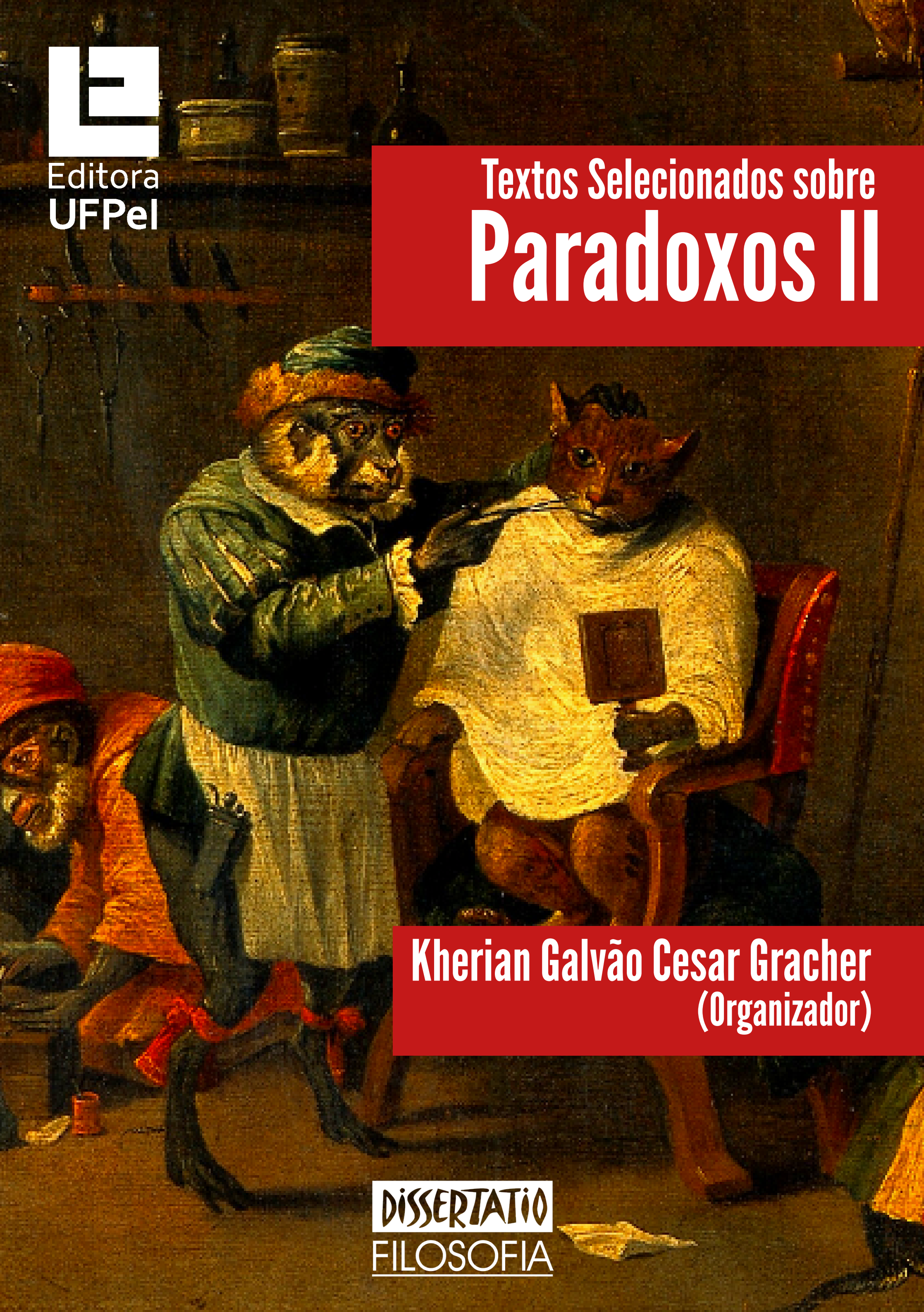

Editora UFPel

Textos Selecionados sobre Paradoxos II

Kherian Galvão Cesar Gracher
(Organizador)

DISSERTATIO FILOSOFIA

TEXTOS SELECIONADOS SOBRE PARADOXOS II

SÉRIE INVESTIGAÇÃO FILOSÓFICA

TEXTOS SELECIONADOS SOBRE PARADOXOS II

Kherian Galvão Cesar Gracher

Pelotas, 2025.

Universidade Federal de Pelotas
Departamento de Filosofia
Núcleo de Ensino e Pesquisa em Filosofia
Editora da Universidade Federal de Pelotas

NEPFil online
Rua Alberto Rosa, 154 – CEP 96010-770 – Pelotas/RS

Primeira publicação em 2025 por NEPFil online e Editora da UFPel.

Dados Internacionais de Catalogação

N123 Textos selecionados sobre paradoxos II. [recurso eletrônico] - Organizador: Kherian Galvão Cesar Gracher – Pelotas: NEPFIL Online, 2025.

208p. - (Série Investigação Filosófica).

Modo de acesso: Internet
<wp.ufpel.edu.br/nepfil>
ISBN: 978-65-998644-7-6

1. Paradoxos 2. Filosofia I. Gracher, Kherian Galvão Cesar. COD 100

Série Investigação Filosófica

A Série Dissertatio Filosofia, uma iniciativa do ***Núcleo de Ensino e Pesquisa em Filosofia*** do Departamento de Filosofia da UFPel e do **Grupo de Pesquisa Investigação Filosófica** do Departamento de Filosofia da UFOP, sob o selo editorial do NEPFil online e da Editora da Universidade Federal de Pelotas, tem por objetivo precípuo a publicação de obras autorais. O objetivo geral da série é disponibilizar materiais bibliográficos relevantes tanto para a utilização enquanto material didático quanto para a própria investigação filosófica.

EDITORES DA SÉRIE

Juliano Santos do Carmo (NEPFIL/UFPEL)

Rodrigo Lastra Cid Reis (SIF/UFOP)

COMISSÃO TÉCNICA

Juliano Santos do Carmo (Capista)

ORGANIZADOR DO VOLUME

Kherian Galvão Cesar Gracher (UFRJ)

CRÉDITOS DA IMAGEM DE CAPA. Disponível em:
https://wellcomeimages.org/indexplus/image/V0017621.html

Em memória de Gibran Almeida,
que ao sonhar em pertencer à comunidade lógica,
mal percebeu que já era um de nós...

Sobre a série Investigação Filosófica

A *Série Investigação Filosófica* é uma série de livros de traduções de verbetes da Enciclopédia de Filosofia da Stanford (*Stanford Encyclopedia of Philosophy*), que intenciona servir tanto como material didático para os professores das diferentes subáreas e níveis da Filosofia quanto como material de estudo para a pesquisa e para concursos da área. Nós, professores, sabemos o quão difícil é encontrar bom material em português para indicarmos. E há uma certa deficiência na graduação brasileira de filosofia, principalmente em localizações menos favorecidas, com relação ao conhecimento de outras línguas, como o inglês e o francês. Tentamos, então, suprir essa deficiência, ao introduzirmos essas traduções ao público de língua portuguesa, sem nenhuma finalidade comercial e meramente pela glória da filosofia.

Essas traduções foram todas realizadas por filósofos ou por estudantes de filosofia supervisionados e revisadas por especialistas na área. Todas as traduções de verbetes da Stanford foram autorizadas pelo querido Prof. Dr. Edward Zalta, editor da Enciclopédia de Filosofia da Stanford; por isso o agradecemos imensamente. Sua disposição para ajudar brinda os países de língua portuguesa com um material filosófico de excelência, que será para sempre disponibilizado gratuitamente no site da Editora da Universidade Federal de Pelotas (Editora UFPel), dado o nosso maior princípio se fundar na ideia de conhecimento livre e a nossa maior intenção ser o desenvolvimento da filosofia em língua portuguesa e do seu ensino. Aproveitamos o ensejo para agradecer também ao editor da Editora UFPel, na figura do Prof. Dr. Juliano do Carmo, que apoiou nosso projeto desde o início. Agradecemos também a todos os organizadores, tradutores e revisores, que participam de nosso projeto. Sem sua dedicação voluntária, nosso trabalho não teria sido possível. Esperamos, com o início desta coleção, abrir as portas para o crescimento desse projeto de tradução e trabalharmos em conjunto pelo crescimento da filosofia em português.

Prof. Dr. Rodrigo Reis Lastra Cid
Prof. Dr. Juliano Santos do Carmo
Editores da Série Investigação Filosófica

Sumário

Introdução de Russell: A Introdução de Todas as Introduções que Não se Introduzem

No volume anterior, *Textos Selecionados de Paradoxos I*, acompanhamos paradoxos que percorriam diferentes territórios: da cognoscibilidade e da vagueza à tradição eleática e às tensões lógico-matemáticas contemporâneas. A cada passo, víamos como esses enigmas funcionam como espelhos que revelam fissuras em conceitos fundamentais e exigem da filosofia a coragem de revisar pressupostos.

O presente volume, entretanto, volta-se para um domínio específico e especialmente instigante: os *paradoxos semânticos e autorreferenciais*, aqueles que não se apoiam em fronteiras imprecisas da experiência ou em intuições sobre movimento e conhecimento, mas que emergem da própria *linguagem* com que raciocinamos. Aqui, as tensões não são externas, mas internas: ligam-se diretamente ao uso do predicado de verdade, aos princípios de compreensão, à autorreferência e às regras formais de inferência.

Esses paradoxos ocupam lugar central na lógica e na filosofia contemporâneas. Foi diante deles que Tarski formulou sua hierarquia de linguagens para definir a verdade sem contradição; que Kripke e Gupta–Belnap buscaram alternativas não hierárquicas com pontos fixos e teorias da revisão; que Russell expôs a insuficiência da teoria ingênua dos conjuntos, conduzindo a Zermelo–Fraenkel, tipos e sistemas alternativos; e que lógicas não clássicas – paraconsistentes, paracompletas, relevantes – passaram a ser consideradas não apenas como curiosidades, mas como respostas possíveis a dilemas semânticos genuínos.

O que está em jogo, portanto, é a própria estabilidade das nossas noções de *verdade*, *definição* e *inferência*. Ao contrário de paradoxos como os de Zenão ou Fitch, que nos desafiam a repensar o espaço-tempo ou a relação entre conhecimento e possibilidade, aqui o desafio é mais profundo: são as próprias engrenagens da coerência lógica que parecem se desajustar. Como observa Barwise e Etchemendy, a importância de um paradoxo nunca está nele mesmo, mas naquilo de que ele é um sintoma – e, neste caso, os sintomas apontam para rachaduras na arquitetura formal que sustenta nossa linguagem e nosso pensamento.[1]

É nesse cenário que se distribuem os capítulos que seguem. Cada um deles examina um ponto nodal dessa constelação de paradoxos semânticos e autorreferenciais, mostrando como raciocínios aparentemente inofensivos podem se converter em dilemas que desafiam a lógica, a semântica e a própria filosofia.

No capítulo (I) (p. 19–84), **Paradoxos e a Lógica Contemporânea**, mapeia-se como as antinomias do fim do século XIX e início do XX (especialmente na teoria dos conjuntos) impulsionaram a virada metateórica da lógica: de "quebras" locais de intuição para diagnósticos estruturais sobre linguagem, prova, verdade e definição. O texto organiza o panorama em fases históricas, apresenta as famílias paradigmáticas (autorreferência e verdade – como Mentiroso e Grelling; tamanho/definibilidade – como Russell, Berry e Richard; e variantes "termais" como o Barbeiro), e reconstrói as respostas *clássicas*: hierarquias de tipos e de linguagem (Tarski), restrições à formação de conjuntos, e esquemas que evitam a diagonalização direta. A partir daí, delineia as *lições das antinomias* (fixpoints, diagonal, fechamento e contagiosidade), e abre o leque das *estratégias não clássicas*: soluções por gluts (paraconsistência), por gaps (paracompletude, supervalorações), modelos de pontos fixos (Kripke) e teorias revisionalistas da verdade. Em seguida, contrasta "lógicas conceituais" – que tomam os paradoxos como guias de engenharia conceitual – com abordagens *revisionalistas* que ajustam regras/axiomas, e esboça uma perspectiva "geométrica" que classifica famílias paradoxais por eixos semânticos e inferenciais. O resultado é um retrato de como a lógica contemporânea transformou os paradoxos em ferramentas para calibrar a fronteira entre linguagem e mundo, e para medir o custo – técnico e filosófico – de cada rota de solução.

No capítulo (II) (p. 85–135), o tema é a **autorreferência** como motor de paradoxos clássicos e contemporâneos. O texto percorre as principais famílias: os *paradoxos semânticos* (Mentiroso e suas variantes fortalecidas, Grelling–Nelson, além dos problemas de definibilidade à la Berry e Richard), os *paradoxos em teoria dos conjuntos* (com destaque para Russell e o papel da compreensão irrestrita) e os *paradoxos algorítmicos/computacionais* (diagonalização, quines e o teorema da recursão de Kleene). Uma seção especial discute o *paradoxo de Yablo*, mostrando como estruturas bem-fundadas – sem laços autorreferenciais

[1] "a importância de um paradoxo nunca está no paradoxo em si, mas naquilo de que ele é um sintoma. Pois um paradoxo demonstra que a nossa compreensão de algum conceito básico ou de um conjunto de conceitos está crucialmente falha, que os conceitos se rompem em casos-limite. [...] Se os conceitos forem importantes, isso não é motivo de riso." *em* BARWISE, Jon; ETCHEMENDY, John. *The Liar: An Essay on Truth and Circularity*. New York; Oxford: Oxford University Press, 1987, pp. 4–5. ISBN 0195059441.

explícitos – também produzem contradições. Em paralelo, o capítulo apresenta os *mecanismos geradores* por trás dessas construções (lema do ponto fixo/diagonalização, quinação, codificação aritmética) e, na sequência, mapeia as *respostas de fundo*: hierarquias e restrições tipadas à maneira de Tarski e Russell, teorias não clássicas de verdade (como os pontos fixos de Kripke) e abordagens revisionistas (Gupta–Belnap), além de limitações axiomáticas a esquemas de compreensão. O resultado é um panorama integrado: da forma lógica comum que viabiliza as antinomias à variedade de estratégias – semânticas, axiomáticas e lógicas – para contê-las sem perder de vista o que elas revelam sobre linguagem, verdade e referência.

No capítulo (III) (p. 136–187), o foco recai sobre o **Paradoxo do Mentiroso** e seu "fenômeno ampliado". O texto inicia com versões canônicas – a sentença que diz de si mesma ser falsa, bem como variações como "não-verdadeira", ciclos e cadeias sem fim – e mostra como princípios muito naturais sobre o predicado de verdade (instâncias do esquema-T, introdução/eliminação de "é verdadeiro/falso") bastam para gerar contradição. Em seguida, apresenta o arsenal de respostas contemporâneas, cada qual preservando diferentes fatias do raciocínio ingênuo sobre a verdade. O capítulo também discute contextualismo, a questão do *poder expressivo* e os argumentos de *revanche* (que reaparecem quando fortalecemos o aparato semântico), além de estender a moldura para outros predicados semânticos (definição, referência) e construções auto-aplicativas aparentadas, mapeando com cuidado o que precisamos ceder – na semântica, na lógica ou na arquitetura do idioma – para domar o Mentiroso sem trivializar a teoria.

No capítulo (IV) (p. 188–206), encontramos o **Paradoxo de Russell**, um dos mais célebres e influentes paradoxos da lógica contemporânea. O texto parte da formulação canônica – a classe de todos os conjuntos que não são membros de si mesmos – e mostra como a aplicação irrestrita do princípio de compreensão na teoria ingênua de conjuntos conduz a uma contradição inevitável. Em seguida, o capítulo apresenta o impacto histórico imediato: o colapso do sistema de Frege, as primeiras tentativas de resposta de Zermelo por meio da axiomatização e a teoria dos tipos proposta pelo próprio Russell. São então discutidas as soluções posteriores, desde os sistemas axiomáticos que moldaram a teoria de conjuntos moderna (ZFC, von Neumann–Bernays–Gödel) até alternativas em lógicas não clássicas que procuram preservar versões ampliadas da compreensão. Por fim, o capítulo evidencia como o paradoxo não é apenas uma curiosidade técnica, mas um ponto de inflexão que redefiniu o estudo dos fundamentos da matemática, a semântica formal e a própria noção de coleção.

No capítulo (V) (p. 207–244), o foco recai sobre o **Paradoxo de Curry**, apresentado nas suas duas faces canônicas: a versão semântica, com um enunciado que diz de si mesmo que "se é verdadeiro, então π", e a versão em teoria ingênua dos conjuntos, via compreensão, que constrói um conjunto C_π tal que "se $C_\pi \in C_\pi$, então π". O texto mostra passo a passo como, a partir de princípios mínimos – identidade, *modus ponens* e contração –, obtém-se π *arbitrária*, sem usar negação: uma rota para a trivialidade que difere do Mentiroso por não requerer contradição explícita. Em seguida, o capítulo enuncia o *Lema de Curry* e a *Afirmação Problemática*, introduz a noção de *teoria Curry-completa* (e o corolário de trivialidade) e mapeia as estratégias de contenção em três frentes: (i) *respostas livres de contração*; (ii) *respostas livres de desprendimento*; e (iii) *respostas livres de transitividade*. O capítulo também discute como princípios de *definição* e de *diluição/afrouxamento* podem reativar o esquema curryano, e estende a análise a predicados como *validade* (a chamada "Curry-validade"). O balanço final explica por que Curry é, em certo sentido, mais exigente que o Mentiroso: para manter verdades ingênuas ou compreensão ingênua sem colapso, é preciso pagar um preço claro – enfraquecer regras estruturais, limitar o desprendimento, rever a transitividade ou regrar estritamente os esquemas ingênuos.

O percurso deste volume reflete a centralidade dos paradoxos autorreferenciais na filosofia e na lógica contemporânea. Iniciamos com uma visão geral sobre o modo como tais enigmas moldaram a própria disciplina lógica, passando à análise da autorreferência como mecanismo gerador de inconsistências. Em seguida, examinamos o *Paradoxo do Mentiroso*, talvez o mais emblemático da tradição, para então avançar aos paradoxos de *Russell* e *Curry*, que desestabilizam tanto a teoria ingênua dos conjuntos quanto noções elementares de implicação e consequência.

Reunidos, esses capítulos mostram que os paradoxos semânticos não são meras curiosidades de linguagem, mas forças estruturantes que definem a fronteira entre o aceitável e o insustentável em nossas teorias formais. Eles evidenciam que não basta falar em verdade, definição ou inferência de modo ingênuo: cada passo exige escolhas delicadas sobre hierarquias, restrições e revisões lógicas. Ao desafiar os alicerces do raciocínio formal, os paradoxos aqui estudados não apenas testam os limites da coerência, mas também impulsionam a contínua reinvenção da lógica e da filosofia da linguagem.

Kherian Gracher
Organizador

(I) Paradoxos e a Lógica Contemporânea[1]

Título Original: Paradoxes and Contemporary Logic
Autor: Andrea Cantini e Riccardo Bruni
Tradução: Caio Cezar Silva
Revisão: Annelyze Reis

Por "paradoxo" comumente se entende uma declaração que afirma algo que vai além (ou mesmo contra) a 'opinião comum' (o que geralmente é acreditado ou mantido). Paradoxos constituem um objeto natural da investigação filosófica desde as origens do pensamento racional; eles foram inventados como parte de argumentos complexos e ferramentas para refutar teses filosóficas (pense nos célebres paradoxos creditados a Zenão de Eleia, sobre o movimento, o continuum, a oposição entre unidade e pluralidade, ou dos argumentos envolvendo as noções de verdade e vagueza, creditados a Escola Megárica, e Eubulides de Mileto). Paradoxos — denominados como *Insolubilia* — constituem também uma parte substancial das investigações lógicas e filosóficas durante a Idade Média.

Este verbete concentra-se no surgimento de temas e noções lógicas não triviais a partir da discussão sobre paradoxos do início do século XX até 1945, e tentativas de avaliar a sua importância para o desenvolvimento da lógica contemporânea. Paradoxos envolvendo vagueza, conhecimento, crença, e o espaço e o tempo são tratados em verbetes próprios.

---

[1]CANTINI, Andrea; BRUNI, Riccardo, "Paradoxes and Contemporary Logic", In: ZALTA, E. N.; NODELMAN, U. (eds.). The Stanford Encyclopedia of Philosophy (Summer 2025 Edition). Stanford, CA: The Metaphysics Research Lab, 2025. Disponível em: https://plato.stanford.edu/archives/sum2025/entries/paradoxes-contemporary-logic/.
A seguir está a tradução da entrada sobre Paradoxos e a Lógica Contemporânea de Andrea Cantini e Riccardo Bruni na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/sum2025/entries/paradoxes-contemporary-logic/. Esta versão traduzida pode diferir da versão atual da entrada, que pode ter sido atualizada desde o momento desta tradução. A versão atual está localizada em https://plato.stanford.edu/entries/paradoxes-contemporary-logic/. Agradecemos aos editores Edward N. Zalta e Uri Nodelman pela permissão para traduzir e publicar esta entrada.

Uma advertência terminológica faz-se necessária. A palavra “antinomia” é usada abaixo como uma alternativa para, e como um sinônimo de, “paradoxo”. A maioria dos paradoxos — mas não todos — envolvem contradições; em tais casos, frequentemente se emprega a palavra “contradição”.[2]

1. Introdução

Entre o fim do século XIX e o início do século XX, os fundamentos da lógica e da matemática foram afetados pela descoberta de uma série de dificuldades — os chamados paradoxos — *envolvendo noções fundamentais e métodos básicos de definição e inferência*, os quais eram comumente aceitos como não problemáticos. Desde então, paradoxos adquiriram um novo papel na lógica contemporânea: na verdade, eles levaram a teoremas (geralmente resultados negativos, tais como improvabilidade e indecidibilidade) e não estão confinados apenas ao reino da dialética estéril. Diversas noções básicas de lógica, tal como correntemente ensinadas, alcançaram sua forma atual ao final de um processo que foi desencadeado mediante várias tentativas de resolver paradoxos. Isto é especialmente verdade para noções como *conjuntos* e *coleções* em geral, para os *conceitos semânticos e sintáticos básicos da lógica clássica padrão* (linguagens lógicas de uma dada ordem, a noção de satisfatibilidade, definibilidade). Após os primeiros quarenta anos, os subprodutos dos paradoxos incluíam axiomatizações da teoria dos conjuntos, um desenvolvimento sistemático da teoria dos tipos, os fundamentos da semântica, uma teoria de sistemas formais (pelo menos em sua fase embrionária), bem como a introdução da dicotomia predicativa/impredicativa que foi importante por razões conceituais, mas também para o futuro dos métodos prova-teórica. [3]

2. Paradoxos: primeiros desenvolvimentos (1897–1917)

Os primeiros trabalhos sobre paradoxos de particular importância tratavam das seguintes noções:

---

[2]Não menos importante, com vistas a facilitar a leitura deste verbete, documentos suplementares, tais como os links que aparecerão nas notas de rodapé abaixo, são usados livremente para explorar questões secundárias ou primárias em maior detalhe. As notas de rodapé destinam-se a breves considerações, explicações notacionais ou terminológicas e esclarecimentos quanto à tradução.

[3]N.T.: “*proof-theoretic*”.

1. Números ordinais e cardinais (Burali-Forti, Cantor);
2. Propriedade, conjunto, classe (Russell, Zermelo);
3. Proposição e verdade (Russell);
4. Definibilidade e o continuum aritmético (ou atomista) (Richard, König, Bernstein, Berry, Grelling).

Algumas dessas contradições já são tratadas em verbetes próprios nesta enciclopédia (como o Paradoxo do Mentiroso e o Paradoxo de Russell); a ênfase aqui será nos problemas de fundo, seus vínculos, suas inter-relações e a interação desses paradoxos com questões fundacionais e filosóficas.[4]

2.1 Dificuldades envolvendo números ordinais e cardinais

Os primeiros paradoxos modernos diziam respeito às noções de número ordinal e cardinal. Burali-Forti, um matemático da Escola de Peano, tentou provar que os números ordinais não são linearmente ordenados. Assumindo por contradição que a classe **ON** de todos os ordinais poderia ser linearmente ordenada, ele observou então que **ON** em si mesma seria bem ordenada e que ela possuiria um ordinal $\Omega \in$ **ON**. Portanto, **ON** seria similar (ordem-isomórfica) a um segmento inicial próprio de si mesma, aquele determinado por Ω, contradizendo um teorema bem conhecido sobre conjuntos bem ordenados. O resultado foi publicado em 1897 e, embora o objetivo original de Burali-Forti seja impossível de ser alcançado, seu argumento mostrou que a coleção **ON** é problemática na melhor das hipóteses (Moore-Garciadiego 1981).

O pai da teoria dos conjuntos, Cantor, já percebera dificuldades semelhantes em 1895 (como testemunhado por Bernstein e em cartas a Hilbert e Dedekind). Na verdade, em uma segunda carta a Dedekind, de 31 de agosto de 1899, Cantor apontou um outro problema, envolvendo a noção de número cardinal e sugerindo que não se pode pensar consistentemente em um "conjunto de todos os conjuntos concebíveis", M. Se M fosse um conjunto genuíno, então possuiria um número cardinal m, que seria o número cardinal máximo. Porém, pode-se considerar também o conjunto $\wp(M)$ de todos os subconjuntos de M e, de

[4] N.T.: https://plato.stanford.edu/entries/liar-paradox/
https://plato.stanford.edu/entries/russell-paradox/

acordo com o teorema de Cantor, a cardinalidade de $\wp M$ deveria ser estritamente maior que o suposto m máximo: contradição.

