明化も起こせると考える人もいるかもしれない。だが、「$\neg\neg\psi$」から「ψ」を導出するということは、二重否定除去則（$\neg\neg A\vdash A$）を受け入れるということである。二重否定除去則は排中律（$\vdash A\vee\neg A$）がなければ成り立たないが、排中律のない論理体系も多く（例えば直観主義論理では成り立たない）、本稿では排中律を自明のものとして受け入れないため、広義の自己言及のパラドクスの結論は「自明化（および否定的自明化）」ではなく「否定的自明化」のみとする。

つので、「$x = x$ ならば、今日は月曜日である」が成り立たないためには、「今日は月曜日である」が成り立っていてはいけない[*26]。同様の方法で、「今日は水曜日である」「コーギーはかわいい」等々についても、それらが成り立たないという結論を得ることができる。

これを形式的に表すと以下のようになる[*27]。

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg\mathbf{Ef}(x)}{\forall P(\neg P(x))}\ \text{def.}}{\neg(\varphi(x) \to \psi)}\ \forall E}{\neg(\neg\varphi(x) \vee \psi)}\ \text{def.} \to}{\neg\neg\varphi(x) \wedge \neg\psi}\ \text{DeMorgan}}{\neg\psi}\ \wedge E}{}$$

この場合も、矛盾許容型論理や縮約規則のない論理が無力なのは変わらない。結論は矛盾ではなく否定的自明化であるし、縮約規則は用いられていないからだ。そのため、語り得ぬもののパラドクスは、（狭義の）自己言及のパラドクスの一般的解決方法を探る既存の研究ではカバーされない。

3.2　広義の自己言及のパラドクスの一般構造

それでは、語り得ぬもののパラドクスをはじめとする広義の自己言及のパラドクスが共通して持つ構造を見てみよう。（本稿では広義の自己言及のパラドクスとして語り得ぬもののパラドクスしか扱っていないが、以下の構造は、語り得ぬもののパラドクスだけでなく、[16, 第 3 章] で示したように、他の広義の自己言及のパラドクスにも当てはまる構造

---

*26 「あらゆるものは自分と同一である」という主張を妥当だと思わない人は、代わりに別の「任意の対象に関して成り立つ性質」を入れて読んでいただきたい。とはいえ、証明図からもわかるように、φ として何を持ってくるかは、結論 $\neg\psi$ を導出するに際して一切関係がない。

*27 「ならば：→」を「$A \to B := \neg A \vee B$」と定義することに異論のある者もいるだろうが、この定義を用いなくとも否定的自明化という結論は導出可能である。

$$\dfrac{\dfrac{\dfrac{\dfrac{\neg\mathbf{Ef}(x)}{\forall P(\neg P(x))}\ \text{def.}}{\neg(\varphi(x) \to \psi)}\ \forall E \qquad \dfrac{[\psi]_v}{\varphi(x) \to \psi}\ \text{Weakening}}{\perp}\ \to E}{\neg\psi}\ \to I_v$$

なお、ここに出てくる「Weakening」は弱化規則（weakening rule）

$$\frac{\beta}{\alpha \to \beta}$$

「β が成り立つ場合、α の時にも β は成り立つ」を意味する。この規則は古典論理や直観主義論理をはじめ、多くの論理体系で採用されている。

であることに注意されたい。）その構造は以下のようなものであると考えられる。

ある性質 Γ に対して、

1. $\exists\Omega(\Omega = \{Y : \exists x Yx \wedge \Gamma(Y)\})$
2. $\exists a \forall Y \in \Omega(a \notin Y)$

この構造を「否定的自明主義図式（negative trivialism Schema）」、以下「NT 図式」と呼ぼう。それぞれの条件について以下で詳しく述べる。

■第一条件　既存の構造たちの第一条件は、Russell's Schema においては

1. $\exists\Omega(\Omega = \{y : \varphi(y)\})$

であり、Inclosure Schema と Curry Schema においては

1. $\exists\Omega(\Omega = \{y : \varphi(y)\} \wedge \psi(\Omega))$

であった。これらと比べると、否定的自明主義図式の第一条件

1. $\exists\Omega(\Omega = \{Y : \exists x Yx \wedge \Gamma(Y)\})$

は似ているが少し異なる。それは、既存の構造においては Ω が「対象の集合」であったのに対して、否定的自明主義図式において Ω は「述語の集合」だからである。この違いは、既存の研究が扱ってきたのは狭義の自己言及のパラドクスであり、本稿が扱うのは広義の自己言及のパラドクスである——述語と関わるパラドクスである——ことに由来する。