Consequentemente, Cantor sugeriu uma distinção crucial — ainda considerada como "subjetiva", isto é, matematicamente imprecisa, por Hilbert (até 1904, veja van Heijenoort 1967, p. 131 — entre totalidades que não podem ser concebidas como um todo (as *inconsistentes*) e aquelas que podem ser consideradas como completas (*fertige Menge*). De forma aproximada, a primeira é uma coleção que não pode ser um elemento de outras coleções, ao passo que a segunda é uma coleção pequena, que pode ser um elemento de outras coleções. Para mais detalhes, consulte o verbete The Early Development of Set Theory.[5] Isso corresponde à distinção entre classes e conjuntos, mais tarde tornada precisa e axiomatizada na abordagem classe-teórica[6] (von Neumann, Bernays, Gödel); trata-se de uma formulação que remete à doutrina de limitação de tamanho de Russell (ver seção 3.1; mas veja também Garciadiego 1992).

No caso da dificuldade descoberta por Burali-Forti, a consequência para Cantor foi que a multiplicidade (*Mannigfaltigkeit*) de números ordinais é ela própria bem-ordenada, mas não é um conjunto: portanto, nenhum ordinal pode ser atribuído a ela e a antinomia é assim resolvida.

2.2 A contradição de Russell

A segunda antinomia famosa publicada (Russell 1903, §§ 78, 101–106; Frege 1903, Apêndice, Outubro de 1902); veja Klement 2010 e o verbete Russell's Paradox)[7] nos leva do paraíso de Cantor para o reino dos fundamentos da lógica e da filosofia da matemática. É surpreendentemente simples, envolve apenas a aplicação de predicados, e possui um caráter autorreferencial (reflexivo) explícito. Nas palavras do próprio Russell (em carta para Frege, de 16 de Junho de 1902),

> seja w o predicado: ser um predicado que não pode ser predicado de si mesmo. Pode w ser predicado de si mesmo? De cada resposta, o oposto se segue. Do mesmo modo, não há classe (como uma totalidade) daquelas classes que, cada uma tomada como uma totalidade, não pertencem a si mesmas. Disto eu

[5] N.T.: https://plato.stanford.edu/entries/settheory-early/

[6] N.T.: "*class-theoretic*".

[7] N.T.: https://plato.stanford.edu/entries/russell-paradox/

concluo que, sob certas circunstâncias, uma coleção definível não forma uma totalidade. (van Heijenoort 1967, p. 124–125)

Russell estava inicialmente envolvido no estudo das "contradições na relação da quantidade contínua com o número e o continuum" (Moore 1995, p. 219), tendo obtido a sua contradição (Maio 1901) como um resultado da reflexão acerca da antinomia que surge com o teorema de Cantor (Russell 1903, nota 7, § 344; § 100, p. 101). Russell provavelmente percebeu a importância da descoberta somente após a resposta de Frege. O efeito da antinomia é que é impossível ter uma operação de abstração $\phi \mapsto \{x \mid \phi\}$ injetivamente mapeando qualquer conceito ϕ (propriedade) em sua extensão (a classe de todos os x tal que $\phi(x)$) (ou seja, se as classes definidas por ϕ e ψ são iguais, então $\phi(a) \leftrightarrow \psi(b)$, para todo objeto a). Consequentemente, também é impossível estabelecer os fundamentos da teoria dos conjuntos a partir de uma *noção puramente lógica dos conjuntos* onde a pertinência espelha fielmente a aplicação de predicados, no sentido em que, à luz de Frege, $x \in y$ significa que (1) y = $\{x \mid \phi(x)\}$ para algum conceito ϕ, e (2) ϕ realmente se aplica a x (para detalhes históricos sobre a descoberta do paradoxo por Russell, veja Moore 1995).

2.3 O paradoxo de Russell envolvendo proposições e verdade: o surgimento da teoria dos tipos

Em seu *Principles of Mathematics*, Russell (1903) apresenta discussões estendidas dos paradoxos de Russell e Burali-Forti em várias formas nas seções 78, 84–85, 101, 301. O paradoxo de Russell é adaptado para mostrar que uma função proposicional ϕ não pode ser um sujeito lógico (isto é, como separada de seu argumento; é *insaturada* em termos fregeanos); de outro modo, ϕ se aplicaria a si mesma, $\neg\phi(\phi)$ seria uma função proposicional e seria possível reproduzir a inconsistência.

No Capítulo 10, seção 102, Russell também apresenta uma forma do teorema de Cantor, que captura a essência lógica da diagonalização (essa versão já é folclórica): nenhuma relação binária pode parametrizar todos os predicados unários sobre um dado domínio U (ou seja, não existe uma relação binária R tal que para todo predicado unário P sobre U, há um objeto a em U para o qual se tem: para todo x em U, $R(a,x) \leftrightarrow P(x)$).

Em suma, a contradição de Russell evidencia o estado crítico de princípios lógicos aparentemente seguros: ou se desiste da hipótese de que "qualquer função proposicional contendo apenas uma variável é equivalente a afirmar a pertinência a uma classe definida pela função proposicional” (em outros termos, o princípio da compreensão); ou se rejeita a ideia de que “toda classe pode ser tomada como um termo” (p. 102–103).

Nas mãos de Russell, o paradoxo se aplica a predicados, classes, e funções proposicionais, e leva a uma nova imagem do universo lógico-matemático, que é delineada na primeira exposição da doutrina dos tipos: para cada função proposicional ϕ é associado um *alcance de significância*[8], ou seja, uma classe de objetos para os quais um dado ϕ se aplica para produzir uma proposição; ademais, são precisamente os alcances de significância que formam os tipos. No entanto, há objetos que não são alcances de significância; estes são apenas átomos (isto é, *urelemente* ou indivíduos)[9] e eles formam o tipo mais baixo. O próximo tipo consiste em classes ou alcance de indivíduos; então, tem-se classes de classes de objetos do tipo mais baixo, e assim por diante (veja também o verbete Type Theory).[10]

Novas dificuldades surgem se se aceita que proposições formam um tipo (pois são os únicos objetos dos quais se pode afirmar significativamente que são verdadeiros ou falsos). Primeiro, há obviamente pelo menos tantas proposições quanto objetos (basta considerar o mapa associando x à proposição expressa por $(x = x)$; p. 527). De outro lado, se é possível formar tipos de proposições, há mais tipos de proposições do que proposições, segundo o argumento de Cantor. Logo, podemos injetar tipos de proposições em proposições através da noção de *produto lógico*. Seja m uma classe de proposições e seja Πm a proposição “toda proposição de m é verdadeira” (tomada como uma conjunção possivelmente infinitária); então, se m e n são classes diferentes, as proposições Πm e Πn são diferentes, ou seja, o mapa associando m ao seu produto Πm é injetivo. Portanto, se consideramos a classe

$$\{p \mid \exists\, m\, (\Pi m = p \;\&\; p \notin m)\} = R$$

temos, por injetividade, uma contradição.

Claro, caso se adote a abordagem extensional, e, portanto, se identificasse proposições equivalentes, a contradição acima não poderia ser derivada. Russell, porém, mantém uma

[8]N.T.: “*range of significance*”.

[9]N.T.: *urelemente* são objetos que não são conjuntos — na medida em que não possuem elementos — e podem ser elementos de um conjunto. Em seu célebre trabalho de 1908, Zermelo incluiu os *urelemente* na primeira versão da sua teoria axiomática de conjuntos, ZFA (Zermelo-Fraenkel com átomos). Na medida em que os desenvolvimentos da lógica foram se consolidando, constatou-se que era possível desenvolver uma teoria axiomática de conjuntos sem o recurso aos átomos.

[10]N.T.: https://plato.stanford.edu/entries/type-theory/

abordagem intensional, enfatizando que proposições equivalentes muitas vezes podem ser bem diferentes. Assim, somos aparentemente forçados a rejeitar a hipótese de que proposições formam *um tipo* e exigir, portanto, que elas devem ter *vários tipos*, enquanto os produtos lógicos devem ter proposições de apenas um tipo enquanto fatores.

Esta foi eventualmente a base da teoria ramificada dos tipos, mas em 1903 Russell ainda considerava a sugestão como dura e artificial. Como uma nota de rodapé na página 527 evidencia, ele acreditava que o conjunto de todas as proposições é um contraexemplo ao teorema de Cantor.

2.4 Os matemáticos e as contradições: Hilbert e Zermelo

Zermelo descobriu independentemente o paradoxo de Russell em Göttingen (como testemunhado por Hilbert e Husserl) da seguinte forma: um conjunto M que compreende como elementos todos os seus subconjuntos é inconsistente. De fato, considere que o conjunto M_0 de todos os elementos de M que não são elementos de si mesmos (por exemplo, o conjunto vazio está em M_0). Este conjunto é um subconjunto de M e, portanto, com base em M, $M_0 \in M$. Se $M_0 \in M_0$, então M_0 não é membro de si mesmo. Portanto, $M_0 \notin M_0$ e uma vez que $M_0 \in M$, $M_0 \in M_0$: contradição.

Além disso, Hilbert notou em um trabalho não publicado (Kahle e Peckhaus, 2002) que contradições adicionais de natureza matemática podem surgir. A primeira deriva da suposição de que há um conjunto bem definido C que satisfaz as seguintes condições de fechamento: (i) o conjunto N dos números naturais é um elemento de C; (ii) $X^X \in C$, sempre que $X \in C$ (onde X^X é o conjunto de todas as funções de X para X); (iii) $\bigcup X \in C$ sempre que $X \subseteq C$. Então, por (iii) $\bigcup C$ = $U \in C$ e finalmente F = $U^U \in C$. Contudo, pela definição de união, $F \subseteq U$; portanto, haveria um mapa de U em F e uma contradição poderia ser derivada por diagonalização.

Ademais, como testemunhado por Hilbert (1905) em seu não publicado *Sommer Vorlesung* (Kahle 2004), ele descobriu uma versão funcional notável do paradoxo de Russell, que mais tarde se tornaria popular no contexto da lógica combinatória, do cálculo lambda e da teoria da recursão. O argumento é baseado na *autoaplicação funcional* e, portanto, na *autorreferência direta*.

A contradição é obtida assumindo que o universo contém tudo, isto é, as variáveis quantificam tanto sobre objetos como funções, e que há pelo menos dois objetos distintos. Então, introduz-se uma nova operação (aplicação universal em nosso sentido): xy é o resultado da aplicação de x em y. Dados dois objetos distintos 0 e 1, e partindo do pressuposto de que o universo é fechado sob definições arbitrárias por casos, existe um objeto f tal que $fx = 0$, se $xx \neq 0$, e $fx = 1$, se $xx = 0$. Então, escolhe-se $x = f$ (já que x quantifica sobre tudo) e facilmente se deriva uma contradição.

Os resultados da Escola de Hilbert não foram publicados porque contradições e paradoxos eram considerados como sintomas de crescimento e como dificuldades temporárias. O diagnóstico era que a lógica tradicional era insuficiente e a teoria de formação e conceitos precisava ser aperfeiçoada. Qualquer conceito C é dado em uma rede de conceitos (carta de Hilbert para Frege, 27 de Dezembro de 1899; Frege 1976, p. 79–80) e esta rede é determinada pelos axiomas. Somente a consistência dos axiomas que definem o conceito garante a legitimidade de C. Em resumo: paradoxos nos dizem que devemos desenvolver uma análise metamatemática das noções de prova e do método axiomático; sua importância é metodológica e epistemológica.

2.5 Por volta de 1905: dificuldades oriundas da definibilidade e do continuum

Como as reações do mundo matemático deixaram claro, quando os problemas elementares da teoria dos conjuntos foram trabalhados, por volta de 1905, os paradoxos foram crucialmente envolvidos. De fato, as novas contradições não afetaram apenas a concepção de conjunto e conceitos lógicos, mas também passaram a abranger a noção de *definibilidade* e sua relação com uma questão fundamental: a estrutura do continuum matemático e, em particular, se o continuum pode ser bem-ordenado e se a Hipótese do Continuum de Cantor é válida.

No Congresso de Heidelberg em 1904, Julius König tentou refutar a hipótese do continuum de Cantor. Devido a um erro descoberto por Zermelo, seu artigo foi imediatamente retirado; porém, no ano seguinte, König elaborou um novo argumento.

Considere os reais que são definíveis em um número finito de palavras. Eles formam uma sequência contável: $E_0, \ldots, E_n, \ldots$. Uma vez que o continuum é incontável, existem reais que não ocorrem na respectiva enumeração. Assumindo que o continuum é bem-ordenado, existe “o menor real E que não está na sequência $E_n \mid n \in \Omega$”; este real não está na sequência, mas a própria expressão “o menor real E que não está na sequência” *define E com um número finito de palavras*; então E ocorre em algum lugar na sequência: contradição!

König também observou que o argumento se estende para a segunda classe de números e um paradoxo semelhante poderia ser obtido se considerada a coleção **FOD** de ordinais contáveis finitamente definíveis. Neste caso, inspirando-se em Cantor, a solução de König é que a segunda classe de números de Cantor não é um conjunto em sentido próprio (uma totalidade completa). Para definir um conjunto, segundo König, deve-se fornecer não apenas uma regra para definir seus elementos, mas também um meio para os distinguir.

Uma contradição relacionada à de König havia sido publicada um pouco antes por Jules Richard, um matemático de um Liceu em Dijon. Usando uma enumeração de todas as permutações com repetições das vinte e seis letras do alfabeto francês, Richard percebeu que o conjunto E dos reais que pode ser definido por um número finito de palavras francesas é enumerável e, portanto, pode-se supor que haja uma enumeração $u_1, u_2, \ldots$ de todos esses números. Mas então podemos definir o seguinte real N: a parte inteira de N é 0, ao passo que o n-ésimo dígito decimal de N é $p+1$ se u_n tem o n-ésimo dígito decimal p diferente tanto de 8 como de 9; do contrário, o n-ésimo dígito decimal de N é 1. Por construção, N não ocorrerá em $u_1, u_2, \ldots$. Por outro lado, se consideramos que N é definido por uma coleção finita de letras, isso deve ocorrer em $u_1, u_2, \ldots$.

Ao contrário de König, Richard não se embasou na boa ordenação do continuum, e a solução proposta é interessante para o debate fundacional que está por vir. Ele apontou que a definição do número N se refere à totalidade de reais definíveis, aos quais o próprio N pertence; mas nenhum objeto deveria ser definível em termos de uma coleção que o contenha. Então parece que a definição é viciosamente circular, e isso a torna ilusória. Essa ideia logo se tornou a base da solução de Poincaré, e eventualmente também de Russell (sobre as coleções impredicativamente definidas, ver seção 3.1).

Então, por que definibilidade? A motivação ficou clara, por exemplo, em “Die Theorie der reellen Zahlen” de Bernstein (1905a), onde a hipótese do continuum de Cantor foi afirmada como estabelecida no positivo. Ele criticou a famosa “noção de Dirichlet” de função arbitrária e afirmou que é possível fornecer fundamentos ao continuum usando somente reais computáveis, reais que possuem uma “lei de formação” explícita (“*Bildungsgesetz*”).

Segundo ele, isso não é uma restrição, pois — ele afirma — há números reais computáveis não enumeráveis segundo sua proposta. Ele também afirma que é possível exibir o novo continuum computável em uma hierarquia (ou seja, uma sequência crescente $\subset$) $\{B_\alpha \mid \alpha < \aleph_1\}$ de subconjuntos de N^N, o qual é definido de tal modo que cada B_α é no máximo de cardinalidade $\aleph_1$ e, portanto, que a cardinalidade da união da sequência é no máximo $\aleph_1$. A ideia é começar com um domínio básico $B_0 \subset N^N$ de funções simples (por exemplo, aquelas que têm valores finitos), e então definir um novo domínio B_1 que estende B_0 com operações que são definidas a partir de elementos de B_0, e assim por diante.

Apesar de suas afirmações não serem justificadas, à luz do trabalho tardio de Gödel sobre construtibilidade, pode-se dizer que a intuição básica de Bernstein era sólida: o problema do continuum pode ser resolvido, conquanto se tenha uma noção apropriadamente geral de definibilidade (ou computabilidade, construtibilidade) e iterando-a ao longo dos ordinais.

Esta questão estava, em última análise, conectada com o problema de entender o continuum atomista. A concepção aritmetizada — no sentido de Dedekind ou Cantor, onde números reais são identificados com conjuntos adequados de racionais — mudou a atenção para as *sequências infinitas arbitrárias de números naturais*. Mas essa noção não era tão fácil de aceitar. De acordo com Bernstein (Bernstein 1905a, p. 449), uma sequência infinita (ou um conjunto infinito) deve ser dada por uma regra genuína. Mas *o que é uma regra*? Dado que devemos ser liberais (para não ter apenas classes especiais de reais, se formos muito restritivos), naturalmente somos levados a pensar em leis arbitrárias finitamente descritas, mudando a atenção para a *sintaxe das regras*. Porém, a menos que se considere a linguagem ordinária, nenhuma sintaxe desse tipo está disponível e isso resulta em indeterminação (este é o diagnóstico de Peano, veja Peano 1902–1906).

A necessidade de uma especificação dos conjuntos infinitos é crucial na discussão relacionada com o problema da boa ordenação. Também afeta o problema relativo à classificação de funções descontínuas e funções analiticamente representáveis de variáveis reais, abordadas pelos franceses semi-intuicionistas Borel, Baire e Lebesgue. Ao mesmo tempo, eles sustentavam que uma entidade matemática (como um conjunto infinito ou uma função) existe apenas na medida em que é “nomeável por um número finito de palavras”, contra as visões platonistas de Hadamard e Zermelo (veja Borel *et al.* 1905).

3. Paradoxos, predicatismo e a doutrina dos tipos: 1905–1913

À luz do debate fundacional, os anos seguintes foram ricos em trabalhos seminais: novos paradoxos foram descobertos (Berry, Grelling-Nelson), um velho paradoxo — o Paradoxo do Mentiroso — apareceu novamente, uma visão abrangente das contradições da lógica e da matemática foi magistralmente delineada na seção de abertura do artigo de Russell de 1908 e no anterior Russell 1906a, enquanto uma distinção conceitual entre dois tipos de paradoxos foi estabelecida no artigo de Peano de 1906. Além disso, as ideias básicas de predicatividade surgiram na discussão entre Poincaré, o principal matemático da época, e Russell. Ainda mais importante para a história da lógica matemática, avanços técnicos fundamentais para resolver os paradoxos e moldar os fundamentos da lógica e da matemática foram realizados: a teoria dos tipos (ramificados) de Russell e a axiomatização da teoria dos conjuntos de Zermelo.

3.1 Poincaré e Russell acerca das contradições

As ideias de Russell e Poincaré para resolver os paradoxos podem ser encontradas em diversos artigos publicados no período de 1905–1912:

1. o longo ensaio "Les mathématiques et la logique, na *Revue de Métaphysique et de Morale* (Poincaré 1905, 1906a), onde Poincaré critica veementemente os fundamentos axiomáticos e logicistas;

2. Russell e seu "On some difficulties in the theory of transfinite numbers and order types" (lido em 1905, publicado em 1907);

3. o contra-ataque de Poincaré (Poincaré, 1906b), e os subsequentes artigos de Russell na *Revue* (Russell 1906) e na *American Journal of Mathematics*(Russell 1908), e os últimos artigos de Poincaré (1909a, 1910, 1912).

Poincaré (1906b) tomou as contradições como base para defender uma concepção intuicionista e kantiana. De acordo com ele, a indução número-teórica[11] e o axioma da escolha constituem intuições independentes, verdadeiros juízos sintéticos *a priori*. Ele então argumenta contra a definição lógica de números naturais de Russell como aqueles números que pertencem a *todas as classes recorrentes* (aquelas classes que contêm 0 e estão fechadas sob um sucessor). Sua objeção é que a definição não é admissível pois refere-se essencialmente a uma totalidade à qual a classe a ser definida pertence — a definição é *impredicativa* — e, portanto, deve ser considerada como *circular*. Porém, segundo Poincaré, *objetos matemáticos não existem sem uma definição própria*, e uma definição própria deve ser *predicativa*, isto é, deve evitar *círculos viciosos*; Poincaré ampliou, de certo modo, o diagnóstico de Richard.

[11] N.T.: "*number-theoretic*".

As visões de Poincaré evoluíram ao longo dos anos e em debate com Russell. No período posterior, ele desenvolveu uma nova abordagem para a predicatividade, que, embora informalmente esboçada, é sugestiva de desenvolvimentos ulteriores em definibilidade e teoria da prova (veja Feferman 1964 e Heinzmann 1985). Ele já não insistia na circularidade viciosa da definição presente nas contradições; em vez disso, ele sustentava a tese de que *uma classificação predicativa é caracterizada por sua invariância*, ou seja, que ela não pode ser afetada pela introdução de novos elementos; em contraste, noções impredicativas estão sujeitas à constante modificação toda vez que novos elementos são introduzidos.

À luz da lógica contemporânea, Poincaré está insinuando alguma forma de *absolutismo ou invariância sob extensão* (que será precisada por Kreisel 1960 via teoria de modelos e teoria da recursão): suas ideias inspirarão a abordagem não ramificada dos fundamentos da análise predicativa.

Em sua última contribuição à *Acta Mathematica* (1909) e em sua quinta palestra em Göttingen (1910), ele também reafirmou o paradoxo de Richard — como um refinamento do teorema de Cantor — na seguinte forma: "não há enumeração definível de reais definíveis".

Enquanto matemáticos e o próprio Poincaré se concentraram nos problemas decorrentes das contradições na medida em que eles envolviam os fundamentos de noções matemáticas específicas (o continuum, os números naturais, a teoria dos números cardinais e ordinais), Russell atacou diretamente o *princípio da compreensão*, isto é, a hipótese de que certas funções proposicionais determinam uma classe (ver seção 2.3). Os paradoxos provam que uma função proposicional pode ser bem definida para cada argumento, mas ainda assim a coleção dos valores para os quais é definida não precisa ser uma classe. Então *o problema crucial se torna lógico*: fornecer um critério para selecionar aquelas funções proposicionais que dão origem a classes (entendidas como objetos bem definidos).

Sob a influência de Poincaré, Russell (1906, p. 634) aceitou o *princípio do círculo vicioso*, para o qual ele usou uma formulação segundo a terminologia e as noções da lógica formal estabelecidas por Peano:

> O que quer que componha uma variável aparente não deve ser um entre os valores possíveis dessa variável.

Em termos lógicos, não é permitido quantificar sobre uma dada classe X ao definir um elemento de X em si mesmo (veja o verbete Definitions)[12].

---

[12]N.T.: https://plato.stanford.edu/entries/definitions

Claro, o princípio do círculo vicioso não é uma teoria em si, mas uma condição que qualquer teoria adequada deve satisfazer. Russell (1906, 1907) propôs provisoriamente três abordagens alternativas: a teoria do zigue-zague, a teoria da limitação de tamanho e a teoria sem-classes. [13] A teoria do zigue-zague tenta capturar a ideia de que as funções certas devem ser 'simples', enquanto, de acordo com a segunda visão, 'predicativa' seria caracterizada por uma certa limitação do tamanho das classes que podem ser definidas predicativamente (por exemplo, a coleção **ON** de todos os ordinais é muito grande). Na teoria sem classes, as classes não são entidades independentes e qualquer coisa dita sobre elas deve ser considerada uma abreviação de uma declaração sobre seus membros e as funções proposicionais que os definem. Isso não está longe da ideia de Poincaré de que objetos matemáticos devem ser especificados em um número finito de palavras. No entanto, Russell desenvolveu o seu próprio aparato técnico — como o 'método de substituição' (Landini, 1998) e a eliminação contextual de descrições definidas (veja o verbete Bertrand Russell) — para implementar as suas ideias.

Em contraste com Poincaré, Russell (1906) não apenas não considerou o infinito atual como um componente essencial do problema fundacional, mas também enfatizou que as contradições surgem mesmo na ausência do infinito. Isso é claramente demonstrado pelo paradoxo do Mentiroso na forma "Estou mentindo". Até onde sabemos, é exatamente neste ponto do tempo que o (provavelmente) mais citado problema semântico na história da lógica recupera uma posição de destaque na análise lógica.

Em sua análise do paradoxo do Mentiroso, Russell assumiu que existe uma entidade verdadeira — a proposição — que é pressuposta por uma declaração genuína (por exemplo, quando digo que Sócrates é mortal, há um fato correspondente à minha afirmação e é esse fato que é chamado de "proposição"). O mesmo vale se a declaração for falsa, mas não no caso em que a declaração em si contém variáveis quantificadas.

O paradoxo é então resolvido ao interpretar o Mentiroso como "há uma proposição que eu afirmo e que é falsa"; portanto, a declaração contém uma quantificação (logo, uma variável aparente) sobre a coleção de todas as proposições, e não é uma proposição em um sentido próprio (Russell 1906, 642–644). Então, a conclusão é que o Mentiroso é falso porque não afirma uma proposição.

Considerações semelhantes se aplicam ao *paradoxo sugerido por Berry*, que é brevemente apresentado em Russell (1906) pela primeira vez em forma publicada, e tem o mérito

[13]N.T.: "*no-classes*".

de não ir além do domínio dos números finitos.

Considere os números naturais que são definíveis com menos de 21 sílabas: este conjunto é não-vazio e finito. Disso se segue que existem números que não são definíveis com menos de 21 sílabas. Considere o menor desses números: evidentemente, por definição, ele não é definível com menos de 21 sílabas.[14]

Por outro lado, tal número é definível com menos de 21 sílabas, uma vez que é determinado unicamente pela expressão “o menor número não definível com menos de 21 sílabas”, a qual tem menos de 21 sílabas.

Por razões de precisão histórica, devemos observar que Beppo Levi, que fez uma contribuição articulada ao debate sobre o axioma da escolha nas duas primeiras décadas de 1900, delineou uma antinomia que é essencialmente uma variante do paradoxo de Berry no contexto da discussão do paradoxo de Richard (veja Levi 1908). Para mais comentários sobre a abordagem de Levi aos paradoxos, veja Lolli 2007 e Bruni 2013.

3.2 Lógica matemática baseada na teoria dos tipos

A teoria russeliana dos tipos é amplamente conhecida e investigada na literatura: é de interesse atual e tem descendentes na lógica e suas aplicações (veja os verbetes Type Theory e Bertrand Russell)[15]. Foi desenvolvida pela primeira vez por Russell em seu memorável trabalho fundamental *Mathematical Logic as based on the theory of types*, de 1908.

A doutrina dos tipos é baseada na observação de que a quantificação universal — entendida como generalidade total, ou seja, quando se abrange ‘todo o universo’ — não faz sentido: quando afirmamos que $\forall x \phi(x)$ é verdadeiro, estamos apenas afirmando que a função $\phi(x)$ tem o valor ‘verdadeiro’ para todos os argumentos para os quais é *significativa*. O ponto essencial é que cada função proposicional tem um alcance de significância, quer dizer, *um tipo*, de modo que *a quantificação é legítima apenas sobre tipos*. Formalmente, cada variável deve ter um tipo pré-designado. Os paradoxos (ou falácias reflexivas) provam que certas coleções, como a totalidade de todas as proposições, de todas as classes e

[14]N.T.: A versão original do paradoxo de Berry utiliza 18 sílabas em vez de 21. Isso se deve ao fato de a expressão equivalente em língua inglesa (“*the least number not definable with less than 18 syllables*”) ter um menor número — 18 — de sílabas. Porém, o paradoxo pode ser obtido em outras línguas e com números de sílaba diferentes.

[15]N.T.: https://plato.stanford.edu/entries/type-theory/
https://plato.stanford.edu/entries/russell/

assim por diante, não podem ser tipos. Assim, podemos quantificar sobre a coleção de homens, mas não podemos afirmar corretamente que 'todas as proposições da forma $p \vee \neg p$ são verdadeiras'. Portanto, as entidades lógicas se dividem em tipos, e, em particular, *toda função proposicional deve ter um tipo superior ao de seus argumentos*. Além disso, à luz do princípio do círculo vicioso, a noção de ordem também deve ser introduzida. Nenhum objeto pode ser definido quantificando sobre uma totalidade que contém o próprio objeto como elemento; portanto, a ordem de cada função proposicional deve ser maior que a ordem das funções proposicionais sobre as quais ela quantifica.

A ideia principal é esclarecida em Russell (1908, p. 163–164) ao considerar como as proposições podem ser organizadas em uma hierarquia "ramificada" adequada, de acordo com sua ordem e de modo a satisfazer o princípio do círculo vicioso. Antes de tudo, existem proposições elementares, ou seja, aquelas que não contêm variáveis ligadas, enquanto o tipo mais baixo consiste de *indivíduos*. Indivíduos são entidades sem estrutura lógica e podem ser considerados como os sujeitos das proposições elementares. O segundo tipo lógico abrange as *proposições de primeira ordem*, isto é, aquelas cujos quantificadores (se houver) abrangem apenas indivíduos. A quantificação sobre proposições de primeira ordem dá origem a um novo tipo, composto exatamente por *proposições de segunda ordem*. De maneira geral, o tipo lógico $(n+1)$-ésimo inclui proposições de ordem n, que contêm quantificação apenas até a $(n-1)$-ésima ordem.

Uma vez que uma função proposicional pode ser obtida a partir de uma proposição 'tratando um ou mais de seus constituintes como variáveis', a hierarquia de tipos e ordens é naturalmente elevada a funções proposicionais, e faz sentido falar da ordem de uma função, sendo sua ordem aproximadamente a ordem do valor (isto é, uma proposição) que a função assume quando é aplicada a um argumento para o qual faz sentido. Assim, por exemplo, uma função que se aplica a indivíduos e toma proposições de primeira ordem como valores é de primeira ordem.

Seguindo a doutrina dos tipos, devemos substituir 'todas as proposições' por 'todas as proposições de ordem n' para um dado n. Assim, a sentença do paradoxo do Mentiroso torna-se "não é verdade que para todas as proposições p de ordem n, se eu afirmar p, p é verdadeira", que é uma proposição de ordem $n+1$. Então, o Mentiroso é simplesmente falso em vez de contraditório; e isso resolve o paradoxo. Argumentos semelhantes resolvem os outros paradoxos.

Nesta teoria, as *funções predicativas* de um argumento — ou seja, aquelas que têm como ordem o sucessor da ordem de seu argumento — desempenham um papel crucial. Por exemplo, uma função predicativa de uma variável individual deve ter ordem 1 (na terminologia atual, é elementarmente definível e quantifica *somente* sobre indivíduos). O *axioma da redutibilidade* (AR) afirma que *toda função proposicional é equivalente, para todos os seus valores, a uma função predicativa das mesmas variáveis*. Assim, por exemplo, podemos ter uma definição de uma propriedade $P(n)$ dos números naturais (considerados como indivíduos) que quantifica sobre, digamos, proposições de segunda ordem. Mas AR implica que existe uma função $P^*(n)$ que é satisfeita exatamente pelos mesmos números que $P(n)$ e é predicativa, isto é, envolve quantificação *apenas* sobre números.

Assim, de acordo com o axioma da redutibilidade, declarações sobre funções arbitrárias podem ser substituídas por declarações sobre funções predicativas; e funções predicativas desempenham o papel de classes, representantes canônicos de conceitos complexos arbitrários (por exemplo, entre as possíveis propriedades de diferentes ordens que têm a mesma extensão que $P(n)$, $P^*(n)$ representa canonicamente a classe de números que satisfazem $P(n)$).

Além do axioma do infinito, AR é uma ferramenta essencial para reconstruir a matemática clássica, mas é um princípio existencial forte, aparentemente em conflito com a ideia filosófica de que entidades lógicas e matemáticas devem ser geradas construtivamente de acordo com o princípio do círculo vicioso. No entanto, foi adotado na (primeira edição do) monumental *Principia Mathematica*, escrito em colaboração com A.N. Whitehead e publicado em 1910 (vol. 1), 1912 (vol. 2) e 1913 (vol. 3).

Curiosamente, a ideia básica subjacente à hierarquia ramificada dos tipos de Russell é um ingrediente crucial na posterior prova de consistência da hipótese do continuum de Gödel através de seu modelo interno L de conjuntos construtíveis. Além disso, como já observado em Gödel (1944), uma forma de AR se torna verdadeira em L no sentido de que, grosso modo, uma função proposicional arbitrária de números naturais é extensivamente equivalente a alguma função de ordem α, para algum ordinal contável α (veja o verbete Kurt Gödel)[16]. Outras aplicações importantes das hierarquias ramificadas têm sido desenvolvidas desde o final dos anos 50 em diferentes campos (desde teoria da recursão até teoria da prova; veja o verbete Type Theory)[17].

[16]N.T.: https://plato.stanford.edu/entries/goedel/
[17]N.T.: https://plato.stanford.edu/entries/type-theory/

3.3 Completando a figura

Já nos primeiros anos do século XX havia uma vasta literatura sobre paradoxos, a qual está longe de se esgotar na discussão anterior. Refinamentos e variações de interesse fundacional e lógico podem ser encontrados nos trabalhos de diversos autores, incluindo matemáticos proeminentes, entre eles Peano, Borel, Schönflies, Brouwer e Weyl. Algumas das propostas mais instigantes e originais são analisadas no restante desta seção.

3.3.1 Peano, Schönflies, Brouwer e Borel

A crítica de Peano (em *Additione a Super Theorema de Cantor-Bernstein*) ao paradoxo de Richard é principalmente conhecida por apontar que "*o exemplo de Richard pertence à linguística, não à matemática*". Essa afirmação abre a distinção entre antinomias matemáticas ou da teoria dos conjuntos e antinomias semânticas: o ponto fraco na definição de Richard é que, até certo ponto, ela é simbólica e formal, mas também faz uso da linguagem natural (*"lingua commune"*), que contém ideias bastante familiares, mas que, ainda assim, não são rigorosamente definidas e podem ser ambíguas (Peano 1906, p. 357–358). Por exemplo, não há um critério preciso para decidir se uma determinada expressão da linguagem natural representa uma regra que define unicamente um número.

Apesar disso, Peano elaborou uma solução formal. Ele tentou eliminar a vagueza e a referência a E, o conjunto dos números reais finitamente definíveis no intervalo unitário, fixando uma "numeração de Gödel" explícita: dado um número natural n, escreve-se n na base B, para B suficientemente grande (de modo a incluir o número de letras do alfabeto mais os sinais de pontuação). Assim, cada número é associado a uma sequência finita de símbolos na linguagem natural, e, em certos casos, essa sequência definirá um número, Val(n) $=$ "o número decimal determinado pela expressão codificada por n e interpretada de acordo com as regras da linguagem natural". Agora, para obter o paradoxo, é necessário provar que existe um único número N em $(0, 1)$ que satisfaz a condição dada por Richard (ver seção 2.5), mas essa condição, e, portanto, o próprio N, depende de Val, que é possivelmente obscuro e não pode ser definido exatamente segundo as regras da matemática (p. 352, p. 358). A conclusão é que tal número real não pode existir e que a definição de Richard é defeituosa, da mesma forma que a expressão "o maior número primo".

Por outro lado, Schönflies e Brouwer reagiram aos paradoxos opondo-se fortemente aos métodos axiomáticos e formais.

Schönflies (1906) manteve uma concepção genética e conteudista dos conjuntos. Segundo ele, os conjuntos são gerados e, uma vez formados, são conceitualmente invariantes. Quando um novo conjunto é construído, ele é adicionado àqueles usados em sua formação, sem alterar sua estrutura preexistente. Dessa forma, a auto pertinência não faz sentido, o conjunto universal não existe e o paradoxo de Russell desaparece. Sua visão pode ser considerada uma espécie de concepção iterativa dos conjuntos. Para Schönflies, as contradições surgem na lógica, e não na matemática, e são resultado da natureza escolástica da lógica. Ele via a lógica como uma influência prejudicial (*"unheilvollen Einflüss"*) sobre a matemática (*"Für den Cantorismus, aber gegen den Russellismus"* é o lema final de Schönflies 1911).[18]

A abordagem de Brouwer aos paradoxos também se baseava em uma concepção conteudista da matemática. Ela pode ser encontrada em sua tese de 1907, capítulo III (confira também van Dalen 1999, p. 105). Quanto à contradição de Russell, Brouwer observou que os princípios lógicos usuais só valem para palavras com conteúdo matemático, e não para sistemas linguísticos, como os de Peano ou Russell. Por exemplo, para decidir se uma classe pertence a uma função proposicional, a classe deve ser uma totalidade completa. As contradições mostram que há funções proposicionais que definem classes complementares (disjuntas) e que, no entanto, não satisfazem o *tertium non datur*.[19] Ideias semelhantes podem ser encontradas em seu artigo filosófico de 1908, "On the unreliability of tertium non datur" (veja van Dalen 1999).

Um resultado positivo do paradoxo de Richard no trabalho de Brouwer (1907, p. 149) é a noção de conjunto *enumeravelmente inacabado*, isto é, um conjunto no qual podemos determinar apenas subconjuntos enumeráveis de elementos, mas onde esses subconjuntos enumeráveis não esgotam o conjunto dado, de modo que sempre se pode produzir novos elementos do conjunto a partir de qualquer subconjunto enumerável dado. Exemplos típicos de conjuntos enumeravelmente inacabados são a totalidade dos ordinais contáveis, os pontos do continuum e, em particular, como pode ser demonstrado com o paradoxo de Richard, o conjunto de todos os pontos definíveis do continuum (1907, p. 150). Brouwer considera a contradição de Burali-Forti não como um paradoxo matemático, pois envolve uma estrutura lógica (a coleção de todos os ordinais), que não é um objeto matemático bem definido e não tem conteúdo matemático adequado. Do ponto de vista matemático, a contradição pode ser

[18]N.T.: Em tradução livre: "A favor do Cantorismo, mas contra o Russelianismo".
[19]N.T.: Expressão clássica em latim para a lei do terceiro excluído.

evitada negando que o maior tipo bem ordenado tenha um tipo de ordem sucessor [20] (p. 153; isso é análogo a Bernstein 1905b).

Entre os matemáticos franceses, o semi-intuicionista Borel (1908) introduziu a distinção entre conjuntos efetivamente enumeráveis e conjuntos denumeráveis. O paradoxo de Richard é então resolvido ao se observar que o conjunto de Richard, E, é certamente denumerável, pois só se pode determinar no máximo um conjunto denumerável de números reais por meios finitos. No entanto, E não é efetivamente enumerável, ou seja, não se pode produzir com um número finito de palavras um procedimento que atribua, de forma inequívoca, uma ordem (ou posição) a cada elemento do conjunto. Para que a enumeração de E fosse efetiva, seria necessário ter resolvido todos os problemas matemáticos possíveis de serem formulados, pois existem expressões que só se tornam definições de um número real com base na prova ou na solução de um determinado problema. Borel tem em mente um exemplo específico: considere a expressão "o único número transcendente cuja expansão decimal é obtida a partir da de π, substituindo-se sempre 7 por 8 e 8 por 7". Naturalmente, essa é uma boa definição apenas se tivermos demonstrado que tal número não é algébrico (veja Borel 1908, p. 446).

Borel, assim como Poincaré, adota um ponto de vista — definibilidade em um número finito de palavras — que é uma extensão do ponto de vista algébrico de Kronecker: somente objetos construtíveis em um número finito de etapas são objetos matemáticos propriamente ditos. No entanto, ao contrário de Poincaré, ele desconsidera o problema das definições predicativas: para ele, todos os paradoxos da teoria dos conjuntos derivam da tese de que a proposição *todo conjunto denumerável é efetivamente enumerável* (*"Tout ensemble dénombrable est effectivement énumérable"*) é evidente, quando, na verdade, é falsa.

3.3.2 Hessenberg, Grelling, Zermelo and Weyl

No campo dos fundamentos da matemática, três capítulos do *Bericht* (1906) de Gerhard Hessenberg sobre os fundamentos da teoria dos conjuntos são dedicados às questões fundacionais. Eles contêm ideias interessantes sobre a filosofia da matemática — as quais, infelizmente, não podem ser discutidas aqui em detalhes. Por exemplo, Hessenberg enfatiza a distinção entre definições da teoria dos conjuntos que fornecem critérios para decidir *efetivamente* a pertinência em um determinado conjunto e aquelas que não fornecem. Com

[20] N.T.: "*successor order type*".

relação à crítica construtiva de Kronecker ao continuum aritmetizado, Hessenberg argumenta que, embora cada número irracional determine uma fração infinita e cada fração infinita possua uma regra de formação ("Bildungsgesetz"), não é verdade que tal regra seja fornecida por *meios finitos explícitos*. Caso contrário, poderíamos derivar uma forma do *paradoxo da denotação finita*, ou seja, se as regras de formação coincidissem com as definíveis, elas seriam, no máximo, enumeráveis e, portanto, os números reais seriam enumeráveis, o que contradiz o teorema de Cantor.

Ao lidar com as contradições conjuntistas,[21] Hessenberg distingue o 'ultrafinito' (parágrafos 96–99) do 'transfinito': esta última noção é um atributo exclusivo dos conjuntos. Em contraste, as coleções envolvidas nos paradoxos (como o conjunto de Russell, o conjunto de todos os conjuntos, de todas as coisas e de todos os números ordinais) são ultrafinitas.

Quanto às soluções propostas para os paradoxos, Hessenberg se inspirou em uma ideia kantiana. Assim como nas ciências naturais surgem antinomias quando a natureza é concebida como um todo fechado, da mesma forma, a coleção **ON** e o conjunto de todos os conjuntos não podem ser concebidos como totalidades completas. Assim, a distinção entre 'ultrafinito' e 'transfinito' parece estar em conformidade com uma abordagem teórica próxima à doutrina da limitação de tamanho de Russell.

Próximo a esta inspiração filosófica, o influente artigo de Grelling e Nelson (1908) tenta unificar os paradoxos e isolar sua estrutura subjacente. O filósofo Leonard Nelson foi uma figura proeminente em Göttingen no início do século XX e contou com o forte apoio de Hilbert (veja Peckhaus 1990). O artigo faz parte de um projeto para desenvolver uma "kritische Mathematik" com um viés filosófico. Ele contém um novo paradoxo (atribuído a Grelling) de natureza semântica (veja também o verbete Self-Reference and Paradox)[22]:

> A cada palavra corresponde um conceito que a própria palavra designa e que se aplica ou não a ela; no primeiro caso, chamamos a palavra de *autológica*, no segundo caso, *heterológica*. Agora, a palavra 'heterológica' em si mesma é autológica ou heterológica. Se assumirmos que a palavra é autológica, então o conceito que ela designa se aplica, e, portanto, 'heterológica' é heterológica. Mas se a palavra for heterológica, o conceito designado não se aplica, então 'heterológica' não é heterológica.

[21] N.T.: "*set-theoretic*".

[22] N.T.: https://plato.stanford.edu/entries/self-reference/

A axiomatização da teoria dos conjuntos por Zermelo (1908) forneceu uma ferramenta eficaz para bloquear contradições (para maiores detalhes, veja os verbetes Set Theory, The Early Development of Set Theory e Alternative Axiomatic Set Theories).[23] As principais ideias da axiomatização podem ser resumidas da seguinte forma: (i) a compreensão ingênua é restrita a um axioma da separação, ou seja, um princípio que garante a existência de subconjuntos suficientes de um conjunto já dado de objetos (números, pontos, funções em um dado espaço); 'suficientes subconjuntos' aqui se refere a todos os subconjuntos especificáveis por meio de *condições definidas* envolvendo as noções primitivas (igualdade de conjuntos e pertinência) e satisfazendo as leis da lógica clássica; (ii) existem axiomas que garantem que as operações básicas de formação de unitário, união, pareamento e conjunto potência sejam bem definidas e que existam pelo menos um conjunto infinito e o conjunto vazio; (iii) assume-se a extensionalidade: dois conjuntos são iguais se e somente se possuem os mesmos elementos; (iv) postula-se o axioma da escolha, que permite selecionar um conjunto de escolhas de qualquer família de conjuntos disjuntos e não vazios.

De imediato, observa-se que a antinomia de Burali-Forti não pode ser derivada no sistema de Zermelo uma vez que a coleção de todos os tipos de ordem não existe como um conjunto, e o paradoxo de Russell simplesmente se torna o teorema de que não existe um conjunto universal.

No entanto, pode-se levantar pelo menos duas objeções contra essa teoria. Em primeiro lugar, a abordagem de Zermelo é *altamente impredicativa* e a impredicatividade era tomada como indispensável por ele (caso contrário, seria necessário rejeitar a matemática padrão). Por exemplo, Zermelo acreditava que isso se aplicava até mesmo à prova de Cauchy-Weierstrass do teorema fundamental da álgebra). Mas a impredicatividade torna a construção de um modelo ou de uma interpretação mais difícil e menos evidente. O segundo ponto é que Zermelo acreditava que o paradoxo da denotação finita (mencionado por Hessenberg) e o de Richard são bloqueados na teoria dos conjuntos, pois o axioma da separação deveria fornecer critérios claros para definir conjuntos. No entanto, isso não ocorre, pois a noção de propriedade definida (*definite Eigenschaft*) de Zermelo é dada informalmente e é, consequentemente, vaga.

Essa última questão foi abordada na palestra de "habilitação" de Weyl (1910), na qual

[23]N.T.: https://plato.stanford.edu/entries/set-theory/
https://plato.stanford.edu/entries/settheory-early/
https://plato.stanford.edu/entries/settheory-alternative/

ele tratou da seguinte questão geral: quando uma relação é explicitamente definível a partir de um conjunto de conceitos primitivos dados na matemática? Primeiramente, ele considera, como estudo de caso concreto, o problema de caracterizar os conceitos explicitamente definíveis da geometria plana elementar: esses conceitos podem ser gerados indutivamente por meio de cinco princípios básicos de definição a partir de dois conceitos primitivos adequadamente escolhidos (por exemplo, a identidade entre pontos $=$, e a relação ternária $E(a, b, c)$, 'a distância do ponto a ao ponto b é a mesma que a distância do ponto a ao ponto c').

Os cinco princípios de definição correspondem a uma *axiomatização finita do princípio da compreensão elementar* e implicam precisamente a existência daqueles conjuntos (relações) que são definíveis por fórmulas na linguagem elementar contendo dois símbolos de predicado para $=$, E. Mais explicitamente, exigem fechamento sob as operações lógicas de negação, conjunção, quantificação existencial e operações combinatórias adequadas de permutação e expansão.

À luz do exemplo geométrico, Weyl critica o conceito "definido por meio de um número finito de palavras" como impreciso e, muito antes de Fraenkel e Skolem, ele consegue tornar o princípio da separação preciso: ele substitui o conceito informal de propriedade definida de Zermelo pela noção de 'relação explicitamente definível a partir da igualdade extensional e da pertinência por meio de princípios lógicos elementares básicos' (ou seja, definível em primeira ordem).

Segundo Weyl, o paradoxo de Richard nos ensina a seguinte distinção: por um lado, só podemos caracterizar um número enumerável de subconjuntos de um dado conjunto por meio de definições explícitas; mas, por outro lado, novos objetos e conjuntos (possivelmente não contáveis) podem ser introduzidos pela aplicação das demais operações da teoria dos conjuntos, como o conjunto potência ou a união.

Weyl abordou o problema da geração de propriedades admissíveis sobre um dado domínio alguns anos depois, em *Das Kontinuum* (1918). Assim como em 1910, um conjunto dos conjuntos de números naturais que são definíveis via operações admissíveis (às quais agora também se adiciona uma forma de *iteração*) é enumerável. Pelo argumento de Cantor, não há uma relação que parametriza todos os subconjuntos dos números naturais (Weyl 1918, seção 5). Weyl aparentemente seguiu uma atitude relativista, segundo a qual a extensão do universo dos conjuntos e suas propriedades dependem das operações aceitas para construí-los (veja o verbete Hermann Weyl).[24]

[24]N.T.: https://plato.stanford.edu/entries/weyl/

Deve-se enfatizar que a atitude de Weyl em relação à antinomia de Grelling é totalmente negativa: ele a considera puro escolasticismo (Weyl 1918, seção 1): segundo ele, não há como atribuir significado a 'heterológico', de modo que, em última análise, tais problemas devem ser resolvidos recorrendo à filosofia.

É interessante observar que, à luz dos recentes desenvolvimentos, o veredito negativo de Weyl deve ser enfraquecido (ver seção 6).

4. Desenvolvimentos lógicos e paradoxos até 1930

No período até 1930, o problema dos paradoxos levou naturalmente à investigação dos cálculos lógicos, e foi, em grande medida, inserido nela (cujo subproduto final foi o livro-texto de Hilbert-Ackermann de 1928). Isso, por sua vez, abriu caminho para a simplificação da teoria dos tipos, para generalizações importantes da noção de conjunto e para uma formulação quase definitiva da teoria axiomática dos conjuntos (seguindo a abordagem de Zermelo, mas também o novo caminho aberto por Johann von Neumann). A ferramenta lógica básica é essencialmente a análise formal axiomática.

4.1 Teoria dos Conjuntos e paradoxos: conjuntos circulares e outras questões

Existem objetos circulares na teoria dos conjuntos? A visão de Zermelo sobre conjuntos, conforme axiomatizada em 1908, não excluía, por si só, a possibilidade de auto pertinência. O problema foi retomado por Mirimanoff (1917a, 1917b, 1920; veja também o verbete Zermelo's Axiomatization of Set Theory).[25] Uma vez permitidos os conjuntos circulares, é preciso *um fortalecimento da igualdade extensional por meio de uma relação de isomorfismo adequada* (bissimulação, na terminologia atual) que essencialmente corresponde ao isomorfismo das árvores que representam os conjuntos dados. O argumento de Russell sugere então uma distinção entre *conjuntos de primeiro tipo*, que não são isomorfos a nenhum de seus próprios elementos, e *conjuntos de segundo tipo*, que são isomorfos a pelo menos um de seus elementos. À luz dessa distinção, *a contradição de Russell evidencia que a coleção R de conjuntos do primeiro tipo não existe como um conjunto*. De fato, um conjunto do segundo tipo sempre contém um conjunto do segundo tipo; portanto, um conjunto de con-

[25] N.T.: https://plato.stanford.edu/entries/zermelo-set-theory/

juntos do primeiro tipo deve ser do primeiro tipo. Se R fosse um conjunto, deveria ser do primeiro tipo; mas então não poderia conter todos os conjuntos do primeiro tipo. Mirimanoff (1917a) introduziu então a distinção fundamental entre conjuntos *ordinários* (bem-fundados) e *extraordinários* (não bem-fundados): um conjunto X é ordinário se toda cadeia descendente $\in$ em X é finita; caso contrário, ele é extraordinário (existem cadeias descendentes $\in$ infinitas). Disso decorre que todos os conjuntos do segundo tipo são extraordinários, mas o inverso não é verdadeiro (por exemplo, considere o conjunto $E = \{e_1, E_1\}$, onde $E_1 = \{e_1, e_2, E_2\}$, $E_2 = \{e_1, e_2, E_3, E_3\}$, etc.).