ちなみに、Γ は Y が満たすべき条件について述べている。語り得ぬもののパラドクスではあまり重要な条件ではない[*28]が、他のパラドクスで効いてくる条件だ[*29]。

■第二条件　次に、第二条件について見ていこう。

2. $\exists a \forall Y \in \Omega(a \notin Y)$

この条件は、「いかなる Ω の要素も満たさないような対象 a が存在する」と述べる条件だ。それはつまり、（ある条件 Γ を満たす）いかなる性質ないし表現を満たさないような対象

[*28] 語り得ぬもののパラドクスの場合、Ω は表現の集合なので $\Gamma(Y)$ は「Y は表現である」という、トリビアルにも思われる条件でしかない。

[*29] 例えば、[13] でも挙げているベリーのパラドクス（に似たパラドクス）、「x は 10000 字以下では表現されない」から生じるパラドクスでは、この条件が重要だ。このパラドクスでは、Ω として「10000 字以下の表現」を持ってきたいため、$\Gamma(Y)$ を「Y は 10000 字以下の表現である」という条件にする。詳しい説明は [16] で行う。

が存在するということである。

上の二つの条件が満たされたとき、「否定的自明化（$\neg\psi$）」という結論は以下のように生じる。「または」文から否定的自明化は以下のように導出される。

$$\dfrac{\exists a \forall Y \in \Omega(a \notin Y) \qquad \dfrac{\dfrac{\dfrac{\forall Y \in \Omega(a \notin Y)}{\neg(\varphi(a) \vee \psi)}\,\forall E}{\neg\varphi(a) \wedge \neg\psi}\,\text{DeMorgan}}{\neg\psi}\,\wedge E}{\neg\psi}\,\exists E$$

$\varphi(x) \vee \psi$ は Ω に含まれるようなものを選ぶ必要があることには注意されたい。なお、ド・モルガン則が好みでなければ以下の証明を参照されたい。

$$\dfrac{\exists a \forall Y \in \Omega(a \notin Y) \qquad \dfrac{\dfrac{\dfrac{\forall Y \in \Omega(a \notin Y)}{\neg(\varphi(a) \vee \psi)}\,\forall E \qquad \dfrac{[\psi]_w}{\varphi(a) \vee \psi}\,\vee I}{\bot}\,\to E}{\neg\psi}\,\to I_w}{\neg\psi}\,\exists E$$

また、「ならば」文から否定的自明化は以下のように導出される。

$$\dfrac{\exists a \forall Y \in \Omega(a \notin Y) \qquad \dfrac{\dfrac{\dfrac{\dfrac{\forall Y \in \Omega(a \notin Y)}{\neg(\varphi(a) \to \psi)}\,\forall E}{\neg(\neg\varphi(a) \vee \psi)}\,\text{def.} \to}{\neg\neg\varphi(a) \wedge \neg\psi}\,\text{DeMorgan}}{\neg\psi}\,\wedge E}{\neg\psi}\,\exists E$$

$$\dfrac{\exists a \forall Y \in \Omega(a \notin Y) \qquad \dfrac{\dfrac{\dfrac{\forall Y \in \Omega(a \notin Y)}{\neg(\varphi(a) \to \psi)}\,\forall E \qquad \dfrac{[\psi]_v}{\varphi(a) \to \psi}\,\text{Weakening}}{\bot}\,\to E}{\neg\psi}\,\to I_v}{\neg\psi}\,\exists E$$

なお、こちらの場合も、$\varphi(x) \to \psi$ は Ω に含まれるようなものを選ぶ必要がある。

3.3 広義の自己言及のパラドクスを止めるには：解決方法の候補とその限界

冒頭でも述べたように、パラドクスとは「正しそうな前提」と「妥当そうな推論」から「受け入れがたそうな結論」が導出されるという困難だ。そのため、パラドクスを解決す

るためには、「正しそうな前提」が正しくないと言うか、「妥当そうな推論」が妥当ではないと言うか、「受け入れがたそうな結論」が受け入れられると言えばよい[*30]。つまり、広義の自己言及のパラドクスの場合、