Para a história dos paradoxos, é importante enfatizar que Mirimanoff (1917a) generalizou a antinomia de Burali-Forti, criando *o paradoxo dos conjuntos fundados*. Esse paradoxo prova que a coleção **WF** de conjuntos ordinários (em um dado conjunto de átomos) não é, ela mesma, um conjunto. De fato, seja **WF** o conjunto dos conjuntos fundados (= ordinários = bem-fundados); então, **WF** em si mesmo é fundado, pois, se **WF** $\ni x_0 \ni x_1 \ni x_2$ ..., então x_0 seria um membro não fundado de **WF**, o que é absurdo. Portanto, **WF** $\in$ **WF**, de modo que **WF** não é fundado, obtendo-se uma contradição (o mesmo paradoxo também aparecerá em Shen-Yuting 1953).

O trabalho de Mirimanoff também é importante para os fundamentos da teoria dos conjuntos. Ele introduziu a noção de classificação ordinal para conjuntos ordinários e percebeu que conjuntos ordinários podem ser organizados em uma hierarquia cumulativa, indexada por sua classificação. No entanto, a existência de uma estrutura cumulativa de conjuntos ordinários não é considerada como um motivo para excluir conjuntos extraordinários. Mirimanoff (p. 212–213) aponta explicitamente para o uso de conjuntos extraordinários para modelar *situações espelhadas*. Ele menciona o caso de um livro B cuja capa está decorada com uma imagem J representando duas crianças olhando para o mesmo livro, ou seja, para a imagem J_1 de B. Em J_1, pode-se perceber novamente as duas crianças e a imagem J_{11} do livro em perspectiva. Agora, J pode ser visto como um conjunto incluindo como elementos as duas crianças e_1 e e_2 e a imagem J_1 de J, que por sua vez se decompõe nas imagens de e_{11}, e_{22} de e_1 e e_2 e na imagem J_{11} de J_1, e assim por diante ad infinitum. Assim, J é isomorfo a um de seus elementos, ou seja, $J_1 : J$ pode ser considerado um conjunto do segundo tipo e, portanto, extraordinário. Esse exemplo não matemático é sugestivo de desenvolvimentos posteriores, como a teoria dos conjuntos não bem-fundados e suas aplicações recentes à semântica. Um conjunto E do segundo tipo também é assimilado a uma coleção impredicativa no sentido de Poincaré (Mirimanoff 1920, p. 34) devido à sua circularidade: de fato, E é dado por uma condição E = (y, z, ... , a, b, c, ...), onde y, z, ... dependem de E.

Por outro lado, em Mirimanoff (1917a) há um uso notável do paradoxo de Burali-Forti que sugere uma condição necessária para que algo seja um conjunto em termos de tamanho, ou seja, se uma coleção está em bijeção com o conjunto de todos os ordinais, então ela não existe como um conjunto. Em Mirimanoff (1917a, 1917b), também está presente a ideia do ordinal de von Neumann (von Neumann 1923, 1925) e uma forma do axioma da substituição está presente.

O sistema de von Neumann de 1925 trata de uma fundação axiomática alternativa da teoria dos conjuntos. Há dois tipos de objetos: objetos do tipo II (funções, correspondentes a classes) e objetos do tipo I (argumentos), ligados pela operação de aplicação de uma função a seus argumentos. Os dois domínios se sobrepõem parcialmente e existem objetos do tipo I–II, correspondentes a conjuntos (como funções que também podem ser argumentos). O axioma fundamental IV-2 estabelece então que um objeto a é uma classe própria (ou seja, não é do tipo I–II) se e somente se a totalidade de seus membros pode ser mapeada sobre a totalidade de todos os argumentos. A antinomia de Burali-Forti mostra que a classe **ON** de todos os ordinais não é um conjunto, o que implica, com o axioma IV-2, que existe uma aplicação de **ON** sobre o universo de todos os conjuntos, e, portanto, que *o universo dos conjuntos é bem ordenado*. Conceitualmente, o sistema resolve o problema de tornar precisa e aplicável a distinção de Cantor entre consistente e inconsistente (contra a crítica inicial de Hilbert); ele também mostra que a escolha global se torna um teorema sob uma visão adequada de conjuntos. Embora objetos circulares não possam existir no modelo hierárquico da teoria dos conjuntos de von Neumann, eles podem ser encontrados nas investigações de outros matemáticos e lógicos, como Finsler. Para Finsler, os paradoxos dependem de noções circulares, mas a circularidade não leva necessariamente a contradições. Em particular, ele considera que a noção de conjunto de Cantor é intrinsecamente circular: conjuntos dependem de conjuntos ou de coisas gerais que também dependem de conjuntos, e *o grafo de dependência associada* pode nos levar de volta a um círculo. Para o leitor contemporâneo, vale mencionar que uma intuição original de Finsler (1926b) foi o *uso da teoria dos grafos para representar estruturas circulares*. Com as setas empregadas para interpretar a pertinência, não é difícil imaginar um conjunto que tenha a si mesmo como único elemento e situações circulares mais complicadas (para maiores esclarecimentos sobre a teoria dos conjuntos de Finsler, veja Holmes 1996). Finsler (1926) aplica o paradoxo de Richard para produzir resultados metamatemáticos, em particular 'proposições formalmente indecidíveis'. No entanto, os argumentos de Finsler não são conclusivos e não podem ser considerados uma antecipação adequada dos teoremas da incompletude de Gödel (sobre o limite de suas ideias, veja a discussão em van Heijenoort 1963, 438–440); mas eles mostram que uma leitura atenta do paradoxo pode ter aplicações inesperadas.

4.2 Desenvolvimentos na teoria dos tipos e os paradoxos

No que diz respeito aos desenvolvimentos subsequentes em conexão com a lógica, o processo de refinamento das ferramentas lógicas continuou de forma constante em Göttingen por meio do trabalho de Hilbert e sua escola. Isso é especialmente evidente em suas anotações de aula não publicadas (por exemplo, aquelas do curso do semestre de inverno de 1917–1918, *Prinzipien der Mathematik*), que, em muitos aspectos, são próximas ao livro-texto de Hilbert-Ackermann de 1928 e contêm formulações da lógica de primeira e segunda ordem, juntamente com a teoria ramificada dos tipos e o axioma da redutibilidade. Os paradoxos são derivados ao permitir uma forma adequada de compreensão irrestrita; a suposição problemática está na admissibilidade de *predicados e proposições como objetos*, isto é, formalmente nas expressões $X(Y)$, $P(P)$. Variantes do paradoxo do Mentiroso tradicional e da antinomia de Berry são introduzidas. Curiosamente, Hilbert se mantém essencialmente na teoria dos tipos (ele não ministra aulas sobre o sistema de Zermelo); ele define a teoria ramificada dos tipos com o axioma da redutibilidade, demonstrando que certas partes da matemática podem ser desenvolvidas no sistema (sobre a influência de Russell, veja Mancosu 2003).

O fato de que a teoria dos tipos e o trabalho de Russell ocupavam um lugar central não apenas em Göttingen, sob Hilbert, é ainda atestado pelo trabalho de Chwistek e Ramsey, que tentaram revisar o *Principia Mathematica* (PM) a partir de perspectivas opostas. Ambos os autores rejeitaram a teoria ramificada dos tipos (TRT, por brevidade) e o axioma da redutibilidade. Seus trabalhos podem ser considerados um resultado típico do processo que levaria a versões simplificadas dos formalismos lógicos. No que diz respeito aos paradoxos, o principal problema é demonstrar que TRT não é necessária para resolvê-los.

A solução proposta por Chwistek foi motivada por uma concepção construtivista/predicativa. Sua posição em 1921 era de que o *Principia Mathematica* não era suficiente para evitar a antinomia clássica de Richard. Por outro lado, Chwistek propôs uma versão do paradoxo do Mentiroso que pode ser reconstruída na teoria simples dos tipos sem o axioma da redutibilidade, desde que seja permitido quantificar sobre todas as proposições. Chwistek adotava uma posição de cunho nominalista e tentou desenvolver uma teoria construtiva dos tipos para os fundamentos da análise (sua tentativa de fundamentar a matemática sem o axioma da redutibilidade foi chamada de “heroica” na introdução da segunda edição do *Principia Mathematica*, de 1925; veja Linsky 2004).

Ramsey (1926) introduziu a distinção — hoje padrão — entre *contradições lógicas* e *epistemológicas* (embora já houvesse indícios dessa distinção em Peano 1906; ver seção 3.3.1). Enquanto as contradições lógicas envolvem termos matemáticos ou lógicos, como classe e número, indicando que nossa lógica ou matemática é problemática, as contradições semânticas envolvem, além de termos puramente lógicos, noções como “pensamento”, “linguagem” e “simbolismo”, que, segundo Ramsey, são termos empíricos (não formais). Assim, essas contradições decorrem de ideias deficientes sobre pensamento ou linguagem e pertencem, propriamente, ao domínio da “epistemologia”.

Segundo Ramsey, as antinomias do primeiro grupo (como a de Russell ou a de Burali-Forti) podem ser evitadas ao se referir a um universo de objetos matemáticos estruturado em tipos de indivíduos, funções de indivíduos, funções de funções de indivíduos, e assim por diante. A quantificação sobre tipos arbitrários é legítima, e os tipos são fechados sob compreensão impredicativa, que é tida como necessária para a matemática. Os tipos são intrínsecos aos objetos lógicos e matemáticos, e os paradoxos lógicos são exatamente aqueles que exigem distinções de tipo para serem resolvidos (por exemplo, a auto pertinência é bloqueada para objetos na hierarquia tipo-teórica)[26]. Para resolver as antinomias semânticas (como o paradoxo do Mentiroso e a antinomia de Berry), Ramsey propõe distinguir várias noções de significado. À luz dos desenvolvimentos posteriores, é interessante notar que, para ele, *a semântica não é uma noção universal viável*: em particular, é impossível obter “uma relação de significado totalmente inclusiva para funções proposicionais. Seja qual for a relação adotada, ainda haverá uma maneira de construir um símbolo que tenha um significado não incluído em nossa relação. Os significados do significado formam uma totalidade ilegítima” (Ramsey 1926, p. 372). Isso leva a indicações para a solução do paradoxo de Grelling (veja seção 3.3.2). Seja R a relação de significado que liga um adjetivo f à função proposicional correspondente denotada por F (de modo que fRF seja válido). Na definição de “heterológico”, utilizamos a relação R: x é heterológico se e somente se existe uma função F tal que F não se aplique a x e xRF. Agora, existe uma função proposicional, digamos

[26]N.T.: “*type-theoretic*”.

H, que é o significado do adjetivo "heterológico". O ponto de Ramsey é que esse sentido de significado não pode ser o mesmo dado por R, e isso bloqueia a contradição quando aplicamos H a "heterológico" (ibid., p. 370). Assim, é necessária uma nova relação de significado dependente da relação R previamente fixada. Essas ideias antecipam as de Tarski (para uma análise das antinomias semânticas em um contexto ramificado, compare também a contribuição posterior de Church 1976, reconsiderada e criticada em Martino 2001).

5. Paradoxos: entre a metamatemática e fundamentos sem tipos (1930–1945)

Com os trabalhos de Gödel e Tarski, os argumentos paradoxais foram reformulados em termos de resultados de ponto fixo, enquanto a concepção semântica da verdade, a formalização da semântica e a aritmetização da sintaxe estabeleceram bases sólidas para investigações metamatemáticas sistemáticas. Ademais, houve um esforço para desenvolver novas grandes lógicas em reação à lógica do *Principia Mathematica*. Conceitualmente, as noções de função-em-intensão/operação (autoaplicável) e propriedade/predicado foram aceitas como primitivas, e o próprio mecanismo de definição/combinação de conceitos foi estudado. Essa linha de pensamento impulsionou a elaboração de métodos sintáticos dentro da lógica combinatória e o surgimento da teoria da recursão. O diagnóstico dos paradoxos foi ainda enriquecido por uma análise mais sutil das características puramente lógicas do raciocínio paradoxal: isso é especialmente verdade para a negação e o papel crucial das propriedades de contração e duplicação incorporadas às leis da implicação padrão. A lógica trivalente foi aplicada à compreensão ingênua.

5.1 Paradoxos e diagonalização

O papel heurístico dos paradoxos é testemunhado pelo próprio Gödel, quando intuitiva e explicitamente relaciona sua construção de sentenças formalmente indecidíveis a paradoxos epistemológicos ("a analogia com a antinomia de Richard salta aos olhos", Gödel 1931, van Heijenoort 1967, p. 599). No entanto, as construções autorreferenciais atingiram um grau adequado de rigor matemático e se tornaram ferramentas matemáticas genuínas somente quando técnicas número-teóricas não triviais foram aplicadas (veja o verbete Recursive Functions),[27] por exemplo, na análise da substituição sintática e na elaboração de

[27] N.T.: https://plato.stanford.edu/entries/recursive-functions/

modelos aritméticos de demonstrabilidade formal (o papel crucial da substituição na produção de contradições já havia sido notado por Russell, embora ele não tenha publicado a respeito; veja Pelham e Urquhart 1994). Conceitualmente, é claro pelas construções de Gödel que a autorreferência, por si só, é inofensiva se for entendida no *sentido indireto*: pode-se ter fórmulas $\varphi(x)$, expressando propriedades de seu próprio "nome" $\ulcorner\varphi\urcorner$, mas nenhuma circularidade problemática surge.

A construção de Gödel foi logo generalizada, como um lema geral da diagonalização, que se refere a propriedades definíveis arbitrárias. Isso pode ser encontrado em Carnap 1934b, p. 91, Carnap 1934a, p. 270, e em Rosser 1939, p. 57, Lema 1:

> Para toda fórmula $\psi(v)$ com apenas v livre, existe uma sentença φ tal que
>
> $$\varphi \leftrightarrow \psi(\ulcorner\varphi\urcorner)$$
>
> é demonstrável (veja o verbete Gödel's Incompleteness Theorems).[28]

De fato, o lema tornou-se a ferramenta padrão para produzir declarações autorreferenciais e para transformar *paradoxos semânticos em resultados de indefinibilidade e indecidibilidade (formal)* (veja o verbete Self-Reference and Paradox).[29] A álgebra subjacente às construções de Gödel só seria plenamente compreendida muito depois, na década de 70. Também é importante destacar que, poucos anos depois (1938), um análogo do lema da diagonalização (o chamado segundo teorema da recursão) foi descoberto por Kleene e logo se tornou uma ferramenta básica nos fundamentos da teoria da recursão e da teoria da computação.

5.2 Paradoxos e os fundamentos da semântica

Fica evidente pelo trabalho feito nas décadas de 20 mencionados acima que o problema de encontrar uma solução formal para os paradoxos semânticos, como o paradoxo do Mentiroso e o paradoxo de Richard, permaneceu essencialmente em aberto. As soluções baseadas na teoria dos tipos não foram desenvolvidas a ponto de fornecer uma análise formal sistemática das noções semânticas (como verdade ou definibilidade). Mas por que

[28]N.T.: https://plato.stanford.edu/entries/goedel-incompleteness/

[29]N.T.: https://plato.stanford.edu/entries/self-reference/

esse problema justificaria o estudo do ponto de vista lógico e matemático? Na realidade, noções semânticas — em particular, a noção de definibilidade — eram utilizadas, de forma mais ou menos explícita, em certas partes da teoria dos conjuntos (teoria descritiva dos conjuntos) e em áreas da teoria das funções mais inclinadas à teoria dos conjuntos, que eram cultivadas pela matemática polonesa nos anos 20. Ao mesmo tempo, o projeto de uma metodologia formal e de uma semântica científica foi desenvolvido por proeminentes filósofos e lógicos poloneses trabalhando em Lvov (atualmente Lviv) e Varsóvia (Lesniewski, Łukasiewicz, Chwistek; veja Wolenski 1995; veja também os verbetes Stanisław Leśniewski, Jan Łukasiewicz e Lvov-Warsaw School).[30] Por exemplo, Chwistek tentou desenvolver uma semântica elementar dentro de um programa fundacional nominalista, no qual conjuntos são identificados com funções proposicionais, a extensionalidade é rejeitada e a noção fundamental da semântica é a relação de substituição "H é o resultado da substituição de G por F em E" (Chwistek 1933, p. 374). Nesse ambiente estimulante, Tarski desenvolveu sua análise fundamental dos paradoxos semânticos, inicialmente entre 1929 e 1930, conforme relatado por Łukasiewicz à Sociedade Polonesa de Ciências em Varsóvia em 1931, e posteriormente detalhado em seu extenso trabalho de 1935 (veja os verbetes Alfred Tarski e Tarski's Truth Definitions).[31]

Primeiramente, a análise do paradoxo do Mentiroso começa com a formulação de um requisito formal que deve ser atendido na investigação semântica da verdade, ou seja, uma "definição materialmente adequada" (*sachlich zutreffende*) do termo "sentença verdadeira" (*wahre Aussage*). Isso equivale ao famoso esquema (T), que pode ser formulado de maneira simplificada como:

(T) x é uma sentença verdadeira se e somente se p.

onde p está para uma sentença e x é um nome de p (a ideia está de acordo com a intuição clássica da correspondência). O resultado que Tarski extrai do paradoxo do Mentiroso é que não pode existir uma linguagem interpretada que seja livre de contradições, obedeça às leis clássicas da lógica e satisfaça os requisitos (I)–(III), onde:

[30] N.T.: https://plato.stanford.edu/entries/lesniewski/
https://plato.stanford.edu/entries/lukasiewicz/
https://plato.stanford.edu/entries/lvov-warsaw/

[31] N.T.: https://plato.stanford.edu/entries/tarski/
https://plato.stanford.edu/entries/tarski-truth/

I. A linguagem possui nomes disponíveis para todas as suas sentenças;

II. Qualquer expressão obtida a partir de (T), substituindo p por uma sentença arbitrária da linguagem e x pelo nome correspondente de p, é aceita como verdadeira na linguagem;

III. Existem sentenças autorreferenciais, ou seja, é legítimo que uma sentença inclua seu próprio nome como um de seus constituintes (de modo que possamos estipular que o nome denota uma sentença $\alpha(\ldots c \ldots)$).

Diante desses obstáculos fundamentais, Tarski desenvolve uma *definição estrutural* das noções semânticas básicas, ou seja, uma definição que depende apenas da forma lógica de uma expressão e do fato de que expressões são definidas recursivamente. No entanto, essa abordagem só é viável para uma linguagem que seja estruturalmente descrita, ou seja, uma linguagem formalizada. Para tais linguagens, geralmente fechadas sob quantificação e contendo fórmulas com variáveis livres, Tarski elabora uma *noção apropriada de satisfação*, que permite introduzir as noções de *definibilidade, denotação, verdade e consequência lógica*. Pode-se então dar uma versão precisa e uma prova da condição de adequação (T) em uma metaciência cujos princípios compreendam: (i) axiomas lógicos gerais, (ii) axiomas especiais que dependam da teoria-objeto que consideramos e (iii) axiomas para lidar com propriedades fundamentais das noções estruturais, ou seja, *princípios de demonstração e definição por indução*. De posse deste aparato semântico, Tarski consegue resolver de forma negativa o problema da existência de (uma contraparte formal de) uma *linguagem universal*, isto é, uma linguagem na qual seja possível definir uma noção adequada de verdade para a própria linguagem. Embora a teoria dos tipos simples (na qual o tipo 0 corresponde ao tipo dos indivíduos, e o tipo $n + 1$ é a coleção de todas as classes de objetos do tipo n) com o axioma do infinito e a extensionalidade pareça uma boa candidata a uma metateoria geral, demonstra-se que, independentemente da definição escolhida para o termo "verdadeiro" dentro da teoria dos tipos, é possível deduzir, na própria teoria, a negação de alguma instância do esquema de adequação (T). Na demonstração desse teorema, Tarski aplica aritmetização e diagonalização, seguindo assim o padrão gödeliano. No lado positivo, o conceito de verdade pode ser definido de maneira adequada para qualquer linguagem formalizada *L* dentro de uma linguagem (a chamada metalinguagem), conquanto esta seja de ordem superior a *L*. Além disso, a semântica de Tarski explicita a observação de Gödel (1931, nota 48) de que “a verdadeira razão da incompletude é que a formação de tipos cada vez mais elevados pode ser continuada no transfinito [...], enquanto em qualquer sistema formal, no máximo, uma quantidade enumerável deles está disponível.” As noções semânticas tarskianas desempenham, portanto, o papel dos tipos superiores mencionados por Gödel. Resumindo, o resultado do trabalho de Tarski é que as noções semânticas são eliminadas em favor das noções (extensionais) de tipo ou conjunto, e uma explicação teórica para os paradoxos semânticos é finalmente alcançada.

5.3 “A inconsistência de certas lógicas formais”

Nas décadas de 1920 e início de 1930, a visão ortodoxa da lógica entre os lógicos matemáticos era, em linhas gerais, de caráter tipo-teórico ou conjuntista. No entanto, houve um esforço para desenvolver novas lógicas como substitutas da lógica do *Principia Mathematica*. Esses sistemas surgiram tanto como tentativas de recuperar a simplicidade da abordagem livre de tipos, derivada do chamado princípio ingênuo da compreensão, quanto para atender necessidades metamatemáticas, como o esclarecimento de conceitos fundamentais subjacentes às noções de “sistema formal”, “formalismo”, “regra”, etc. Em particular, Church e Curry propuseram teorias que (i) assumem como primitivas as noções de função em intensão autoaplicável (operação), e (ii) enfatizam o próprio mecanismo de definição/combinação de conceitos. Ao observar de perto o desenvolvimento desses sistemas, percebe-se que *construções paradoxais tornaram-se ferramentas essenciais para definir objetos e demonstrar fatos matemáticos e lógicos não triviais*. Seguindo ideias de Schönfinkel e visando uma análise matemática do processo de substituição, a tese de Curry de 1930 introduziu uma linguagem formal baseada em operadores gerais básicos, os chamados combinadores B (composição), C (permutação), W (duplicação), K (cancelamento), Q (igualdade), além de constantes lógicas como o quantificador universal e a implicação. Expressões são então geradas indutivamente por aplicação a partir de constantes; intuitivamente, um termo M representa uma função, e o termo aplicativo MN (onde a justaposição desempenha o papel da aplicação e os parênteses são associados à esquerda) denota o valor do termo obtido ao substituir a primeira variável de M por N. A autoaplicação MM é permitida, e essa característica indica que *os objetos da lógica combinatória não podem simplesmente ser interpretados como funções na teoria dos conjuntos*. O sistema formal consiste em equações padrão sobre combinadores (por exemplo, $Bxyz = x(yz)$, $Wxy = xyy$ ou $Cxyz = xzy$), regras para a igualdade e constantes lógicas; seu objetivo principal é derivar igualdades

$X = Y$ e fazer afirmações da forma $\vdash X$ (ou seja, X é demonstrável). A lógica combinatória é uma teoria que analisa os modos de combinação de objetos formais, a substituição e as noções de proposição e função proposicional (para uma introdução mais aprofundada às variantes do formalismo e um panorama das propriedades dos cálculos relacionados, veja o verbete Combinatory Logic).[32] Para Curry, a raiz dos paradoxos encontra-se na suposição de que *combinações de conceitos são sempre proposições*. A noção de proposição torna-se um *conceito teórico* cuja validade é decidida pela própria teoria. Tipos não são atribuídos às expressões do sistema formal desde o início, mas são inferidos pelo próprio sistema, que possui uma natureza dual: ele pode derivar identidades, mas também verdades. Em particular, se $\vdash MN$ é derivado, isso pode ser interpretado como "N é de tipo M" ou "N é um elemento de M". Essas ideias antecipam desenvolvimentos fundamentais, como a chamada *interpretação fórmulas como tipos* (veja Howard 1968).

O formalismo de Church — originalmente introduzido em Church 1932 e Church 1933 como um conjunto de postulados para a fundamentação da lógica formal — inclui regras de conversão (isto é, regras computacionais), que permitem a substituição de termos por outros intensionalmente equivalentes, e regras para afirmar certos termos como "verdadeiros". A sintaxe proporciona um sistema geral de notação para funções, baseado em uma linguagem aplicativa na qual há uma única categoria básica de termos (fórmulas bem formadas, segundo sua terminologia). Alguns termos são formalmente demonstráveis (ou afirmáveis) e são classificados como verdadeiros. Os termos são definidos indutivamente a partir de um conjunto de constantes básicas e variáveis por meio de aplicação e do respectivo *operador de abstração lambda*: se M é um termo contendo a variável x, $\lambda x.\ M$ é um termo, nomeando a função definida por M. As constantes básicas designam operações lógicas: (uma forma restrita de) implicação formal, quantificador existencial, conjunção, negação, operação de descrição e abstração generalizada (ou seja, se F é equivalência lógica formal, $A(F, M)$ é "o que M tem em comum com qualquer N formalmente equivalente a M"). Verifica-se que a lógica de Church pode interpretar a teoria ingênua das classes, evidenciando que o sistema é suspeitamente forte e expressivo (a força e expressividade são herdadas pelo formalismo derivado do sistema de Church; para maiores detalhes, veja o verbete The Lambda Calculus).[33] A esperança de Church era que as contradições pudessem ser evitadas garantindo-se a possibilidade de que uma função proposicional fosse *indefinida* para alguns argumentos.

[32]N.T.: https://plato.stanford.edu/entries/logic-combinatory/

[33]N.T.: https://plato.stanford.edu/entries/lambda-calculus/

No entanto, as teorias de Curry e Church foram quase imediatamente demonstradas inconsistentes em 1934, por Kleene e Rosser, que (essencialmente) provaram uma versão do paradoxo de Richard (ambos os sistemas *podem enumerar demonstravelmente* suas próprias funções número-teóricas totalmente definíveis e demonstráveis). Esse resultado foi motivado pelo próprio Church em 1934, quando ele utilizou o paradoxo de Richard para demonstrar uma espécie de teorema de incompletude (no que diz respeito às afirmações que expressam a totalidade de funções número-teóricas).

A razão das inconsistências foi finalmente esclarecida no ensaio de Curry de 1941. Nele, Curry distingue duas noções básicas de completude: um sistema S é *dedutivamente completo* se, sempre que ele demonstra uma proposição B a partir da hipótese A, então ele também deriva a implicação $A \to B$ (teorema da dedução ou regra de introdução da implicação); S é *combinatoriamente completo* (Curry 1941, p. 455) se, sempre que M for um termo do sistema possivelmente contendo um x indeterminado, há um termo (o $\lambda x.M$ de Church) que nomeia a função de x definida por M. Curry então observa que o paradoxo de Kleene-Rosser surge porque os sistemas de Church e Curry satisfazem ambos os tipos de completude, demonstrando assim que *essas duas propriedades são incompatíveis*. Na parte mais técnica do artigo, Curry axiomatiza cuidadosamente os principais elementos explorados por Kleene e Rosser e realiza um trabalho substancial tanto no aspecto lógico quanto no aspecto matemático (por exemplo, ao desenvolver uma parte da aritmética recursiva e ao definir a existência de um enumerador, um termo T tal que, se α é o número de Gödel de um termo fechado M e Z_α é o termo que representa formalmente α, então $TZ_\alpha = M$ é demonstrável em S, etc.). A demonstração de Curry sobre a inconsistência dos sistemas combinatórios é insatisfatória porque depende fortemente de um desvio pela teoria dos números e pela gödelização, o que, na verdade, é desnecessário, como o próprio Curry logo percebeu e apresentou em um artigo com o mesmo título do de Kleene e Rosser, "em deferência aos descobridores originais da contradição" (Curry 1942). O principal resultado desse trabalho é o seguinte teorema (o paradoxo de Curry; veja o verbete Curry's Paradox):[34]

1. assuma que temos um sistema combinatoriamente completo, ou seja, essencialmente um sistema de lógica combinatória que contém as propriedades padrão de igualdade e os axiomas básicos que garantem a definibilidade do lambda de Church, juntamente com os axiomas correspondentes que definem os combinadores B, C, I, W;

[34] N.T.: https://plato.stanford.edu/entries/curry-paradox/

2. assuma também que o sistema contém um operador de implicação $\supset$ que satisfaz, para termos arbitrários M e N

$$\vdash M \supset M$$
$$\vdash M \supset (M \supset N) \Rightarrow \vdash M \supset N$$
$$\vdash M \text{ e} \vdash M \supset N \Rightarrow \vdash N.$$

Então $S \vdash M$, para todo termo M.

Para a demonstração, basta encontrar, para um dado termo B, um termo A tal que A = $A \supset B$. Curry observa que é possível construir a prova de duas maneiras. Por autorreferência direta, podemos escolher: $A = HH$, onde $H = \lambda Y(N(YY))$ e $N = \lambda X(X \supset B)$. Por outro lado, pode-se aplicar autorreferência indireta e explorar a maquinaria de Curry (1941) e Kleene-Rosser: utilizando um enumerador T, define-se $U = \lambda X(TXX \supset B)$ e $A = UZ_u$, sendo u o número de Gödel de U. Curry sugere que essas duas abordagens são análogas, respectivamente, ao paradoxo de Russell e ao paradoxo do Mentiroso. É interessante notar que esses dois métodos correspondem às ferramentas que hoje são conhecidas como o *primeiro* e o *segundo teorema do ponto-fixo* da lógica combinatória e do cálculo lambda (Barendregt 1984, pp. 131 e 143; para maiores detalhes sobre a distinção entre o primeiro e o segundo teorema da recursão na teoria clássica da recursão, veja o verbete Recursive Functions).[35]

A análise de Curry sobre as soluções do paradoxo nos levaria ao campo da teoria da funcionalidade, à história da lógica combinatória e à teoria da demonstração. Aqui, basta lembrar que, segundo ele, um possível remédio seria formular dentro do próprio sistema a própria noção de proposição, e uma maneira de evitar as contradições levaria a uma hierarquia de proposições canônicas (ou a uma teoria de níveis de implicação, já esboçada por Church). Ideias relacionadas têm sido desenvolvidas desde os anos 1970 por Scott (1975), Aczel (1980), Flagg e Myhill (1987), entre outros.

5.4 Criticando a implicação padrão e a negação

Nos anos de 1930, surgiu uma rota alternativa para resolver as antinomias. Essa abordagem explora o combinador de duplicação (contração) W, que satisfaz $Wfx = fxx$; se

[35]N.T.: https://plato.stanford.edu/entries/recursive-functions/

N representa a negação e $Babc = a(bc)$, então $W(BN)W(BN))$ é um ponto fixo da negação, bem como um análogo funcional da classe de Russell. No nível lógico, atribuir um tipo a W leva à inferência essencial aplicada na derivação do paradoxo de Curry, ou seja, a regra de contração: $A \to (A \to B) \Rightarrow (A \to B)$. O papel da contração foi notado por Fitch (1936), que observou que, para derivar o paradoxo de Russell, considera-se uma função de duas variáveis, depois diagonaliza-se e trata-se tal objeto como uma nova função proposicional unária. Mas esse passo só funciona se W for aceito. Fitch então propôs uma lógica "não-contrativa", mas seu artigo não passou de um breve esboço de um fragmento da lógica clássica com predicados. Foi apenas em meados dos anos 1980 que as lógicas sem contração passaram a ser usadas sistematicamente na teoria da demonstração e na ciência da computação teórica (para maiores detalhes, veja o verbete Linear Logic).[36]

Em 1942, Fitch propôs uma nova abordagem para o problema de encontrar sistemas consistentes de lógica combinatória, que foram progressivamente expandidos e refinados ao longo de muitos anos (até 1980). O método de Fitch para evitar paradoxos consiste na *construção de modelos sintáticos adequados, dotados de noções autorreferenciais de classe, pertinência e verdade*. Verdade e pertinência são gerados indutivamente por regras iterativas que correspondem a condições naturais de fechamento lógico e podem ser formalizadas por meio de cláusulas *positivas* (isto é, sem negação ou implicação). Isso implica que o processo de geração é cumulativo e se torna saturado em um certo ponto, resultando em interpretações consistentes e não triviais para verdade e pertinência. Matematicamente, uma coleção de cláusulas positivas sempre dá origem a um operador, G, que mapeia conjuntos de expressões em conjuntos de expressões e preserva a relação de inclusão (isto é, é monótono); os conjuntos saturados correspondem então a pontos fixos do operador monótono G (ou seja, a conjuntos X que satisfazem $G(X) = X$), os quais existem de acordo com um teorema clássico sobre reticulados completos (veja Birkhoff 1967).

No início da década de 40, Fitch explorou um sistema combinatório K puramente positivo (sem negação), com o objetivo explícito de definir uma espécie de *sistema formal universal*, no qual *qualquer sistema lógico* pudesse ser representado. Posteriormente, ele conseguiu fortalecer sua abordagem para incluir formas de negação e implicação, à medida que forneceu uma *geração simultânea de verdade e falsidade*, o que equivale, na prática, a conceber a verdade como um *predicado parcial*. A abordagem de Fitch é radicalmente intensional: classes são sempre classes de expressões M em alguma linguagem (por exemplo,

[36]N.T.: https://plato.stanford.edu/entries/logic-linear/

a lógica básica) e identificadas com atributos, enquanto a pertinência é essencialmente reduzida à verdade no sentido de K. Assim, a afirmação $M \in T$ significa essencialmente que M realmente se enquadra na propriedade especificada (ou expressa) por T (veja também o verbete Combinatory Logic).[37] Sistemas lógicos similares também foram propostos e demonstrados como consistentes através de métodos prova-teóricos por Schütte no início dos anos 1950 (para um tratamento abrangente, veja Schütte 1960).

Até certo ponto, as ideias de Fitch podem ser vistas como a introdução da visão de que os predicados básicos de verdade e pertinência devem ser *parciais* ou, se preferir, *trivalorados*. Bochvar (1937) delineou uma proposta baseada na introdução de uma lógica trivalorada, na qual, além dos valores de verdade padrão T (verdadeiro) e F (falso), há um terceiro valor N, interpretado como "sem significado". Sua análise lógica leva à conclusão de que os paradoxos envolvem proposições desprovidas de significado. Uma característica marcante do formalismo de Bochvar é a distinção entre dois tipos de conectivos, correspondendo aproximadamente a dois modos diferentes de asserção. Uma afirmação A, por si só, assume exatamente um dos valores prescritos (verdadeiro, falso ou sem significado); mas as operações lógicas *internas* trivaloradas também possuem operações lógicas *externas*, que correspondem a afirmações em um nível meta e permitem o uso da lógica clássica para lidar com proposições não clássicas. Formalmente, Bochvar descreveu tabelas de verdade trivaloradas para os principais conectivos proposicionais internos: $\&$ (conjunção), $\sim$ (negação), $\vee$ (disjunção), $\rightarrow$ (implicação) e $\leftrightarrow$ (equivalência lógica). Os valores de verdade de $\sim A$, $A \mathbin{\&} B$, etc., coincidem com seus valores clássicos se A, B assumirem valores clássicos; porém, eles são sem significado (assumem valor N) se pelo menos um dentre A, B tiver valor N (estriticidade). Nenhuma fórmula construída com os conectivos padrão pode ser válida (ou tautológica, isto é, verdadeira em todas as atribuições possíveis), pois $A \rightarrow A$ possui valor N se A não tem significado. No entanto, Bochvar introduziu conectivos que permitem a formulação de *afirmações metateóricas*, como $\vdash A$, $\neg A$, $\downarrow A$, interpretadas, respectivamente, como "A é verdadeira", "A é falsa" e "A é sem significado". O valor de $\vdash A$($\neg A$) é verdadeiro se A for verdadeiro (falso) e falso (verdadeiro) caso contrário. Bochvar descreveu uma versão expandida do cálculo lógico sem tipos de Hilbert–Ackermann (1928) e, para eliminar os paradoxos, restringiu a substituição e, consequentemente, o esquema de compreensão da forma

[37] N.T.: https://plato.stanford.edu/entries/logic-combinatory/

$\exists F \forall x(F(x) \leftrightarrow \varphi)$ onde φ é qualquer fórmula na qual F não ocorre livremente e que pode ter ou não ter x como variável livre.

para condições φ somente com operações lógicas *internas*. Aparentemente, a solução de Bochvar não é simplesmente uma *solução de lacuna* na qual a lógica é enfraquecida. Em vez disso, ele formaliza a distinção entre *nível objeto* e *nível meta* dentro da própria lógica. Isso torna sua teoria bastante expressiva (por exemplo, ela consegue lidar com a própria noção de "sem sentido", "sem significado").

5.5 Processos Não Terminativos, Ciclos e Ambiguidade Típica

O artigo de Behmann (1931), apresentado em 1929, localiza a origem dos paradoxos na maquinaria definicional. O artigo começa com uma análise da contradição de Russell na forma da aplicação de predicados. Nele, Behmann observa que, uma vez que o predicado F de Russell é definido pela estipulação

$$F(\varphi) =_{df} \neg\varphi(\varphi),$$

deve ser possível eliminar F de qualquer argumento que o envolva. No entanto, se tentarmos substituir F por seu definiens, obtemos $F(F) \equiv \neg F(F)$, ficando assim presos em um regresso infinito uma vez que não há uma expressão livre de F que possa substituir $F(F)$. Dessa forma, a contradição é atribuída a um erro na teoria das definições, a saber, o uso de definições que dão origem a uma cadeia infinita de substituições sem convergir para um resultado. A proposta técnica de Behmann consiste em uma lógica reformada sem tipos, mas com um operador adicional ! que, quando aplicado a um predicado χ, seleciona exatamente aqueles argumentos x para os quais χ se aplica de maneira significativa. Por exemplo, o silogismo Barbara, geralmente expresso na forma

$$(\forall x)(A(x) \to B(x)) \,\&\, (\forall x)(B(x) \to C(x)) \to (\forall x)(A(x) \to C(x)),$$

é corrigido para

$$(\forall x)(A(x) \to B(x)) \,\&\, (\forall x)(B(x) \to C(x)) \to (\forall x)(C(x))! \to (A(x) \to C(x))$$

onde o domínio do quantificador final é restrito apenas àqueles elementos para os quais a proposição faz sentido. Behmann não desenvolveu uma teoria sistemática até muito mais tarde, e a interpretação precisa de seu operador especial ! permanece incerta. No entanto, seu trabalho inspirou as pesquisas de Aczel e Feferman (1980).

Lewis e Langford (1932) chegaram a conclusões que não são muito diferentes das de Behmann. Segundo eles, os paradoxos demonstram que certas expressões *não expressam proposições*. Eles adotam a notação p : α para indicar que p é um nome cujo significado é a proposição α (de modo que p e "α" denotam a mesma entidade e podem ser substituídos um pelo outro). Tipicamente, o paradoxo do Mentiroso assume a forma "$p : p$ é falso", mas também podemos imaginar situações autorreferenciais mais complexas, como:

$(p1)$	$p2$ é falso;
$(p2)$	$p1$ é falso.

Nesse caso, não há uma contradição, mas ficamos presos em um regresso vicioso (p. 440), e, portanto, *nenhuma proposição* surge. Em geral, pode-se construir ciclos arbitrariamente complicados e verificar que eles podem levar tanto a contradições quanto a regressos infinitos. Mas, em ambos os casos, a expressão falha em convergir para uma proposição definida.

Mesmo depois que as lógicas desenvolvidas por Russell, Zermelo e Tarski forneceram os meios teóricos para lidar com as dificuldades relacionadas às noções de classe, conjunto, verdade e definibilidade, os paradoxos continuaram a ser um problema relevante. Isso se deve provavelmente ao interesse persistente por paradigmas formais alternativos, às características e axiomas controversos do *Principia Mathematica*, bem como ao lugar problemático que a autorreferência ocupa na lógica matemática. Nesse contexto, vale mencionar o artigo de Quine de 1937 sobre o sistema NF (para maiores detalhes, veja o verbete Quine's New Foundations), [38] inspirado na noção russelliana de *ambiguidade típica*, isto é, no dispositivo sistemático de suprimir os índices de ordem das funções proposicionais e de seus argumentos, permitindo que sejam restaurados à vontade, quando necessário, conforme a disciplina da teoria dos tipos (ver seção 3.2). A ideia é restringir o esquema de compreensão ingênuo apenas às instâncias *estratificadas*, onde em geral uma fórmula φ é estratificada se e somente se é possível atribuir um número natural (um tipo, por brevidade) a cada ocorrência

[38] N.T.: https://plato.stanford.edu/entries/quine-nf/

de termo de modo que a fórmula resultante seja bem formada no sentido da teoria dos tipos. Por exemplo, se $t \in s$ é uma sub-fórmula de φ, então o tipo de s deve ser exatamente um a mais do que o tipo de t, e assim por diante. Claramente, a estratificação bloqueia a formação de conjuntos quando fórmulas como $x \in x$, $\neg x \in x$ estão presentes. Além disso, no sistema NF, o conjunto universo existe. O problema da consistência para NF continua em aberto (embora existam resultados parciais sobre fragmentos do sistema que impõem restrições na estratificação ou na extensionalidade). Notavelmente, NF refuta o axioma da escolha, conforme um teorema clássico de Specker. Outro resultado clássico de Specker estabelece a existência de um modelo de NF em uma versão apropriada da teoria dos tipos simples, na qual há um equivalente formal da ambiguidade típica. Os paradoxos não estão tão distantes de NF. Em 1942, Rosser e Lyndon publicaram independentemente uma versão da antinomia de Burali-Forti dentro de uma extensão aparentemente natural de NF (chamada ML) obtida pela adição de "classes últimas". O sistema ML foi concebido com o objetivo de evitar certas fraquezas de NF (por exemplo, no que diz respeito à indução número-teórica). No entanto, o resultado de Lyndon-Rosser revelou, mais uma vez, a presença inesperada de um paradoxo na teoria dos conjuntos e nos fundamentos da lógica matemática.

6. Um olhar sobre as investigações atuais

6.1 Dos paradoxos aos teoremas

Como observado há muitos anos por Kreisel e muito apropriadamente lembrado por Dean (2020, p. 541), a questão correta a ser abordada não é *como se livrar dos paradoxos* ou resolvê-los: em vez disso, o problema *frutífero* é *como extrair algo deles*. De fato, paradoxos podem ser convertidos em simples teoremas da indecidibilidade/indefinibilidade, e isso é consequência de uma metodologia sistemática: com o trabalho de Gödel e Tarski, argumentos paradoxais são reformulados em resultados de pontos fixos, enquanto a concepção semântica da verdade dá origem *a uma formalização da própria semântica*, fornecendo bases sólidas para investigações metamatemáticas sistemáticas, como pode ser observado em várias contribuições iniciais de diversos lógicos, por exemplo, Kreisel (1950) e Wang (1955). Por exemplo, a aritmetização da semântica permite um refinamento do teorema da completude, conhecido como *teorema da completude aritmetizado* (ART): toda teoria recursiva consistente possui um modelo no qual os símbolos de função são substituídos por funções primitivas recursivas e os símbolos de predicado são substituídos por predicados definíveis *com apenas dois quantificadores* em uma versão da teoria formal dos números (cf. Hilbert e Bernays 1939, p. 293, e Feferman 1960). Ao adicionar uma sentença aritmética Con(S), que expressa a consistência de uma teoria dos conjuntos S como um novo axioma da teoria elementar dos números, pode-se provar no sistema resultante *traduções aritméticas* de todos os teoremas de S. Um resultado dessa formalização metamatemática é uma unificação *de fato* entre os paradoxos da teoria dos conjuntos e os paradoxos da teoria semântica, no sentido de que paradoxos de ambos os tipos tornam-se ferramentas para provar incompletude e indecidibilidade. Tipicamente, uma dada noção paradoxal é formalizada como um predicado na linguagem de uma teoria que interpreta (pelo menos um fragmento da) a teoria dos números Z; então, aplica-se diagonalização, autorreferência etc., para obter afirmações correspondentes a sentenças número-teóricas que se tornam indecidíveis ou improváveis, dada uma hipótese de consistência.

6.2 Estruturas fundacionais e paradoxos

Motivações filosóficas exercem uma forte influência na investigação lógica contemporânea sobre paradoxos, e, portanto, é natural perguntar o que ainda sobrevive da teoria fregeana original dos conceitos, baseada em um princípio inconsistente de abstração e a visão logicista. Bem, a resposta é que um legado significativo ainda está vivo: isso é especialmente evidente na abordagem neofregeana, baseada no princípio de Hume e em suas variações. Subsistemas consistentes do Grundgesetze de Frege foram isolados e são estudados atualmente: veja Burgess (2005) para uma lista abrangente de referências a trabalhos anteriores de Boolos, Wright e Hale, Heck, Wehmeier, Ferreira, Antonelli e May (veja também os ensaios contidos em Reck e Cook (2016), assim como o verbete Logicism and Neologicism).[39] Além disso, pesquisas dentro da tradição neofregeana demonstraram que a inconsistência do princípio da abstração extensional (*Lei Básica V*) no Grundgesetze é apenas um caso particular de um resultado mais geral sobre inconsistência na lógica de segunda ordem (estendida com um símbolo para uma função f que mapeia conceitos para objetos) de quaisquer princípios de abstração que satisfaçam a condição chamada de "parte-todo": se *A* está estritamente incluído em B, então $f(A) \neq f(B)$ (veja Mancosu e Siskind 2019).

Por outro lado, se deixarmos de lado a inspiração ideológica do logicismo, poderíamos crer que o desenvolvimento da lógica e da teoria dos conjuntos no século XX eliminou com-

[39]N.T.: https://plato.stanford.edu/entries/logicism/

pletamente os paradoxos e que contradições em sistemas lógicos sejam fenômenos restritos aos anos da crise fundacional. Mas isso não é verdade: paradoxos foram descobertos em sistemas lógicos relacionados à ciência da computação. Por exemplo, há meio século, Girard demonstrou que uma teoria construtiva dos tipos desenvolvida por Per Martin-Löf (1971), baseada na correspondência de Curry-Howard, é inconsistente quando se assume a existência de um tipo de todos os tipos. A contradição decorre de uma reconstrução tipo-teórica de um argumento relacionado à antinomia de Burali-Forti e ao paradoxo de Mirimanoff dos conjuntos fundados. Mais tarde, Coquand (1986) provou que certas extensões do Cálculo C de construções são inconsistentes. De maneira simplificada, C é uma teoria dos tipos impredicativa de ordem superior que estende o Sistema F de Girard, um poderoso cálculo lambda de segunda ordem e com abstração sobre tipos, que é adequado para representar provas da lógica intuicionista impredicativa de segunda ordem. Coquand (1994) apresentou um novo paradoxo que afeta a teoria dos tipos, refinando o resultado de Reynolds, segundo o qual não existe um modelo conjuntista clássico para o polimorfismo (ou seja, o Sistema F de Girard; veja os verbetes Type Theory e Set Theory: Constructive and Intuicionistic ZF).[40] Por outro lado, um desenvolvimento geral e sem tipos da teoria das construções como uma base para a demonstrabilidade construtiva em lógica e matemática foi originalmente proposto por Kreisel e Goodman, mas revelou-se sujeito a uma antinomia, recentemente reexaminada por Dean e Kurokawa (2016).

Na fronteira entre questões fundacionais e aplicações em ciência da computação, argumentos de natureza paradoxal aparecem na investigação da *matemática explícita* de Feferman (EM), uma teoria de operações (autoaplicáveis) e classificações não-extensionais. Por exemplo, a existência de uma construção de tipo potência forte [41] leva à inconsistência nas chamadas “teorias dos tipos e nomes”, um desenvolvimento de EM introduzido por Jäger, 1997 (para uma bibliografia detalhada sobre EM de 1975–2013, veja o sítio eletrônico Bibliografia da Matemática Explícita;[42] para resultados mais recentes conectados com EM, veja Hayashi 2025). Argumentos paradoxais também são úteis para avaliar o papel de universos e *refutar a não-extensionalidade* em EM, quando presentes formas de compreensão ingênua *uniforme* (veja Cantini e Minari 1999). De fato, o papel da uniformidade é essencial nessas

[40] N.T.: https://plato.stanford.edu/entries/type-theory/
https://plato.stanford.edu/entries/set-theory-constructive/

[41] N.T.: “*strong power type*”.

[42] N.T.: https://home.inf.unibe.ch/~til/em_bibliography/

investigações. No que diz respeito à compreensão ingênua, já se sabe desde os anos 70 (Malitz 1976) e 80 (Weydert 1988, Forti e Hinnion 1989) que existem bons modelos topológicos da *extensionalidade* e da compreensão ingênua *não uniforme* restrita a *condições de positividade generalizadas*. Isso levou ao estudo dos chamados *hiperuniversos*. Resultados adicionais de consistência/inconsistência sobre a relação entre extensionalidade e princípios de compreensão uniforme vs. não uniforme podem ser encontrados em Hinnion e Libert (2003), assim como em Libert e Esser (2005). Em uma direção semelhante, propostas teóricas recentes combinam novamente ideias da lógica combinatória e do cálculo lambda com reinterpretações “indutivas” da compreensão ingênua e do esquema de verdade irrestrito, em uma linha de pesquisa já iniciada por Fitch no final da década de 40 (veja Scott 1975, Flagg e Myhill 1987, Aczel 1980 e Feferman 1984). A partir de 1992, houve uma tentativa, por parte de K. Grue, de ressuscitar o cálculo lambda de Church como um fundamento para a matemática. Grue (2002) apresenta uma extensão muito forte do cálculo lambda, a chamada *teoria dos mapas*, na qual a teoria axiomática dos conjuntos padrão torna-se interpretável, e que pode ser utilizada para esclarecer a diferença entre os paradoxos de Russell e Burali-Forti.

6.3 Circularidade e autorreferência

Desde os trabalhos de Mirimanoff, Finsler e outros, lógicos têm estudado universos da teoria dos conjuntos nos quais conjuntos circulares existem. No entanto, foi apenas a partir do início da década de 80 que uma matemática genuína dos conjuntos não bem-fundados (veja o verbete Non-wellfounded Set Theory)[43] passou a ser desenvolvida. Utilizando o axioma AFA de antifundação, permite-se a autorreferência direta na teoria dos conjuntos, e existem inúmeros conjuntos que solucionam equações autorreferenciais gerais (AFA foi introduzido por Forti e Honsell em 1983; para um desenvolvimento sistemático e histórico, veja Aczel, 1988). Em particular, conjuntos não bem-fundados são aplicados à análise de paradoxos, à semântica das línguas naturais e à ciência teórica da computação (veja Barwise e Etchemendy 1984, Barwise e Moss 1996).

Sobre a questão de se a autorreferência pode ser evitada na derivação de paradoxos e, portanto, se existem contradições genuínas decorrentes da ausência de fundação, uma resposta positiva foi dada pelo paradoxo semântico de Yablo (1993): há infinitos agentes

[43]N.T.: https://plato.stanford.edu/entries/nonwellfounded-set-theory/

etc., cada um afirmando a mesma sentença: "pelo menos um agente que vem depois de mim está mentindo"; no entanto, isso leva a uma contradição (veja os verbetes Self-reference and Paradox e Liar Paradox).[44] A questão levantada por essa construção — se a circularidade e a autorreferência são condições necessárias e suficientes para o surgimento de paradoxos — foi posteriormente analisada em Yablo (2006) (para um estudo abrangente sobre o tema, veja Cook 2014; para uma prova sem o lema da diagonalização, veja Halbach e Zhang 2017). Além disso, a conexão entre os fenômenos da incompletude e dos paradoxos foi estendida de modo a incluir o paradoxo de Yablo como um caso especial dos paradoxos do tipo Mentiroso (Kurahashi 2014, Kikuchi e Kurahashi 2016).