1. 正しそうな前提（否定的自明主義図式の二条件）を棄却する
2. 妥当そうな推論を棄却する
3. 受け入れがたそうな結論（否定的自明化）を受け入れる

のいずれかを行えばよいということだ。

紙面の都合上、すべてのオプションについて見ていくことはできないため、冒頭でも述べたように、可能な解決方法の一つを本稿では紹介する。また、その解決方法の問題点についても考察する[*31]。

その解決方法は、「2. 妥当そうな推論を棄却する」に当たる。先ほど、NT 図式の三条件から否定的自明主義が導出される証明を四つ記載した。それぞれの証明において使用されている規則、使用されていない規則は異なるものの、四つの証明すべてで用いられている規則も存在する。そのうちの一つが、一階の存在量化子 $\exists$ の除去規則だ。四つの証明全てにおいて、「$\exists a \forall Y \in \Omega(a \notin Y)$」と「$\forall Y \in \Omega(a \notin Y) \vdash \neg\psi$」から「$\neg\psi$」への移行が行われている。これはつまり、一階の存在量化子 $\exists$ の除去規則という妥当そうな推論規則を棄却することで、否定的自明化という結論を回避することができるということだ。

■一階の存在量化子 $\exists$ の導入と自由論理　ちなみに、一階の存在量化子 $\exists$ の（除去ではなく）導入をブロックすることは見た目に反しそう難しくない。これが一般的に成り立たないとする論理体系があるからだ。その論理体系は自由論理（free logic）と呼ばれる。詳細は省くが[*32]、自由論理の特徴は、「自由論理において表すことができるもの、自由論理において言及することができるものは、存在するものよりも多い」というものだ。そのため、自由論理では、存在量化子 $\exists$ や普遍量化子 $\forall$ が走る領域（ドメイン）、つまり存在するものの領域の外側にも言及できるものがあることになる。そしてその結果として、自由論理においては「t は A である：At」から「A なものが存在する：$\exists xAx$」への移行が一般

[*30] Priest[8] は「受け入れがたそうな結論（矛盾）」は受け入れられるものだと言うことで、Pleitz[6] は「妥当そうな推論（縮約規則を用いた推論）」が妥当ではないと言うことでパラドクスの解決を目指したと言える。ちなみに Russell の解決方法、タイプ理論は、「正しそうな前提」を正しくないと言うものだと考えられる。詳しくは [16] で述べるが、タイプ理論によって、Russell's Schema の第一条件 $\exists\Omega(\Omega = \{y : \varphi(y)\})$ が成り立たなくなるからだ。

[*31] [16, 第 4 章、付録] ではすべてのオプションについて考察をしているので、興味のある方はそちらを参照されたい。

[*32] [5, p.123-124] を参照されたい。

には成り立たなくなる。例えば「ゼウスは雷を操る神である」が成り立っていても、「ゼウス」という存在しないものは（存在）量化子が走るドメインに含まれないため、「雷を操る神が存在する」は成り立たなくなる。そして「t は A である：At」から「A なものが存在する：$\exists xAx$」への移行が一般には成り立たないということは、存在量化子 $\exists$ の導入が一般には成り立たないということだ。このようにして、自由論理ではこの移行がブロックされる[*33]。

■問題の移行をブロックするために　自由論理は強力で、存在量化子の導入（$\exists I$）をブロックすることができる。だが、その自由論理でも、存在量化子の除去（$\exists E$）をブロックすることはできない。つまり、本論で問題となっている移行、「$\exists a\forall Y \in \Omega(a \notin Y)$」と「$\forall Y \in \Omega(a \notin Y) \vdash \neg\psi$」から「$\neg\psi$」への移行をブロックすることはできない。

だが、存在量化子の除去（$\exists E$）をブロックすることは不可能ではない。先ほど述べたように、自由論理では「言及することができるものは、存在するもの（量化のドメインに含まれるもの）よりも多い」とすることで、存在量化子の導入（$\exists I$）をブロックした。逆に言えば、Fujikawa[2] が示唆するように、「存在するもの（量化のドメインに含まれているもの）は、言及できるものよりも多い」とすることで、存在量化子の除去をブロックすることができる[*34]。