A análise da autorreferência e da diagonalização motivou a aplicação de técnicas algébricas e topológicas: considere os modelos de Scott para o cálculo lambda extensional (Scott 1972) e sua subsequente interpretação categórica. Por outro lado, a teoria das categorias tem sido utilizada para novas abordagens aos paradoxos desde Lawvere (1969). Uma abordagem matemática à questão geral de "autorreferência vs. não fundação" pode ser encontrada em Bernardi (2001, 2009). Além do paradoxo de Yablo e de uma versão baseada em teoria dos jogos do paradoxo de Mirimanoff, diversos resultados clássicos (como a existência de um conjunto não recursivamente enumerável e os teoremas de Cantor sobre a não enumerabilidade dos reais) podem ser transformados em teoremas de existência para certas cadeias não fundadas (e, formalmente, tais cadeias são vistas como pontos fixos generalizados).

6.4 Dos paradoxos à incompletude

Na metamatemática padrão, um papel importante para a compreensão aprofundada do segundo teorema da incompletude é desempenhado pelo *teorema de Löb* (Löb 1955). A prova desse teorema está relacionada ao *paradoxo de Curry* (veja o verbete Curry's Paradox)[45] e a um argumento informal devido a Geach (1955). Além disso, o teorema de Löb é essencial para definir estruturas matemáticas, estas adequadas para fornecer versões de autorreferência e incompletude (veja as chamadas *álgebras de Magari* e a análise modal da demonstrabilidade formal em Boolos 1993). Na mesma direção, há aplicações do *paradoxo de Berry*. Por exemplo, em 1966, Vopěnka provou o segundo teorema da incompletude para

[44] N.T.: https://plato.stanford.edu/entries/self-reference/
https://plato.stanford.edu/entries/liar-paradox/

[45] N.T.: https://plato.stanford.edu/entries/curry-paradox/

a teoria dos conjuntos e classes de Bernays-Gödel utilizando uma forma desse paradoxo. Boolos (1989) explora um argumento do tipo Berry[46] para provar a incompletude na forma "*não existe um algoritmo cuja saída contenha todas as sentenças verdadeiras da aritmética e nenhuma falsa*" sem recorrer à diagonalização. O paradoxo de Berry também foi relacionado aos fenômenos da incompletude devido a trabalhos (desde os anos 1960 e 1970) na complexidade de Kolmogorov e na *teoria algorítmica da informação*. Em particular, Chaitin demonstrou, em diversos artigos, como explorar a aleatoriedade para provar certas limitações dos sistemas formais (veja Chaitin 1995). Em conexão com os resultados de Chaitin, Kritchman e Raz (2011) fornecem uma prova do segundo teorema da incompletude baseada em um argumento semelhante ao *paradoxo do teste surpresa* (veja o verbete Epistemic Paradoxes).[47] Esse paradoxo, por sua vez, pode ser explicitamente relacionado ao teorema da completude de Solovay para a lógica da demonstrabilidade (veja Montagna 1994). Mais recentemente, Egré (2005), De Voos, Kooj e Verbrugge (2018) aplicaram a lógica da demonstrabilidade para resolver o *paradoxo do Conhecedor*. Vale lembrar que — novamente no campo da lógica epistêmica — a autorreferência pode ser aplicada para provar a *incompletude em modelos de crença*. Brandenburger e Keisler (2006) identificaram um paradoxo de autorreferência no âmbito das crenças em jogos, que resulta em um *teorema da impossibilidade jogo-teórico*, análogo ao paradoxo de Russell. Uma versão informal desse paradoxo é a de que a seguinte configuração de crenças é impossível: Ann acredita que Bob assume que Ann acredita que a suposição de Bob está errada. Essa versão é formalizada para demonstrar que qualquer modelo de crença de certo tipo deve conter uma "lacuna". Uma interpretação desse resultado é que "se as ferramentas do analista estiverem disponíveis para os jogadores em um jogo, *então haverá sentenças sobre as quais os jogadores podem pensar, mas que não podem assumir*". Além disso, o simples fato de paradoxos como esse, com um distinto caráter lógico, surgirem em contextos que envolvem noções epistêmicas é notável, podendo ser considerado um indício de que essa linha de pesquisa, baseada nos muitos exemplos já disponíveis na literatura, tende a se expandir e se tornar sistemática no futuro graças à aplicação de métodos formais (ver seção 6.5). Embora não esteja diretamente conectado ao fenômeno da incompletude, há diversos relatos de antinomias que afetam conjuntos de postulados aparentemente legítimos e naturais envolvendo certas noções epistêmicas. Um exemplo importante desse tipo de fonte é o trabalho de Leitgeb (2021),

[46]N.T.: "*Berry-type argument*".

[47]N.T.: https://plato.stanford.edu/entries/epistemic-paradoxes/

apresentando uma solução para o paradoxo da loteria (veja Kyburg 1961; veja também o verbete Epistemic Paradoxes)[48], que afeta princípios que relacionam crença categórica e crença gradativa (para maiores detalhes, veja Douven 2021).

6.5 Sobre os fundamentos da semântica, novamente

No último quarto do século passado, uma série de artigos em lógica surgiu a partir da discussão sobre os fundamentos da semântica, com diversas propostas que generalizam ou modificam a concepção semântica da verdade de Tarski. Não obstante, desde 1975 a abordagem hierárquica de Tarski foi, em certa medida, suplantada por novas ideias que tornaram o ideal de fechamento lógico e semântico acessível em muitos aspectos (especialmente por meio dos métodos de ponto fixo usados por Kripke 1975 e Martin-Woodruff; veja Martin 1984). Merece destaque também a abordagem derivada de Herzberger, Gupta e Belnap (1993) (veja o verbete Revision Theory of Truth)[49], que possui conexões com partes não elementares da teoria da definibilidade, teoria dos conjuntos e teoria da recursão superior[50] (Welch 2001, 2009, 2011, 2019). Isso levou ao estudo axiomático geral das *definições revisão-teóricas*[51] e teorias das *definições circulares* (veja Bruni 2009, 2013b, 2015, 2019). Em Standefer (2015) é estabelecida uma conexão entre um teorema do tipo Solovay e definições circulares da teoria da revisão, e uma lógica modal particular RT (teoria da revisão): um teorema de completude para essa lógica, análogo ao teorema de completude de Solovay para GL, é provado. A lógica modal em questão é construída sobre um operador naturalmente conectado à construção revisão-teórica (por isso o nome), como explicado em Gupta e Standefer (2017). A teoria da demonstração desse sistema é estudada em Standefer (2018).

Os avanços recentes na teoria da revisão contribuíram para demonstrar a fertilidade dessa abordagem (em uma direção relacionada à última parte da seção 6.4 deste verbete): Gupta (2011) apresenta uma aplicação ao conceito de racionalidade estratégica em um certo tipo de jogos finitos, que está ligado a uma forma comum de entender a escolha racional em contextos estratégicos. Embora desenvolvida de maneira independente, a abordagem de Gupta lembra o trabalho anterior de H. Gaifman sobre a racionalidade sendo afetada por paradoxos semelhantes aos paradoxos da verdade como o paradoxo do Mentiroso (veja

[48]N.T.: Veja a nota 49.
[49]N.T.: https://plato.stanford.edu/entries/truth-revision/
[50]N.T.: ”*higher recursion theory*”.
[51]N.T.: ”*revision-theoretic definitions*”.

Gaifman 1999). Uma extensão dessa abordagem para a classe de todos os jogos finitos é apresentada em Bruni e Sillari (2018). O aspecto notável da aplicação de Gupta é o emprego exclusivo da parte finita da teoria da revisão (ou seja, não requer a iteração transfinita dessa construção) e a eliminação da regra para lidar com estágios correspondentes a ordinais limite. Essa regra do limite, essencial para abordar o conceito de verdade, revelou-se o aspecto mais crítico da abordagem revisão-teórica para conceitos circulares, tanto do ponto de vista da complexidade quanto do ponto de vista conceitual (para uma abordagem recente e inovadora sobre o tema, veja Campbell-Moore 2019).

A riqueza e variedade de ferramentas semânticas têm desencadeado uma espécie de experimentação com diversas propostas *mistas*. Refinamentos da teoria da revisão e suas generalizações podem ser encontrados em Rivello (2019a, 2019b) onde é desenvolvida uma nova abordagem para uma teoria formal da verdade, que implementa características da teoria do ponto fixo de Kripke com a teoria da revisão de Herzberger-Gupta. Uma combinação análoga, estudada sob um viés mais voltado à teoria da demonstração, foi considerada em Standefer (2017). De forma semelhante, combinações da semântica de mundos possíveis para o predicado de necessidade com os modelos parciais supervaloracionais de Kripke para o predicado de verdade também são estudadas por Nicolai, 2018 (para uma tentativa recente de completar a figura das teorias axiomáticas de verdade supervaloracional tomando por base as contribuições de Cantini 1990, veja Dopico e Hayashi 2024). [52]

Field (2003, 2008) propôs soluções influentes para os paradoxos semânticos que combinam técnicas kripkeanas e técnicas da teoria da revisão. Segundo ele, as soluções atuais para os paradoxos não são satisfatórias devido aos seguintes pontos: (i) a ausência de uma condicional (e bicondicional) adequada; (ii) a falha (de algumas instâncias) do esquema–T, $A \leftrightarrow T(\ulcorner A \urcorner)$; (iii) a falha da intersubstitutividade entre A e $T(\ulcorner A \urcorner)$; (iv) a impossibilidade de fornecer uma análise interna da defeituosidade das sentenças paradoxais. Como consequência, Field (2008) desenvolveu uma teoria da verdade com um *operador condicional não clássico*, permitindo expressar uma noção de *verdade determinada* e afirmar que o paradoxo do Mentiroso não é determinadamente verdadeiro. A análise da construção de Field requer desenvolvimentos sofisticados da teoria dos conjuntos e da teoria da recursão (veja Welch 2008, 2009, 2011). Além disso, a possibilidade de enriquecimento lógico da teoria da verdade de Kripke com *novos condicionais* abriu novas direções de pesquisa. Por exemplo, Rossi (2016) propõe um método interessante para incorporar uma condicional que satisfaça

[52] N.T.: https://arxiv.org/abs/2410.12471

as exigências da lógica trivalente de Łukasiewicz em uma construção de ponto fixo para a verdade. Na mesma direção, considerável atenção tem sido direcionada na literatura recente ao chamado *problema da revanche*: soluções típicas, por exemplo, do paradoxo do Mentiroso baseiam-se em noções que, se expressáveis na linguagem objeto, geram novas versões do paradoxo. Assim, a solução seria apenas uma ilusão.[53] O problema da revanche pode ser exemplificado pelo Mentiroso Reforçado: informalmente, uma vez que temos um modelo que faz com que *a sentença mentirosa M* não seja nem verdadeira nem falsa e podemos expressar esse fato, então *M* não é verdadeira. Mas essa é exatamente a afirmação feita por *M*, logo *M* é verdadeira. Assim, o paradoxo parece ressurgir (para mais detalhes, veja o verbete Liar Paradox;[54] veja também a coletânea de ensaios em Beall 2007).

Soluções "indexicais" para o paradoxo do Mentiroso foram desenvolvidas em diversas contribuições, como as de Burge, Gaifman e Simmons. A ideia é que o paradoxo do Mentiroso não envolve sentenças, mas *ocorrências específicas de sentenças*, isto é, tokens de sentenças (essa ideia pode ser encontrada já em soluções escolásticas). Por questão de precisão histórica, vale mencionar que, em 1913, Leśniewski, mais tarde orientador de Tarski, já havia proposto uma solução indexical de inspiração nominalista para o paradoxo do Mentiroso em seu artigo *A Critique of the Logical Principle of the Excluded Middle* (veja Betti 2004).

Além do aspecto modelo-teórico, *investigações axiomáticas sobre a verdade* e paradoxos relacionados tornaram-se cada vez mais importantes desde os artigos pioneiros de Friedman e Sheard (1987) e Feferman (1991). Desde o ano 2000, essa linha de pesquisa tem sido intensamente estudada com diversos objetivos, desde *análises prova-teóricas* até discussões filosóficas sobre *minimalismo* (para um panorama das variedades de sistemas teóricos da verdade e referências apropriadas, *veja* as monografias recentes de Halbach, 2011 e Horsten, 2011, bem como o verbete Axiomatic Theories of Truth;[55] veja também os artigos de Feferman 2008, Fujimoto 2010, Leigh e Rathjen, 2010).

Por fim, mas não menos importante, o estudo axiomático de noções epistêmicas tem sido consideravelmente beneficiado pela aplicação de técnicas usadas para provar resultados da incompletude e da indefinibilidade desde os anos 1960. Essas técnicas levaram a resultados negativos (Kaplan e Montague 1960, Montague 1963, e Thomason 1980) e esta-

[53]N.T.: Também encontrado na literatura sob como "problema da vingança".
[54]N.T.: https://plato.stanford.edu/entries/liar-paradox/
[55]N.T.: https://plato.stanford.edu/entries/truth-axiomatic/

beleceram um elo interessante com o paradoxo do teste surpresa. A situação pode também ter mudado pelo estudo da semântica de mundos possíveis para noções modais, concebidas como predicados em Halbach, Leitgeb e Welch (2003). No entanto, esse tema ainda está em aberto para debate e experimentação: por exemplo, argumenta-se em Halbach e Welch (2009) que a abordagem predicativa da necessidade é um caminho viável — conquanto considerado o poder expressivo — caso se utilize linguagens que envolvem tanto um *predicado de verdade* quanto um *operador de necessidade*.

6.6 Mantendo-se não clássico

Diversas soluções foram propostas baseando-se no uso de lógicas paraconsistentes (Priest) ou lógicas subestruturais (vide os verbetes Paraconsistent Logic e Substructural Logics,[56] bem como Mares e Paoli 2014).

A investigação de paradoxos semânticos e conjuntistas no contexto da lógica infinitamente valorada — inicialmente explorada por Mow Shaw-Kwei (1954) e Skolem (1957) — recebeu um novo impulso com as contribuições de Hájek, Shepherdson e Paris (2000), bem como Hájek (2005, 2010) (veja Fisher, Halbach, Kriener e Stern 2015 para uma discussão crítica sobre a interação entre as abordagens axiomática e a semântica da verdade). Tipicamente, artigos como esses aplicam resultados básicos da *análise matemática* (por exemplo, o teorema do ponto fixo de Brouwer). Vale mencionar que Leitgeb (2008) forneceu uma prova da consistência para uma *teoria probabilística da verdade* com o esquema–T irrestrito empregando o teorema de Hahn-Banach. Um desdobramento dessa abordagem, que se conecta com pesquisas relacionadas à teoria da revisão (ver seção 6.5) é encontrado em Campbell-Moore, Horsten e Leitgeb (2019).

Teorias da verdade ingênua — baseadas na bicondicional irrestrita e em uma lógica sem contração — podem ser encontradas na literatura, por exemplo, em Cantini (2002), Zardini (2011), Bacon (2013) e Standefer (2016). Por outro lado, Ripley (2012) propõe uma abordagem alternativa baseada em um sistema lógico não transitivo (veja também Cobreros *et al.* 2012 e Cobreros *et al.* 2015, que tentam estender essa abordagem aos paradoxos da vagueza).

Grišin (1981) demonstrou que um sistema baseado no princípio paradoxal *por exce-*

[56]N.T.: https://plato.stanford.edu/entries/logic-paraconsistent/
https://plato.stanford.edu/entries/logic-substructural/

lência — o esquema uniforme da compreensão ingênua — e em alguma forma de lógica não contrativa possui a eliminação do corte, e, portanto, é consistente. Por outro lado, a consistência do sistema é destruída pela extensionalidade, o que pode ser visto como um paradoxo adicional! Aqui, "uniforme" significa que um operador de abstração que vincula variáveis[57] $\{x \mid \varphi(x,a)\}$ e nomeia o conjunto definido por φ, dependendo de uma lista de parâmetros a, está disponível. Curiosamente, foi demonstrado que sistemas intimamente relacionados possuem aplicações inesperadas na caracterização de classes de complexidade (Girard 1998, Terui 2004). Por outro lado, o sistema é *computacionalmente completo* (ele pode interpretar a lógica combinatória, Cantini 2003).

A escolha de uma lógica de base não clássica frequentemente não é isenta de consequências, no que tange à facilidade de uso de uma teoria. Uma tendência recente na literatura (baseada em Halbach e Nicolai 2018) é a tentativa de comparar teorias clássicas e não clássicas (quando cabível, como a axiomatizações da semântica de ponto fixod e Kripke) para determinar o custo relacionado a uma escolha em relação à outra (ver também Castaldo 2023a, 2023b, e também Castaldo 2024, que tenta argumentar a favor das teorias clássicas por meio da recaptura).

Finalmente, uma vez que o caráter da verdade clássica (tarskiana) é a sua natureza infinitária, poder-se-ia classificar como não clássica a investigação de Sato e Walker 2023, a qual tenta esclarecer o relacionamento entre verdade e finitismo respondendo à pergunta: Faz sentido investigar uma versão da "verdade axiomática finitista"?

6.7 Em direção a uma abordagem "geométrica"

Além do uso de ferramentas da álgebra e da análise, investigações lógicas recentes sobre paradoxos têm aplicado a teoria dos grafos (veja Cook 2004, Rabern, Rabern e Macauley 2013, Beringer e Schindler 2017, Hsiung 2017). A ideia central é tentar capturar em termos geométricos *os padrões do paradoxo*, *suas características estruturais*. Por exemplo, é possível associar um *grafo de referência* (rfg) às sentenças da linguagem da aritmética com verdade através da noção de dependência de Leitgeb (Leitgeb 2005). Ainda, conjectura-se que uma solução para o problema de caracterização de rfgs perigosos equivale à afirmação de que, basicamente, *o grafo do Mentiroso e o grafo de Yablo são os únicos rfgs paradoxais*. Essa linha de pesquisa é desenvolvida independentemente em Rossi (2019), explorando

[57]N.T.: "*variable-binding*".

uma ampla gama de comportamentos semânticos exibidos por sentenças paradoxais, propondo uma teoria unificada da verdade e do paradoxo (veja também Bruni e Rossi 2023, que, ao estender esses meios, fornece uma unificação de paradoxos do tipo Mentiroso e do tipo sorites). O resultado é uma teoria da verdade que fornece uma classificação tríplice das sentenças paradoxais (sentenças do tipo Mentiroso, sentenças do tipo contador de verdade, e sentenças do tipo revanche) e propõe um modo de interpretar esses três tipos de sentenças paradoxais, bem como sentenças livres de paradoxo, dentro de um único modelo no qual a ferramenta fundamental é a noção de “grafo semântico”. Por sua vez, em Hsiung (2020), a abordagem de Leitgeb é combinada com a de Beringer e Schindler (2017) para estudar um tipo de paradoxos de característica finitária,[58] definido em termos da própria relação de dependência de Leitgeb.

Notavelmente, a teoria dos grafos não parece ser a única maneira pela qual se tenta alcançar um tratamento uniforme dos paradoxos (na forma de uma teoria da “paradoxicalidade”; ver, por exemplo, Castaldo 2021).

Referências Bibliográficas

P. Aczel. Frege structures and the notion of proposition, truth and set. In J. Barwise et al., editor, *The Kleene Symposium*, pages 31–59. North-Holland, Amsterdam, 1980.

P. Aczel. *Non-Well-Founded Sets*. Center for the Study of Language and Information, Stanford, 1988.

C. A. Anderson and M. Zelëny, editors. *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*. Kluwer, Dordrecht, 2001.

A. Bacon. Curry’s paradox and omega inconsistency. *Studia Logica*, 101:1–9, 2013.

H. Barendregt. *The Lambda Calculus. Its Syntax and Semantics* (Studies in Logic and the Foundations of Mathematics, vol. 103). Elsevier, Amsterdam, 1984.

J. Barwise and J. Etchemendy. *The Liar*. Oxford University Press, New York, 1984.

J. Barwise and L. Moss. *Vicious Circles. On the Mathematics of Non-Wellfounded Phenomena*. Center for the Study of Language and Information, Stanford, 1996.

J. C. Beall, editor. *Liars and Heaps*. Clarendon Press, Oxford, 2003.

J. C. Beall, editor. *Revenge of the Liar. New Essays on the Paradox*. Oxford University Press, New York, 2007.

[58] N.T.: Estes paradoxos são também conhecidos como paradoxos localmente finitos.

H. Behmann. Zu den Widersprüchen der Logik und der Mengenlehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 40:37–48, 1931.

T. Beringer and T. Schindler. A graph-theoretic analysis of the semantic paradoxes. *Bull. Symb. Log.*, 23:442–492, 2017.

C. Bernardi. Fixed points and unfounded chains. *Annals of Pure and Applied Logic*, 109: 163–178, 2001.

C. Bernardi. A topological approach to yablo's paradox. *Notre Dame Journal of Formal Logic*, 50:331–338, 2009.

F. Bernstein. Die Theorie der reellen Zahlen. *Jahresbericht der Deutschen MathematikerVereinigung*, 14:447–449, 1905a.

F. Bernstein. Über die Reihe der transfiniten Ordnungszahlen. *Mathematische Annalen*, 60: 187–193, 1905b.

F. Bernstein. Zum Kontinuumproblem. *Mathematische Annalen*, 60:463–464, 1905c.

A. Betti. Lesniewski's early liar, tarski and natural language. *Annals of pure and applied logic*, 127:267–287, 2004.

G. Birkhoff. *Lattice Theory*. American Mathematical Society, Providence, 3rd edition, 1967.

D. A. Bochvar. On a three-valued logical calculus and its applications to the analysis of the paradoxes of the classical extended functional calculus. *History and Philosophy of Logic*, 2(1):87–112, 1981. Tradução em inglês de M. Bergmann do original em russo, publicado em Mathematicheski Sbornik, 4 (46): 287–308, 1937.

G. Boolos. A new proof of the gödel incompleteness theorem. *Notices of the American Mathematical Society*, 36:388–390, 1989.

G. Boolos. *The logic of provability*. Cambridge University Press, Cambridge, 1993.

E. Borel. Les paradoxes de la théorie des ensembles. *Annales scientifique de l'École Normale Supérieure*, 25:443–448, 1908. Reimpresso em Borel, E., *Leçons sur la théorie des functions*, Paris: Gauthier Villars, 2. ed, 1914, 162–166.

E. Borel, R. Baire, J. Hadamard, and H. Lebesgue. Cinq lettres sur la théorie des ensembles. *Bulletin de la Société Mathématique de France*, 33:261–273, 1905.

A. Brandenburger and H. J. Keisler. An impossibility theorem on beliefs in games. *Studia Logica*, 84:211–240, 2006.

L. E. J. Brouwer. *Over die Grondslagen der Wiskunde*. Tese, Amsterdam. Tradução em inglês em Brouwer, L.E.J., *Collected Works I*, Amsterdam: North Holland, 1975., 1907.

R. Bruni. A note on theories for quasi-inductive definitions. *Review of Symbolic Logic*, 2:

684–699, 2009.

R. Bruni. Analytic calculi for circular concepts by finite revision. *Studia Logica*, 101:915–932, 2013a.

R. Bruni. Beppo levi's analysis of the paradoxes. *Logica Universalis*, 7:211–231, 2013b.

R. Bruni. Some remarks on the finite theory of revision. In D. Achourioti et al., editor, *Unifying the philosophy of truth*, Logic, Epistemology, and the Unity of Science: Volume 36, pages 169–187. Springer, Dordrecht, 2015.

R. Bruni. Addressing circular definitions via systems of proofs. In S. Centrone et al., editor, *Mathesis Universalis. Computability and Proof*, pages 75–100. Springer Verlag, 2019.

R. Bruni and L. Rossi. Truth meets vagueness. Unifying the semantic and the soritical paradoxes. *Journal of Philosophical Logic*, 52(6):1637–1671, 2023.

R. Bruni and G. Sillari. A rational way of playing: revision theory for strategic interaction. *Journal of Philosophical Logic*, 47:419–448, 2018.

O. Bueno, C. Menzel, and E. N. Zalta. Worlds and propositions set free. *Erkenntnis*, 79: 797–820, 2014.

C. Burali-Forti. Una questione sui numeri transfiniti. *Rendiconti del Circolo Matematico di Palermo*, 11(260):154–164, 1897. Tradução em inglês In van Heijenoort 1967, 104–111.

A. G. Burgess and J. P. Burgess. *Truth*. Princeton University Press, Princeton, 2011.

J. Burgess. *Fixing Frege*. Princeton University Press, Princeton, 2005.

C. Campbell-Moore. Limits in the revision theory: More than just definite verdicts. *Journal of Philosophical Logic*, 48:11–35, 2019.

C. Campbell-Moore, L. Horsten, and H. Leitgeb. Probability for the Revision Theory of Truth. *Journal of Philosophical Logic*, 48:87–112, 2019.

A. Cantini. Partial truth. In L. Horsten and V. Halbach, editors, *Principles of Truth*, pages 183–202. Hansel-Hohenhausen, Frankfurt, 2002.

A. Cantini. The undecidability of grišin's set theory. *Studia Logica*, 74:345–368, 2003.

A. Cantini. On a russellian paradox about propositions and truth. In G. Link, 259–284, 2004.

A. Cantini and P. Minari. Uniform inseparability in explicit mathematics. *The The Journal of Symbolic logic*, 64:313–326, 1999.

G. Cantor. *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. In E. Zermelo, editor. Olms: Berlin–Hildesheim, 1962.

G. Cantor. *Briefe*. In H. Meschkowski and W. Nilson, editors. Olms: Berlin–Hildesheim, 1991.

R. Carnap. Die Antinomien und die unvollständigkeit der Mathematik. *Monatshefte für Mathe-*

matik und Physik, 41:263–284, 1934a.

R. Carnap. *Die logische Syntax der Sprache*. Springer, Berlin, 1934b.

L. Castaldo. Fixed-point models for paradoxical predicates. *Australasian Journal of Logic*, 18(7):688–723, 2021.

L. Castaldo. Notes on models of (partial) Kripke–Feferman truth. *Studia Logica*, 111:83–111, 2023a.

L. Castaldo. On the costs of classical logic. *Erkenntnis*, 88:1157–1188, 2023b.

L. Castaldo. Saving logic from paradox via nonclassical recapture. *Philosophical Studies*, 181:1547–1563, 2024.

G. Chaitin. The Berry paradox. *Complexity*, 1:26–30, 1995.

A. Church. A set of postulates for the foundation of logic. *Annals of Mathematics*, 33:346–366, 1932.

A. Church. A set of postulates for the foundation of logic. *Annals of Mathematics*, 34:839–864, 1933.

A. Church. The Richard paradox. *American Mathematical Monthly*, 41:356–361, 1934.

A. Church. A comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski. *The Journal of Symbolic logic*, 41:747–760, 1976.

L. Chwistek. Antinomje logiki formalnej. *Przegląd Filozoficzny*, 24:164–171, 1921. Tradução em inglês por Z. Jordan: Antinomies of formal logic. In S. McCall, editor, *Polish Logic 1920–1939*, Oxford: Clarendon Press, 338–345, 1967.

L. Chwistek. Über die Antinomien der Prinzipien der Mathematik. *Mathematische Zeitschrift*, 14:236–243, 1922.

L. Chwistek. Die nominalistische Grundlegung der Mathematik. *Erkenntnis*, 3:367–388, 1933.

P. Cobreros, P. Egré, D. Ripley, and R. van Rooij. Vagueness, truth and permissive consequence. In K. Fujimoto, J. Martínez Fernández, H. Galinon, and T. Achourioti, editors, *Unifying the Philosophy of Truth*, pages 409–430. Springer Verlag, 2002.

P. Cobreros, P. Egré, D. Ripley, and R. van Rooij. Tolerant, classical, strict. *Journal of Philosophical Logic*, 41:347–385, 2012.

R. T. Cook. Patterns of paradox. *The Journal of Symbolic logic*, 69:767–774, 2004.

R. T. Cook. *The Yablo Paradox*. Oxford University Press, New York, 2014.

T. Coquand. An analysis of girard's paradox. *Proceedings of the IEEE Symposium on Logic in Computer Science*, pages 227–236, 1986.

T. Coquand. A new paradox in type theory. In D. Prawitz, B. Skyrms, and D. Westerstahl, editors, *Logic, Methodology and Philosophy of Science IX* (Studies in Logic and the Foundations of Mathematics, vol. 134), pages 555–570. North-Holland, Amsterdam, 2002.

H. B. Curry. Grundlagen der kombinatorischen Logik. *American Journal of Mathematics*, 52: 509–536, 1930.

H. B. Curry. The paradox of Kleene and Rosser. *Transactions of the American Mathematical Society*, 41:454–516, 1941.

H. B. Curry. The inconsistency of certain formal logics. *The Journal of Symbolic logic*, 7: 115–117, 1942.

M. De Vos, B. Kooi, and R. Verbrugge. Provability Logic meets the knower paradox. In G. Bezhanishvilii, editor, *Advances in Modal Logic 2018*, pages 31–35. Rijksuniversiteit Groningen, 2018.

W. Dean. Incompleteness via paradox. *Review of Symbolic Logic*, 13(3):541–592, 2020.

W. Dean and H. Kurokawa. Kreisel's Theory of Constructions, the Kreisel-Goodman Paradox, and the Second Clause. In T. Piecha and P. Schroeder-Heister, editors, *Advances in Proof-Theoretic Semantics* (Trends in Logic, Volume 43), pages 27–63. Springer International Publishing, 2016.

I. Douven, editor. *Lotteries, Knowledge, and Rational Belief: Essays On the Lottery Paradox*. Cambridge University Press, Cambridge, 2021.

S. Eberhard and T. Strahm. Unfolding Feasible Arithmetic and Weak Truth. In T. Achourioti, H. Galinon, J.M. Fernandez, and K. Fujimoto, editors, *Unifying the Philosophy of Truth*, pages 153–167. Springer, Dordrecht, 2015.

P. Egré. The Knower Paradox in the Light of Provability Interpretations of Modal Logic. *Journal of Logic, Language, and Information*, 14:13–48, 2005.

A. Enayat and A. Visser. New constructions of satisfaction classes. In T. Achourioti, H. Galinon, J.M. Fernandez, and K. Fujimoto, editors, *Unifying the Philosophy of Truth*, pages 321–335. Springer, Dordrecht, 2015.

S. Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta mathematicae*, 49:35–92, 1960.

S. Feferman. Systems of predicative analysis i. *The Journal of Symbolic logic*, 29:1–30, 1964.

S. Feferman. Constructive theories of functions and classes. In M. Boffa and D. van Dalen, editors, *Logic Colloquium '78*, pages 159–224. North-Holland, Amsterdam, 1979.

S. Feferman. Towards useful type-free theories. i. *The Journal of Symbolic logic*, 49:75–111, 1984a.

S. Feferman. Reflecting on incompleteness. *The Journal of Symbolic logic*, 56:1–49, 1984b.

S. Feferman. Axioms for determinatess and truth. *The Journal of Symbolic logic*, 1(2):204–217, 2008.

H. Field. A revenge-immune solution to the semantic paradoxes. *Journal of Philosophical Logic*, 32:139–177, 2003.

H. Field. *Saving Truth from Paradox*. Oxford University Press, New York, 2008.

P. Finsler. Gibt es Widersprüche in der Mathematik? *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 34:143–155, 1925.

P. Finsler. Formale Beweise und die Unentscheidbarkeit. *Mathematische Zeitschrift*, 25: 676–682, 1926a.

P. Finsler. Über die Grundlagen der Mengenlehre. *Mathematische Zeitschrift*, 25:683–713, 1926b.

F.B. Fitch. A system of formal logic without an analogue to Curry W-operator. *The Journal of Symbolic logic*, 1:92–100, 1936.

F.B. Fitch. A basic logic. *The Journal of Symbolic logic*, 7:105–114, 1942.

R. Flagg and J. Myhill. Implications and analysis in classical frege structures. *Annals of Pure and Applied Logic*, 34:33–85, 1987.

M. Forti and R. Hinnion. The consistency problem for positive comprehension principles. *The Journal of Symbolic logic*, 54:1401–1418, 1989.

M. Forti and F. Honsell. Set theory with free construction principles. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze*, Serie IV(10):493–522, 1983.

G. Frege. *Grundgesetze der Arithmetik. Begriffschriftlich Abgeleitet (Volume 2)*. Jena, 1903; reprinted, Hildesheim: Olms, 1962.

G. Frege. *Wissenschaftslicher Briefswechsel*. F. Meiner, Hamburg, 1976.

G. Frege. *Nachgelassene Schriften*. F. Meiner, Hamburg, 1984.

H. Friedman and M Sheard. An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33:1–21, 1987.

K. Fujimoto. Relative truth and definability of axiomatic truth theories. *Bulletin of Symbolic Logic*, 16(33):305–344, 2010.

K. Fujimoto. Autonomous progression and transfinite iteration of self-applicable truth. *The Journal of Symbolic logic*, 76:914–945, 2011.

K. Fujimoto. Classes and truths in set theory. *Annals of Pure and Applied Logic*, 163:1484–

1523, 2012.

H. Gaifman. Operational pointer semantics: Solution to selfreferential puzzles i. In M. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 43–59. Morgan Kaufmann, San Francisco, 1988.

H. Gaifman. Self-reference and the acyclity of rational choice. *Annals of Pure and Applied Logic*, 96:117–140, 1999.

A. Garciadiego. *Bertrand Russell and the Origins of Set-theoretic "Paradoxes"*. Birkhäuser, Basel, 1992.

P. Geach. On insolubilia. *Analysis*, 15:71–72, 1955.

J. Y. Girard. Light linear logic. *Information and Computation*, 143:175–204, 1998.

K. Grelling and L. Nelson. Bemerkungen zu den Paradoxien von Russell und Burali-Forti. *Abhandlungen der Fries'chen Schule*, 2:301–334, 1908.

V. N. Grišin. Predicate and set theoretic calculi based on logic without contraction rules. *Math. USSR Izv.*, 18(1):41–59, 1982. Tradução em inglês do original em russo, publicado em Izvestiya Akademii Nauk SSSR Seriya Matematicheskaya, 45(1): 47–68, 1981.

K. Grue. Lambda calculus as a foundation of mathematics. In C. A. Anderson and M. Zelëny, editors, *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*, pages 287–311. Kluwer, Dordrecht, 2001.

A. Gupta. On circular concepts. In *Truth, Meaning and Experience*, pages 95–134. Oxford University Press, Oxford, 2011.

A. Gupta and N. Belnap. *The Revision Theory of Truth*. MIT Press, Cambridge, MA, 1993.

A. Gupta and S. Standefer. Conditionals in theories of truth. *Journal of Philosophical Logic*, 46:27–63, 2017.

K. Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme i. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.

P. Hajek. On the arithmetic in the Cantor-Łukasiewicz set theory. *Archive for Mathematical Logic*, 44:763–782, 2005.

P. Hajek. On white's expansion of Łukasiewicz logic. *Journal of Logic and Computation*, 20: 389–397, 2010.

P. Hajek, J. Paris, and J. Shepherdson. The liar paradox and fuzzy logic. *The Journal of Symbolic logic*, 65:339–346, 2010.

V. Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, Cambridge, 2011.

V. Halbach and C. Nicolai. On the costs of non-classical logic. *Journal of Philosophical Logic*,

47:227–257, 2018.
V. Halbach and P. Welch. Necessities and necessary truths: a prolegomenon to the use of modal notions in the analysis of the intensional notions. *Mind*, 118:71–100, 2009.
V. Halbach and S. Zhang. Yablo without Gödel. *Analysis*, 77:53–59, 2017.
V. Halbach, H. Leitgeb, and P. Welch. Possible world semantics for modal notions conceived as predicates. *Journal of Philosophical Logic*, 32:179–223, 2003.
D. Hayashi. Theories of Frege structure equivalent to Feferman's system T_0. *Annals of Pure and Applied Logic*, 176(1):1–36, 2025.
G. Hessenberg. *Grundbegriffe der Mengenlehre*. Göttingen, Vandenhoek und Ruprecht, 1906; Também publicado em Abhandlungen der Fries'chen Schule, Neue Reihe, 1: 479–706, 1906.
D. Hilbert. Über die Grundlagen der Logik und der Arithmetik. In *Verhandlungen des Dritten Internationalen-Mathematiker Kongresses*, pages 174–185. Teubner, Leipzig, 1904. Tradução em inglês In van Heijenoort 1967, 129–138.
D. Hilbert. *Prinzipien der Mathematik, Vorlesung von D. Hilbert*. Mathematisches Institut der Georg-August Universität Göttingen, 1918.
D. Hilbert and W. Ackermann. *Grundzüge der theoretischen Logik*. Springer, Berlin-Heidelberg, 1928.
D. Hilbert and P. Bernays. *Grundlagen der Mathematik (Volumes I–II)*. Springer, Berlin, 1934/1939.
R. Hinnion and T. Libert. Positive abstraction and extensionality. *The Journal of Symbolic logic*, 68:828–836, 2003.
M. R. Holmes. Tarski's Theorem and NFU. In C. A. Anderson and M. Zelëny, editors, *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*, pages 469–478. Kluwer, Dordrecht, 2001.
L. Horsten. *The Tarskian Turn. Deflationism and Axiomatic Truth*. MIT Press, Cambridge, MA, 2011.
L. Horsten and G. Leigh. Truth is simple. *Mind*, 126:195–232, 2017.
M. Hsiung. Boolean paradoxes and revision periods. *Studia Logica*, 105:881–914, 2017.
M. Hsiung. What paradoxes depend on. *Synthese*, 197:887–913, 2020.
A. Irvine. Bertrand Russell's Logic. In D. M. Gabbay and J. Wood, editors, *Handbook of the History of Logic (Volume 5: Logic from Russell to Church)*, pages 1–28. Elsevier and North-Holland, Amsterdam, 2009.

G. Jäger. Power types in explicit mathematics. *The Journal of Symbolic logic*, 62:1142–1146, 1997.

R. Kahle. David Hilbert über Paradoxien. *Preprint Number 06–17, Departamento de Matemática, Universidade de Coimbra*, pages 1–42, 2004.

R. Kahle and V. Peckhaus. Hilbert's paradox. *Historia Mathematica*, 29:127–155, 2002.

D. Kaplan and R. Montague. A paradox regained. *Notre Dame Journal of Formal Logic*, 1: 79–90, 1960.

M. Kikuchi. A note on Boolos' proof of the incompleteness theorem. *Mathematical Logic Quarterly*, 40:528–532, 1994.

M. Kikuchi and T. Kurahashi. Liar-type paradoxes and the incompleteness phenomena. *Journal of Philosophical Logic*, 45:381–398, 2016.

K. Klement. Russell, His Paradoxes, and Cantor's Theorem: Part i. *Philosophy Compass*, 5: 16–28, 2010.

S. C. Klement and J. B. Rosser. The inconsistency of certain formal logics. *Annals of Mathematics*, 36:630–637, 1935.

G. Kreisel. Note on arithmetic models for consistent arithmetic formulae of the predicate calculus. *Fundamenta Mathematicae*, 37:265–285, 1950.

G. Kreisel. La prédicativité. *Bulletin de la Société Mathématique de France*, 88:371–391, 1960.

S. Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–716, 1975.

S. Kritchman and R. Raz. The surprise examination paradox and the second incompleteness theorem. *Notices of the American Mathematical Society*, 57:1454–1458, 2011.

T. Kurahashi. Rosser-type undecidable sentences and yablo's paradox. *Journal of Philosophical Logic*, 43:999–1017, 1961.

H. Kyburg. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown, 1961.

J. König. Über die Grundlagen der Mengenlehre und das Kontinuumsproblem. *Mathematische Annalen*, 61:156–160, 1905.

F.W. Lawvere. Diagonal arguments and cartesian closed categories. In P. Hilton, editor, *Category Theory, Homology Theory and their Applications, II*, (Volume 92: Lecture Notes in Mathematics), pages 134–145. Springer, Berlin-Heidelberg, 1969.

G. Leigh. A proof-theoretic account of classical principles of truth. *Annals of Pure and Applied Logic*, 164:1009–1024, 2013.

G. Leigh. Conservativity for theories of compositional truth via cut elimination. *The Journal of Symbolic logic*, 80:845–865, 2015a.

G. Leigh. Some Weak Theories of Truth. In T. Achourioti, H. Galinon, J.M. Fernandez, and K. Fujimoto, editors, *Unifying the Philosophy of Truth*, pages 281–292. Springer, Dordrecht, 2015b.

G. Leigh and M. Rathjen. An ordinal analysis for theories of self-referential truth. *Archive for Mathematical Logic*, 49:213–247, 2010.

G. Leigh and M. Rathjen. The Friedman-Sheard programme in intuitionistic logic. *The Journal of Symbolic logic*, 77:777–806, 2012.

H. Leitgeb. What truth depends on. *Journal of Philosophical Logic*, 34(2):155–192, 2005.

H. Leitgeb. What theories of truth should be like (but cannot be). *Blackwell Philosophy Compass*, 2:276–290, 2007.

H. Leitgeb. On the probabilistic convention T. *Review of Symbolic Logic*, 1:218–224, 2008.

H. Leitgeb. Stability and the lottery paradox. *Review of Symbolic Logic*, 1:218–224, 2021a.

H. Leitgeb. Stability and the lottery paradox. In I. Douven, editor, *Lotteries, Knowledge, and Rational Belief: Essays On the Lottery Paradox*, pages 147–170. Cambridge University Press, Cambridge, 2021b.

B. Levi. Intorno alla teoria degli aggregati. *Reale Istituto Lombardo di Scienze e Lettere, Rendiconti (2nd Series)*, 35:863–868, 1902.

B. Levi. Antinomie logiche? *Annali di Matematica (3rd Series)*, 15:188–216, 1908.

C. I. Lewis and C. H. Langford. *Symbolic Logic*. New York: The Century Co., 1932; reprinted, New York: Dover, 1952.

T. Libert and O. Esser. On topological set theory. *Mathematical Logic Quarterly*, 51:263–273, 2005.

G. Link, editor. *One Hundred Years of Russell's Paradox*. De Gruyter, Berlin, 2004.

B. Linsky. Leon Chwistek on the no-classes theory. *History and Philosophy of Logic*, 25: 53–71, 2004.

G. Lolli. A Berry-type paradox. In C. S. Calude, editor, *Randomness and Complexity. From Leibniz to Chaitin*, pages 155–160. World Scientific, Singapore, 2021.

M. H. Löb. Solution of a problem of Leon Henkin. *The The Journal of Symbolic logic*, 20: 115–118, 1955.

R. J. Malitz. *Set Theory in which the axiom of foundation fails*. Ph.d. thesis, Philosophy Department, University of California, Los Angeles, 1976.

P. Mancosu. The Russellian Influence on Hilbert and his School. *Synthèse*, 137:59–101, 2003.

P. Mancosu and B. Siskind. Neologicist Foundation: Inconsistent abstraction principles and part-whole. In G. M. Mras, P. Weingartner, and B. Ritter, editors, *Philosophy of Logic and Mathematics*, Proceedings of the 41st International Wittgenstein Symposium, pages 215–248. De Gruyter, Berlin, 2019.

P. Mancosu, R. Zach, and C. Badesa. The development of mathematical logic from russell to tarski: 1900–1935. In L. Haaparanta, editor, *The History of Modern Logic*, pages 1–187. Oxford University Press, Oxford, 2004.

E. Mares and F. Paoli. Logical consequence and the paradoxes. *Journal of Philosophical Logic*, 43:439–469, 2014.

R.L. Martin. *Recent Essays on Truth and the Liar Paradox*. Oxford University Press, Oxford, 1984.

P. Martin-Löf. A theory of types. Technical report, 71–3, University of Stockholm, 1971.

P. Martin-Löf. An intuitionistic theory of types. In G. Sambin and J.A. Smith, editors, *Twenty-five years of constructive type theory*, pages 127–172. Oxford University Press, Oxford, 1998.

E. Martino. Russellian type theories and semantical paradoxes. In C. A. Anderson and M. Zelëny, editors, *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*, pages 491–506. Kluwer, Dordrecht, 2001.

V. McGee. *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth*. Hackett Publishing, Indianapolis and Cambridge, 1991.

T. Meadows. Infinitary Tableau for Semantic Truth. *Review of Symbolic Logic*, 8:217–235, 2015.

D. Mirimanoff. Les antinomies de Russell et de Burali-Forti et le problème fondamentale de la théorie des ensembles. *L'Enseignement Mathématique*, 19:37–52, 1917a.

D. Mirimanoff. Remarques sur la théorie des ensembles et les antinomies cantoriennes i. *L'Enseignement Mathématique*, 19:209–217, 1917b.

D. Mirimanoff. Remarques sur la théorie des ensembles et les antinomies cantoriennes ii. *L'Enseignement Mathématique*, 21:29–52, 1920.

S. K. Moh. Logical paradoxes for many-valued systems. *The Journal of Symbolic logic*, 19: 37–39, 1954.

R. Montague. Syntactical treatment of modality, with corollaries on reflection principles and finite axiomatizability. *Acta Philosophica Fennica*, 16:153–167, 1963.

G. H. Moore. The origin of Russell's paradox: Russell, Couturat and the antinomy of infinite number. In J. Hintikka, editor, *From Dedekind to Gödel. Essays on the development of the foundation of mathematics*, pages 215–239. Kluwer, Dordrecht, 1995.

G. H. Moore and A. Garciadiego. Burali-Forti's paradox: a reappraisal of its origins. *Historia Mathematica*, 8:319–350, 1981.

J. Murzi and M. Carrara, editors. Paradox and logical revision. *Topoi*, 34(1):1–131, 2015.

J. Myhill. A system which can define its own truth. *Fundamenta Mathematicae*, 37:190–192, 1950.

C. Nicolai. Necessary truths and supervaluations. In C. de Florio and A. Giordani, editors, *From arithmetic to metaphysics*, (Philosophical Analysis: 73), pages 309–329. De Gruyter, Berlin, 2018.

G. Peano. Additione. In *G. Peano, Opere Scelte (Volume I)*. First published In Revista de Matematica, 8: 143–157, 1906; Roma Cremonese, 1957, 344–358.

J. Pelham and A. Urquhart. Russellian propositions. In D. Prawitz, B. Skyrms, and D. Westerstahl, editors, *Logic, Methodology and Philosophy of Science IX*, pages 307–326. Elsevier, Amsterdam, 1994.

H. Poincaré. Les mathématiques et la logique. *Revue de Métaphysique et de Morale*, 13: 815–835, 1905.

H. Poincaré. Les mathématiques et la logique. *Revue de Métaphysique et de Morale*, 14: 17–34, 1906.

H. Poincaré. La logique de l'infini. *Revue de Métaphysique et de Morale*, 17:461–482, 1909a.

H. Poincaré. Réflexions sur les deux notes precedents. *Acta Mathematica*, 32:195–200, 1909b.

H. Poincaré. Über transfiniten Zahlen. In *Sechs Vorträge über ausgewählte Gegenstände der reinen Mathematik und Physik*, pages 43–48. Teubner, Leipzig-Berlin, 1910.

H. Poincaré. "la logique de l'infini. *Scientia*, 12:1–11, 1912.

W. V. O. Quine. "new foundations for mathematical logic. *American Mathematical Monthly*, 12:70–80, 1937.

L. Rabern, B. Rabern, and M. Macauley. Dangerous reference graphs and semantic paradoxes. *Journal of Philosophical Logic*, 42:727–765, 2013.

F. P. Ramsey. The foundations of mathematics. *Proceedings of the London Mathematical Society (Series 2)*, 25:338–384, 1926.

F. P. Ramsey. *The Foundations of Mathematics and Other Logical Essays*. Routledge and Kegan Paul, London, 1931.

E. H. Reck and R. T. Cook, editors. *Special Issue: Reconsidering Frege's Conception of Number*. Philosophia Mathematica, 24:1–116, 2016.

W. N. Reinhardt. Some Remarks on Extending and Interpreting Theories with a Partial Predicate for Truth. *Journal of Philosophical Logic*, 15:219–251, 1986.

J. Richard. Les principes des mathématique et le problème des ensembles. *Revue Générale des Sciences Pures et Appliquées*, 16:541–543; also in Acta Mathematica, 30: 295–296, 1906; Tradução em inglês In van Heijenoort 1967, 142–144, 1905.

D. Ripley. Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, 5:354–378, 2012.

E. Rivello. Revision without revision sequences: circular definitions. *Journal of Philosophical Logic*, 48:57–85, 2019a.

E. Rivello. Revision without revision sequences: self-referential truth. *Journal of Philosophical Logic*, 48:523–551, 2019b.

J. B. Rosser. An informal exposition of proofs of Gödel's and Church's theorem. *The Journal of Symbolic logic*, 4:53—60, 1939.

L. Rossi. Adding a conditional to Kripke's theory of truth. *Journal of Philosophical Logic*, 45: 485–529, 2016.

L. Rossi. A unified theory of truth and paradox. *Review of Symbolic Logic*, 45:209–254, 2019.

B. Russell. *The Principles of Mathematics (Volume 1)*. Cambridge University Press, Cambridge, 1903.

B. Russell. Les paradoxes de la logique. *Revue de Métaphysique et de Morale*, 14:627–650, 1906.

B. Russell. On some difficulties in the theory of transfinite numbers and order types. *Proceedings of the London Mathematical Society (2nd series)*, 4:29–53, 1907.

B. Russell. Mathematical logic as based on the theory of types. *American Journal of Mathematics*, 30:222–262; reprinted in B. Russell, Logic and Knowledge, London: Allen and Unwin, 1956, 59–102; reprinted in van Heijenoort 1967, 152–182, 1908.

K. Sato and J. Walker. Finitist axiomatic truth. *The Journal of Symbolic logic*, 88(1):22–73, 2023.

T. Schindler and T. Beringer. Reference graphs and semantic paradox. In A. Arazim and M. Dancak, editors, *Logica Yearbook 2015*, pages 1–15. College Publications, London, 2016.

A. Schönflies. Über die logischen paradoxien der mengenlehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 15:19–25, 1906.

K. Schütte. *Beweistheorie*. Springer, Berlin-Heidelberg, 1960.

D. Scott. Continuous lattices. In W. Lawvere, editor, *λ-Toposes, Algebraic Geometry and Logic*, (Lecture Notes in Mathematics: 274), pages 97–136. Springer, Berlin-Heidelberg, 1972.

D. Scott. Combinators and classes. In C. Böhm, editor, *λ-calculus and computer science*, (Lecture Notes in Computer Science: 37), pages 1–26. Springer, Berlin, 1975.

M. Sheard. A guide to truth predicates in the modern era. *The Journal of Symbolic logic*, 59: 132–154, 1994.

K. Simmons. *Universality and the Liar*. Cambridge University Press, New York, 1993.

E. Specker. Typical ambiguity. In E. Nagel, P. Suppes, and A. Tarski, editors, *Logic, Methodology and Philosophy of Science*, pages 116–123. Stanford University Press, Stanford, 1962.