A なものが存在する（$\exists xAx$）と仮定しよう。「A なものが存在する」とはわかっていても、その「A なもの」が言及できるものではないとき、「この…が A なのだ」と言うことはできない。つまり「（ある）t は A である：At」と言うことはできない。存在量化子の除去とは、「$\exists xAx$」と「$At \vdash \varphi$」から「φ」を導出するものだった。「At」が言えないとき、当然、「$At \vdash \varphi$」を得ることもできない。そのため存在量化子 $\exists$ の除去を行うことができなくなるのだ。

この方法を採用して問題となっている移行をブロックすることを考えよう。上で見たように、存在量化子の除去が行われないのは、「A なもの」が（存在はすれど）言及できるものではない場合だった。つまり、「$\exists a\forall Y \in \Omega(a \notin Y)$」と「$\forall Y \in \Omega(a \notin Y) \vdash \neg\psi$」から「$\neg\psi$」が導出されないのは、「$\forall Y \in \Omega(a \notin Y)$ である a」が（存在はしていても）言及できるものではない場合だ。

[*33] ちなみに、同様の理由から「あらゆるものは A である：$\forall xAx$」から「t は A である：At」への移行も自由論理では成り立たない。

[*34] 自由論理と Fujikawa で「存在するもの（量化のドメインに含まれるもの）」の意味は異なる。自由論理の量化のドメインは「存在する対象」のみに制限されているが、Fujikawa においてそのような制限は一切ないだろう。むしろ、彼においては、「自分自身と同一ではないような対象」といった不思議な対象すら含まれている。

語り得ぬもののパラドクスにおいて、Ω は表現の集合だった。そのため「$\forall Y \in \Omega(a \notin Y)$」な a とは、「いかなる表現でも表すことができないもの」である。このようなものが「（存在はしていても）言及できるものではない」という主張に異論のある者はいないだろう。また、[16] でも述べたように、他の広義の自己言及のパラドクスにおいても、「$\forall Y \in \Omega(a \notin Y)$」な a が「（存在はしていても）言及できるものではない」ということは不自然な主張ではない。[*35]

■「存在するものは言及できるものよりも多い」というアイデアの問題　「存在するものは言及できるものよりも多い」というアイデアにより、一階の存在量化子 ∃ の除去が行えなくなるため、広義の自己言及のパラドクスは解消される。だが、この解決方法を採用すれば万事解決というわけではない。というのも、「存在するものは言及できるものよりも多い」という一見すると自然で当たり前のアイデアは、いくつかの問題を孕んでいるからだ。

■問題１：任意の対象についての主張が証明できなくなる　「存在するものは言及できるものよりも多い」とすることでブロックされるのは、広義の自己言及のパラドクスに用いられる「存在量化子 ∃ の除去」だけではない。存在量化子 ∀ の導入

$$\frac{Px}{\forall x Px}$$

も同じくブロックされるのだ。この規則は、「条件なしに x を選んできたとき、x が何であっても P が成り立つなら、任意の（すべての）対象について P が成り立つ」という規則だ。この規則は、「存在するものが言及できるものよりも多い」とき、成り立たない。条件なしに選んできたいかなる x に関しても P が成り立つとしても、その「条件なしに選んできた x」は言及できる対象の範囲から選ばれてきたものでしかない。そのため、「存在するが言及できない対象」について、それが P であるとは限らないのだ。よって、（存在するが言及できない対象を含めた）任意の対象について P が成り立つとは言い切れない。よって ∀ の導入はブロックされるのだ。これにより、「任意の対象は…である」という形の主張を証明によって導出することができなくなり、任意の対象についての言明は、証明されたものではありえず、せいぜい「筆者の主張の宣言」でしかなくなる。

[*35] とはいえ、[16] や [13] でも紹介している、ベリーのパラドクス（に似たパラドクス）においては、[16] で記したように、「『$\forall Y \in \Omega(a \notin Y)$』な a が、（存在はしていても）言及できるものではない」という主張に問題がないとは言い切れない。だが、詳しくは別稿に譲るが、この問題は比較的容易に解消されると筆者は考える。

とはいえ、ほとんどの場合において、任意の対象、すべての対象についての主張を証明することはそもそも難しいため、これは深刻な問題ではないと考える人もいるだろう。例えば、「すべてのカラスは黒い（任意の対象について、それがカラスならば、それは黒い）」を示すためにはすべてのカラスを調べなければならないが、それは実質的に不可能である[*36]。

■問題２：再出　語り得ぬもののパラドクス　存在量化子 ∀ の導入がブロックされることが問題ではないとしても、「存在するものは言及できるものよりも多い」というアイデアには他の、より深刻な問題がある。

存在するものは言及できるものよりも多いということは、存在するが言及できない対象があるということだ。だが、我々はその言及できない対象について、「（存在するが）言及できない対象」という仕方で言及をしてしまっている。つまり、問題となっている対象は、言及できない対象であり、かつ、言及できる対象ということになる。語り得ぬもののパラドクスを含めた広義の自己言及のパラドクスを解消するためのアイデアが、語り得ぬもののパラドクスを再出させてしまっているのだ。

この問題に対し、「『言及』の階層化を考えればよいのではないか」という考えが出るのは自然だろう。すなわち、「ある対象が言及できない」ということを、「ある対象が、ある言語において言及できない」と捉えて、その対象に対する言及（「言及できない対象である」という言及）は、その言語に対するメタ的な言語からの言及であるとするアイデアだ。つまり、言及できない対象の言及不可能性はある言語においてのものであり、言及可能性はそれのメタ言語においてのものであるというアイデアだ。詳しい議論は [15, sec.2] を参照されたいが、このアイデアを採用すると、あらゆるものが「言及できない対象」となる。それゆえ、「このリンゴは赤い」「明日は金曜日である」といった言明すらできなくなる（少なくとも、そのような言明が真であることはありえなくなる）。「このリンゴ」や「明日」も言及できない対象となるからだ。このような破壊的な結論が生じる以上、「言及の階層化」というアイデアは妙手ではないだろう。

[*36] 正確には、「すべてのカラスは黒い」を示すためにすべてのカラスを調べなくてもよく、「すべての黒くないものはカラスではない」を示すことで代用できる。とはいえこちらを示すことも実質的には不可能である。

4　おわりに

本稿では、（狭義の）自己言及のパラドクスの一般的解決方法を探る先行研究を概観し、それらの研究では語り得ぬもののパラドクスのような広義の自己言及のパラドクスはカバーしきれないということを確認した。そして、広義の自己言及のパラドクスの一般構造を提示した。この一般構造を分析することで、どのような解決方法のオプションがあるのかが明らかになる[*37]。すなわち、パラドクスを解決するためには、(1) 前提が実は正しくないと示すか、(2) 推論が実は妥当ではないと示すか、(3) 結論が実は受け入れられると示せばよいため、広義の自己言及のパラドクスの解決方法としては、(1) NT 図式の二条件のいずれかが正しくないと示すか、(3) 否定的自明化を受け入れるか、(2) NT 図式の二条件から否定的自明化を導出する際に用いられる推論規則を棄却するか、の三つのオプションが存在するのである。

だが、構造の分析によって解決方法の候補がわかるということと、その「解決方法の候補」の採用が妙手であるかどうかは別のことである。実際、本稿では、構造を分析することで「存在量化子 ∃ の除去を制限する」ことが解決方法の候補であることを示したが、この解決方法の採用を正当化するアイデアである「存在するものは言及できるものよりも多い」というアイデアは、いくつかの問題を孕むものであった。そのため、少なくともこのアイデアを通して「存在量化子 ∃ の除去を制限する」という解決方法を採用することが妙手であると判断するのは拙速であろう。

だが、注意されたいが、「あるアイデアを通して『広義の自己言及のパラドクスの共通の解決方法の候補』の一つを採用すること」に慎重にならねばならないことは、広義の自己言及ののパラドクスの共通の解決方法の探求が振り出しに戻ったことを全く意味しない。我々は広義の自己言及のパラドクスの構造をすでに把握しているし、その構造の分析により、それらパラドクスの解決方法の候補もすでに把握している。つまり、共通の解決方法の探求において今後すべきは、一から解決方法を考えることではなく、解決方法（の候補）の採用を正当化するアイデアの発明とそのアイデアの検討だ。

冒頭でも述べたように、広義の自己言及のパラドクスについての研究（および「狭義・広義両方の」自己言及のパラドクスについての研究）は十分とは言い難い。本稿がその研究の一助となれば幸いである。

[*37] また、これまで提案されてきた「パラドクスの解決方法」が、どのオプションに分類されるかがわかるようになるため、（本稿では行わなかったが、）パラドクスの解決方法を整理することも可能となる。

References

[1] Otavio Bueno. Troubles with trivialism. *Inquiry*, 2007.

[2] Naoya Fujikawa. Ineffability and nonobjecthood. In *What's so bad about dialetheism?*, 2018.

[3] Paul Douglas Kabay. *A Defense of Trivialism*. PhD thesis, The University of Melbourne, 2008.

[4] Andre Kukla. *Ineffability and Philosophy*. Routledge, 2005.

[5] Karel Lambert. *Free Logic: Selected Essays*, chapter The Philosophical Foundations of Free Logic. Cambridge University Press, 2002.

[6] Martin Pleitz. Curry's paradox and the inclosure schema. *Academia*, 2014.

[7] William Poundstone. *Labyrinths of Reason: Paradox, Puzzles, and the Frailty of Knowledge*. Anchor Press, 1988.

[8] Graham Priest. *Beyond the Limits of Thought*. Oxford University Press, 2 edition, 2002.

[9] Graham Priest. *Doubt Truth to Be a Liar*. Oxford University Press, 2 edition, 2008.

[10] Frank R. Ramsey. The foundations of mathematics. In *Foundations — Essays in Philosophy, Logic, Mathematics and Economics*. Routledge and Kegan Paul, 1931.

[11] Bertrand Russell. *The Principles of Mathematics*. W.W. Norton and Co. Inc, 1996.

[12] Bertrand Russell. On some difficulties in the theory of transfinite numbers and order types. In Gregory Moore, editor, *The Collected Papers of Bertrand Russell*, volume 5. Routledge, 2014.

[13] Maiko Yamamori. Critique of previous comprehensive studies of self-referential paradoxes. *Tetsugaku*, volume 3, 2019

[14] 山岡悦郎. うそつきのパラドックス. 海鳴社, 2004.

[15] 山森真衣子. 語り得ぬものに関する普遍主義と真矛盾主義. 哲学論叢, 2016.

[16] 山森真衣子. 広義の自己言及のパラドクスの解決方法とそのコスト. PhD thesis, 京都大学, 2019.

[17] 廣松渉 他, editor. 岩波哲学・思想事典. 岩波書店, 1998.

[18] 久木田水生 池田真治, 伊藤遼. タイプ理論の起源と発展. 哲学論叢, 38(別冊):S49–60, 2011.

（筆者　やまもり・まいこ　京都大学大学院文学研究科／哲学）

General Structure and Solution(s) of Broad self-referential Paradoxes

by

Maiko YAMAMORI

Department of Philosophy
Graduate School of Letters
Kyoto University

This paper will handle with self-referential paradoxes, especially paradoxes called 'broad self-referential paradoxes' here.

There are some paradoxes that seem to have self-reference in some senses, but have not been taken as self-referential ones in terms of existed criterion of 'self-reference'. In order to save such our intuition, I will introduce new criterion of 'self-reference' and categorize self-referential paradoxes into two groups; I will call paradoxes that fit existed criterion of self-reference 'narrow self-referential paradoxes', and paradoxes that seem to be self-referential but does not fit the criterion 'broad self-referential paradoxes'.

The goal of my study is to find the best solution applicable to all self-referential paradox (not only narrow ones but both of narrow and broad ones). To find such general solution requires to find the general structure of all self-referential paradoxes which is essential for the paradoxes occurring; we can find candidates of solutions by analyzing the structure, and we can consider and decide the best solution among these candidates by comparing them. And to find the general structure of all self-referential paradoxes requires to find in advance both of 'the general structure of narrow self-referential paradoxes' and 'the general structure of broad self-referential paradoxes'. Some existed studies try to find general structures (and general solutions) of narrow self-referential paradoxes; there is, however, no study about general structures (and general solutions) of broad ones. So this paper aims at finding it. Moreover, in this paper I will analyze the structure and show candidates of general solutions to broad self-referential paradoxes.

This paper consists of three sections. First, before getting into the main topic, I will show some self-referential paradoxes in the first section, including broad one. Then, in the next section, I will introduce existed works try to find general structures of self-referential para-doxes. In the last section, firstly I will show that the works introduced in the second section can't cover broad self-referential paradoxes. Then I will show the general structure of them, and by analyzing it I will show the candidates of its solutions.