S. Standefer. Solovay-type theorems for circular definitions. *Review of Symbolic Logic*, 8: 467–487, 2015.

S. Standefer. Contraction and revision. *Australasian Journal of Logic*, 13:58–77, 2016.

S. Standefer. Non-classical circular definitions. *Australasian Journal of Logic*, 14:147–180, 2017.

S. Standefer. Proof theory for functional modal logic. *Studia Logica*, 106:49–84, 2018.

A. Tarski. Sur les ensembles définissables de nombres réels i. *Fundamenta Mathematicae*, 17:210–239, 1931.

A. Tarski. The concept of truth in the languages of the deductive sciences (Polish). *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizyczych*, 34, 1933. Reprinted in Zygmunt, J. (ed.), Alfred Tarski, Pisma Logiczno-Filozoficzne, 1 Prawda, Warsaw: Wydawnictwo Naukowe PWN, 1995, 13–172.

A. Tarski. Der Wahrheitsbegriff in der formalisierten Sprachen. *Studia Philosophica*, 1:261–405, 1935. Versão estendida da tradução em alemão de Tarski 1933; Tradução em inglês In A. Tarski, Logic, Semantics, Metamathematics, 2nd edition, Indianapolis: Hackett, 1983, 152–278.

K. Terui. Light Affine Set Theory: a naïve set theory of polynomial time. *Studia Logica*, 77: 9–40, 2004.

R. Thomason. A note of syntactical treatments of modality. *Synthèse*, 44:391–395, 1980.

J. van Heijenoort. *From Frege to Gödel. A source book in mathematical logic 1879–1931*. Harvard University Press, Cambridge, MA, 1967.

A. Visser. Semantics and the liar paradox. In *Handbook of Philosophical Logic*, (Volume IV), pages 617–706. Kluwer, Dordrecht, 1989.

J. von Neumann. Eine axiomatisierung der mengenlehre. *Journal für die reine und angewandte Mathematik*, 154:219–240; Com correções In Volume 155 (1926), p. 128, 1925.

H. Wang. Undecidable sentences generated by semantic paradoxes. *The Journal of Symbolic logic*, 20:31–43, 1955.

P. Welch. On gupta-belnap revision theories of truth, kripkean fixed points and the next stable. *Bulletin of Symbolic Logic*, 7:345–360, 2001.

P. Welch. Ultimate truth vis-a-vis, stable truth. *Review of Symbolic Logic*, 1:126–142, 2008.

P. Welch. Games for truth. *Bulletin of Symbolic Logic*, 15:410–427, 2009.

P. Welch. Weak systems of determinacy and arithmetical quasi-inductive definitions. *The Journal of Symbolic logic*, 76:418–436, 2011.

P. Welch. Some observations on truth hierarchies. *Review of Symbolic Logic*, 7:1–30, 2014.

P. Welch. Rethinking revision. *Journal of Philosophical Logic*, 48:137–154, 2019.

E. Weydert. *How to approximate the naive comprehension scheme inside classical logic*. Dissertation, Rheinische Friedrich-WilhelmsUniversität, Bonn, 1988; Published in Bonner Mathematische Schriften, 194, Universität Bonn, Mathematisches Institut, Bonn, 1989.

H. Weyl. Über die Definitionen der mathematischen Grundbegriffe. *Mathematisch-naturwissenschaftliche Blätter*, 7:93–95; 109–113, 1910.

H. Weyl. *Das Kontinuum*. Veit, Leipzig, 1918; New York, Dover, 1994.

A. N. Whitehead. Über die Stellung der Definition in der Axiomatik. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 20:222–255, 1911.

A. N. Whitehead and B. Russell. *Principia Mathematica (3 volumes)*. Cambridge University Press, Cambridge, 1910, 1912, 1913; 2nd edition, 1925 (Vol. 1), 1927 (Vols. 2 and 3).

J. Wolenski. On tarski's background. In J. Hintikka, editor, *From Dedekind to Gödel, Essays on the development of the foundation of mathematics*, pages 331–341. Kluwer, Dordrecht, 1995.

S. Yablo. Paradox without self-reference. *Analysis*, 53:251–252, 1993.

S. Yablo. Circularity and paradox. In T. Bolander, V. F. Hendricks, and S. A. Pedersen, editors, *Self-Reference*, pages 165–184. Center for the Study of Language and Information

Publications, Stanford, 2006.

S. Yuting. The paradox of the class of all grounded sets. *The Journal of Symbolic logic*, 18: 114, 1953.

E. Zardini. Truth without contra(di)ction. *Review of Symbolic Logic*, 4:498–535, 2011.

E. Zermelo. Untersuchungen über die Grundlagen der Mengenlehre i. *Mathematische Annalen*, 65:261–281, 1908.

J. Zygmunt, editor. *Alfred Tarski, Pisma Logiczno-Filozoficzne, 1 Prawda*. Wydawnictwo Naukowe PWN, Warsaw, 1995.

(II) Autorreferência e Paradoxo[1]

Título Original: Self-Reference and Paradox
Autores: Thomas Bolander
Tradução: Benedito Monteiro
Revisão: Bruna Silva

No contexto da linguagem, a *autorreferência* é usada para denotar uma afirmação que se refere a si mesma ou ao seu próprio referente. O exemplo mais famoso de uma sentença autorreferencial é a *sentença do mentiroso*: "Esta sentença não é verdadeira". A autorreferência também é frequentemente usada em um contexto mais amplo. Por exemplo, uma imagem pode ser considerada autorreferencial se contiver uma cópia de si mesma (veja a imagem acima)[2]; e uma obra literária pode ser considerada autorreferencial se incluir uma referência à própria obra. Na filosofia, a autorreferência é de interesse específico em relação à análise da linguagem, mas também à análise da mente. A autorreferência também é um campo de especial interesse em matemática e ciência da computação, particularmente em relação aos fundamentos dessas ciências.

Também falamos de autorreferência ao usarmos o pronome de primeira pessoa "eu", como em "Eu sinto dor" ou "Eu conheço essa música". A autorreferência no sentido de um

[1] BOLANDER, Thomas, "Self-Reference and Paradox", In: ZALTA, E. N.; NODELMAN, U. (eds.). The Stanford Encyclopedia of Philosophy (Fall 2024 Edition). Stanford, CA: The Metaphysics Research Lab, 2024. Disponível em: https://plato.stanford.edu/archives/fall2024/entries/self-reference/.
A seguir está a tradução da entrada sobre Autorreferência e Paradoxo de Thomas Bolander na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/fall2024/entries/self-reference/. Esta versão traduzida pode diferir da versão atual da entrada, que pode ter sido atualizada desde o momento desta tradução. A versão atual está localizada em https://plato.stanford.edu/entries/self-reference/. Agradecemos aos editores Edward N. Zalta e Uri Nodelman pela permissão para traduzir e publicar esta entrada.

[2] A imagem mencionada pelo autor consiste em um loop infinito, em que a reflexão espelhada se repete recursivamente, criando um efeito de repetição sem fim. (N. T.)

agente referir-se a si mesmo, assim como a diferença entre a autoatribuição e autoidentificação em relação à atribuição ou identificação de outros, é abordada em várias outras entradas desta enciclopédia, incluindo autoconhecimento[3], conhecimento de si[4], autoconsciência[5], epistemologia de Descartes[6] e introspecção[7].

Uma parte importante do interesse filosófico na autorreferência está centrada nos paradoxos. Um *paradoxo* é um raciocínio aparentemente sólido, baseado em suposições aparentemente verdadeiras, que ainda levam a uma contradição (Quine, 1976). Considere novamente a sentença do mentiroso mencionada anteriormente. É uma sentença L que expressa: "A sentença L não é verdadeira." Somos levados a uma contradição quando tentamos determinar se L é verdadeira ou não. Se primeiro assumirmos que L é verdadeira, então ele deve estar expressando uma afirmação verdadeira sobre o mundo. Como L expressa "A sentença L não é verdadeira", agora temos que L não é verdadeira, o que é uma contradição. Suponha, contrariamente, que L não seja verdadeira. Então, a expressão "A sentença L não é verdadeira" é verdadeira. Mas a expressão entre aspas é exatamente a afirmação expressa por L, então L deve ser verdadeira, novamente uma contradição. Assim, independentemente de assumirmos que L é verdadeira, seja assumindo que não, chegamos a uma contradição. Portanto, agora temos uma contradição obtida por meio de um raciocínio aparentemente sólido, baseado em suposições aparentemente verdadeiras. Isso, portanto, qualifica-se como um paradoxo. Esse paradoxo é conhecido como o *paradoxo do mentiroso*. A sentença do mentiroso leva a um paradoxo *porque* é autorreferencial, mas a autorreferência não é uma condição suficiente para a paradoxalidade. A *sentença veraz* "Esta sentença é verdadeira" não é paradoxal, e tampouco a sentença "Esta sentença contém quatro palavras" (embora seja falsa).

A maioria dos paradoxos de autorreferência pode ser categorizada como *semântica*,

[3]Gertler, Brie, "Self-Knowledge", The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2024/entries/self-knowledge.

[4]Gertler, Brie, "Self-Knowledge", The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2024/entries/self-knowledge/.

[5]Smith, Joel, "Self-Consciousness", The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2024/entries/self-consciousness/.

[6]Newman, Lex, "Descartes' Epistemology", The Stanford Encyclopedia of Philosophy (Winter 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/win2023/entries/descartes-epistemology.

[7]Schwitzgebel, Eric, "Introspection", The Stanford Encyclopedia of Philosophy (Fall 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/fall2024/entries/introspection/.

conjuntista ou *epistêmicos*. Os paradoxos semânticos, como o paradoxo do mentiroso, são principalmente relevantes para as teorias da verdade. Os paradoxos conjuntistas são relevantes para os fundamentos da matemática, e os paradoxos epistêmicos são relevantes à epistemologia. Embora esses paradoxos sejam diferentes no assunto no qual se relacionam, eles compartilham a mesma estrutura subjacente e, muitas vezes, podem ser abordados usando os mesmos meios matemáticos.

Nesta introdução, apresentaremos primeiramente alguns dos paradoxos de autorreferência mais conhecidos e discutiremos a estrutura subjacente em comum. Em seguida, abordaremos as profundas consequências que esses paradoxos têm em diferentes áreas: teorias da verdade, teoria dos conjuntos, epistemologia, fundamentos da matemática e computabilidade. Por fim, apresentaremos as abordagens mais proeminentes para a resolução desses paradoxos.

1. Paradoxos da Autorreferência

1.1 Paradoxos Semânticos

Os paradoxos de autorreferência são conhecidos desde a antiguidade. A descoberta do paradoxo do mentiroso é frequentemente creditada a Eubúlides de Mégara, que viveu no século IV a.C. O paradoxo do mentiroso pertence à categoria dos paradoxos *semânticos*, já que é baseado na noção semântica de verdade. Outros paradoxos semânticos bem conhecidos incluem o paradoxo de Grelling, o paradoxo de Berry e o paradoxo de Richard.

O *paradoxo de Grelling* envolve um predicado definido da seguinte maneira. Diz-se que um predicado é *heterológico* se não é verdadeiro em relação a si mesmo, isto é, se ele próprio não possui a propriedade que expressa. Assim, o predicado "alemão" é heterológico, uma vez que não é uma palavra alemã, mas o predicado "deutsch" não é heterológico. A questão que leva ao paradoxo é:

> O "heterológico" é heterológico?

É fácil perceber que chegamos a uma contradição, independentemente de respondermos "sim" ou "não" a essa questão (o raciocínio segue de maneira semelhante ao paradoxo do mentiroso). O paradoxo de Grelling é autorreferencial, pois a definição do predicado "heterológico" se refere a *todos* os predicados, incluindo o próprio predicado heterológico.

Definições como essa, que dependem de um conjunto de entidades, sendo que pelo menos uma delas é a entidade que está sendo definida, são chamadas de *impredicativas*.

O *paradoxo de Berry* é outro paradoxo baseado em uma definição impredicativa, ou melhor, em uma descrição impredicativa. Algumas expressões da língua inglesa são descrições de números naturais, por exemplo, “a soma de cinco e sete” é uma descrição do número 12. O paradoxo de Berry surge ao tentar determinar a denotação da seguinte descrição:

> O menor número que não pode ser referido por uma descrição contendo menos de 100 símbolos.

A contradição é que essa descrição com 93 símbolos denota um número que, por definição, não pode ser denotado por nenhuma descrição contendo menos de 100 símbolos. É claro que a descrição é impredicativa, já que se refere implicitamente a todas as descrições, incluindo a si mesma.

O *paradoxo de Richard* considera frases da língua inglesa que definem números reais, em vez de números naturais. Por exemplo, “a razão entre a circunferência e o diâmetro de um círculo” é uma frase que define o número π. Suponha que uma enumeração de todas essas frases seja fornecida (por exemplo, colocando-as em ordem lexicográfica). Agora considere a frase:

> O número real cujo *n*-ésimo decimal é 1 sempre que o *n*-ésimo decimal do número denotado pela *n*-ésima frase for 0; caso contrário, é 0.

Essa frase define um número real, então deve estar entre as frases enumeradas, digamos que seja o número k nessa enumeração. Mas, ao mesmo tempo, por definição, ele difere do número denotado pela k-ésima frase no k-ésimo decimal. Assim, temos uma contradição. A frase definidora é obviamente impredicativa. A construção particular empregada nesse paradoxo é chamada de *diagonalização*. A diagonalização é uma construção geral e um método de prova originalmente inventado por Georg Cantor (1891) para provar a não-enumerabilidade do conjunto das partes dos números naturais. Ela também foi usada como base para o paradoxo de Cantor, um dos paradoxos da teoria dos conjuntos que será considerado a seguir.

1.2 Paradoxos da Teoria dos Conjuntos

Os paradoxos mais conhecidos da teoria dos conjuntos são o paradoxo de Russell e o paradoxo de Cantor. O *paradoxo de Russell* surge ao considerar o *conjunto de Russell* R de todos os conjuntos que não são membros de si mesmos, ou seja, o conjunto definido por $R = \{x \mid x \notin x\}$. A contradição é derivada ao perguntar se R é um membro de si mesmo, isto é, se $R \in R$ é verdadeiro. Se $R \in R$, então $R \in R$ é um membro de si mesmo e, portanto, $R \notin R$, pela definição de R. Se, por outro lado, $R \notin R$, então R não é um membro de si mesmo e, assim, $R \in R$, novamente pela definição de R.

O *paradoxo de Cantor* é baseado em uma aplicação do teorema de Cantor. O *teorema de Cantor* afirma que, dado qualquer conjunto finito ou infinito S, o conjunto das partes de S tem uma cardinalidade (tamanho maior) estritamente maior do que S. O teorema é provado por uma forma de diagonalização, a mesma ideia subjacente ao paradoxo de Richard. O paradoxo de Cantor considera o conjunto de todos os conjuntos. Chamemos esse conjunto de *conjunto universal* e denotemos-o por U. O conjunto das partes de U é denotado por $\mathcal{P}(U)$. Como U contém *todos* os conjuntos, ele conterá, em particular, todos os elementos de $\mathcal{P}(U)$. Assim, $\mathcal{P}(U)$ deve ser um subconjunto de U e deve, portanto, ter uma cardinalidade (tamanho) que é menor ou igual à cardinalidade de U. No entanto, isso contradiz imediatamente o teorema de Cantor.

O *paradoxo do hiperjogo* é uma adição mais recente à lista de paradoxos da teoria dos conjuntos, inventado por Zwicker (1987). Chamaremos um jogo de dois jogadores de *bem-fundamentado* se ele tiver que terminar em um número finito de jogadas. O xadrez em torneio é um exemplo de jogo bem-fundamentado. Agora definimos o *hiperjogo* como o jogo em que o jogador 1, na primeira jogada, escolhe um jogo bem-fundamentado a ser jogado, e o jogador 2, subsequentemente, faz a primeira jogada no jogo escolhido. Todas as jogadas restantes são, então, jogadas do jogo escolhido. O hiperjogo deve ser um jogo bem-fundamentado, uma vez que qualquer partida durará exatamente uma jogada a mais do que algum jogo bem-fundado dado. No entanto, se o hiperjogo é bem-fundamentado, então ele deve ser um dos jogos que podem ser escolhidos na primeira jogada do hiperjogo, ou seja, o jogador 1 pode escolher o hiperjogo na primeira jogada. Isso permite que o jogador 2 escolha o hiperjogo na jogada subsequente, e os dois jogadores podem continuar escolhendo o hiperjogo *ad infinitum*. Assim, o hiperjogo não pode ser bem-fundamentado, contradizendo nossa conclusão anterior.

1.3 Paradoxos Epistêmicos

O paradoxo epistêmico mais conhecido é o *paradoxo do conhecedor* (Kaplan & Montague, 1960; Montague, 1963). Esse paradoxo tem muitas formulações equivalentes, uma delas baseada na frase “Esta sentença não é conhecida por ninguém”. Vamos chamar essa frase de sentença do conhecedor, abreviada como KS. KS é obviamente bastante semelhante à sentença do mentiroso, exceto que o conceito central envolvido é o conhecimento, em vez da verdade. O raciocínio que leva a uma contradição a partir de KS é um pouco mais complexo do que no paradoxo do mentiroso. Primeiro, KS é demonstrada como verdadeira pelo seguinte raciocínio:

> Assuma, para obter uma contradição, que KS não é verdadeira. Então, o que KS expressa não pode ser o caso, ou seja, KS deve ser conhecida por alguém. Como tudo o que é conhecido é verdadeiro (isso faz parte da definição do conceito de conhecimento), KS é verdadeira, o que contradiz nossa suposição. Isso conclui a prova de que KS é verdadeira.

A parte do raciocínio que acabou de ser realizado para provar a veracidade de KS deve estar disponível para qualquer agente (pessoa) com capacidades de raciocínio suficientes. Ou seja, um agente deve ser capaz de provar a veracidade de KS e, assim, saber que ao conhecimento que KS é válido. No entanto, se KS é conhecida por alguém, então o que expressa não é o caso, e, portanto, não pode ser verdadeira. Isso é uma contradição, e assim temos um paradoxo. O papel da autorreferência nesse paradoxo é óbvio, já que se baseia em uma sentença, KS, referindo-se diretamente a si mesma.

O paradoxo do conhecedor é apenas um dos muitos paradoxos epistêmicos envolvendo autorreferência. Veja o artigo sobre paradoxos epistêmicos[8] para mais informações sobre a classe dos paradoxos epistêmicos. Um paradoxo epistêmico mais recente, situado no contexto de crenças e suposições em um jogo de dois jogadores, é o paradoxo de Brandenburger-Keisler (Brandenburger & Keisler, 2006), descrito em detalhes no artigo sobre fundamentos epistêmicos da teoria dos jogos[9]. Para uma discussão detalhada e história

[8] Sorensen, Roy, “Epistemic Paradoxes”, The Stanford Encyclopedia of Philosophy (Fall 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/fall2024/entries/epistemic-paradoxes/.
[9] Pacuit, Eric and Olivier Roy, “Epistemic Foundations of Game Theory”, The Stanford Encyclopedia of Philosophy (Fall 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/fall2024/entries/epistemic-game/.

dos paradoxos de autorreferência em geral, veja o artigo sobre paradoxos e lógica contemporânea[10].

1.4 Estruturas Comuns nos Paradoxos

Os paradoxos acima são todos bastante semelhantes em sua estrutura. No caso dos paradoxos de Grelling e Russell, isso pode ser visto da seguinte maneira. Defina a *extensão* de um predicado como o conjunto de objetos dos quais ele é verdadeiro. Para um predicado P, denotamos sua extensão por $\text{ext}(P)$. O paradoxo de Grelling envolve o predicado heterológico, que é verdadeiro para todos aqueles predicados que não são verdadeiros para si mesmos. Assim, a extensão do predicado heterológico é o conjunto $\{P \mid P \notin \text{ext}(P)\}$. Compare isso ao conjunto de Russell R dado por $\{x \mid x \notin x\}$. A única diferença significativa entre esses dois conjuntos é que o primeiro é definido em predicados, enquanto o segundo é definido em conjuntos. As provas de contradições baseadas nesses dois conjuntos também compartilham a mesma estrutura, conforme visto abaixo (onde "het" é a abreviação de "heterológico"):

$$\begin{aligned}
\text{(Grelling)} \quad & \text{het} \in \text{ext}(\text{het}) \Leftrightarrow \text{het} \in \{P \mid P \notin \text{ext}(P)\} \\
& \Leftrightarrow \text{het} \notin \text{ext}(\text{het}). \\
\text{(Russell)} \quad & R \in R \Leftrightarrow R \in \{x \mid x \notin x\} \\
& \Leftrightarrow R \notin R
\end{aligned}$$

Aqui temos dois paradoxos de uma estrutura quase idêntica pertencentes a duas classes distintas de paradoxos: um é semântico e o outro é teórico-conjuntista. O que isso nos ensina é que, mesmo que os paradoxos pareçam diferentes por envolverem assuntos diferentes, eles podem ser quase idênticos em sua estrutura subjacente. Assim, em muitos casos, faz mais sentido estudar os paradoxos da autorreferência como um único tipo, em vez de estudá-los, digamos, os paradoxos semânticos e os paradoxos da teoria dos conjuntos separadamente.

[10]Cantini, Andrea and Riccardo Bruni, "Paradoxes and Contemporary Logic", The Stanford Encyclopedia of Philosophy (Fall 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/fall2024/entries/paradoxes-contemporary-logic/.

Os paradoxos de Russell e Cantor também são mais semelhantes do que parecem à primeira vista. O paradoxo de Cantor é baseado em uma aplicação do teorema de Cantor ao conjunto universal U (cf. Seção 1.2 acima). Abaixo, damos a prova do teorema de Cantor para um conjunto arbitrário S.

> Precisamos provar que $\mathcal{P}(S)$ tem maior cardinalidade que S. Assumimos, para obter uma contradição, que isso não é o caso. Então deve existir uma função (potencialmente parcial) f de S sobre $\mathcal{P}(S)$. Agora considere o conjunto $C = \{x \in \text{dom}(f) \mid x \notin f(x)\}$. Claramente, $C \subseteq S$ então $C \in \mathcal{P}(S)$. Como f é sobrejetora em $\mathcal{P}(S)$, deve então existir um conjunto $c \in \text{dom}(f)$ tal que $f(c) = C$. No entanto, agora obtemos uma contradição, pois o seguinte vale:

$$\begin{aligned} c \in f(c) &\Leftrightarrow c \in \{x \in \text{dom}(f) \mid x \notin f(x)\} \\ &\Leftrightarrow c \notin f(c). \end{aligned}$$

Note a semelhança entre essa sequência de equivalências e as sequências de equivalências correspondentes derivadas para os paradoxos de Russell e Grelling acima. Agora considere o caso especial do teorema de Cantor onde S é o conjunto universal. Então podemos simplesmente escolher f como a função identidade em $\mathcal{P}(S)$, já que S é o conjunto universal e, portanto, $\mathcal{P}(S) \subseteq S$ (qualquer conjunto deve ser um subconjunto do conjunto universal). Então, f é a função parcial $f : \mathcal{P}(S) \to \mathcal{P}(S)$ definida por $f(x) = x$. Mas então C acima se torna o conjunto de Russell, e a sequência de equivalências se torna a prova da contradição no paradoxo de Russell! Assim, o paradoxo de Cantor nada mais é do que uma ligeira variação do paradoxo de Russell; o argumento central que leva à contradição é o mesmo em ambos.

Priest (1994) fornece evidências ainda mais firmes sobre a semelhança entre os paradoxos da autorreferência ao mostrar que todos eles se encaixam no que ele originalmente chamou de *Esquema de Russell Qualificado*, agora denominado *Esquema de Inclusão*. A ideia por trás disso remonta ao próprio Russell (1905), que também considerou que os paradoxos da autorreferência têm uma estrutura subjacente comum. Dado dois predicados P e Q, e uma função possivelmente parcial δ, o Esquema de Inclusão consiste nas seguintes duas condições:

1. $w = \{x \mid P(x)\}$ existe e $Q(w)$ é verdadeiro;
2. Se y é um subconjunto de w tal que $Q(y)$ é verdadeiro, então:
    2.1. $\delta(y) \notin y$,
    2.2. $\delta(y) \in w$.

Se essas condições forem satisfeitas, temos a seguinte contradição: como w é trivialmente um subconjunto de w e como $Q(w)$ é verdadeiro pela condição 1, temos tanto $\delta(w) \notin w$ quanto $\delta(w) \in w$, pelas condições 2a e 2b, respectivamente. Assim, qualquer tripla (P, Q, δ) que satisfaça o Esquema de Inclusão produzirá um paradoxo. Priest mostra como a maioria dos paradoxos bem conhecidos da autorreferência se encaixam no esquema. Abaixo, consideraremos apenas alguns desses paradoxos, começando com o paradoxo de Russell. Nesse caso, definimos a tripla (P, Q, δ) da seguinte forma:

- $P(x)$ é o predicado "$x \notin x$".
- $Q(y)$ é o predicado universal verdadeiro para qualquer objeto.
- δ é a função identidade.

Então, w no Esquema de Inclusão se torna o conjunto de Russell e a contradição obtida a partir do esquema se torna o paradoxo de Russell.

No caso do paradoxo de Richard, definimos a tripla da seguinte maneira:

- $P(x)$ é o predicado "x é um real definível por uma frase em inglês."
- $Q(y)$ é o predicado "y é um conjunto denumerável de reais definíveis por uma frase em inglês."
- δ é a função que mapeia qualquer conjunto denumerável y de reais para o real z, cujo enésimo decimal é 1 sempre que o enésimo decimal do enésimo real em y for 0; caso contrário, é 0. (Qualquer enumeração dos elementos em y servirá.)

Aqui, $w = \{x \mid P(x)\}$ torna-se o conjunto de todos os reais definíveis por frases em inglês. Para qualquer subconjunto denumerável y de w, $\delta(y)$ é um real que, por construção, diferirá de todos os reais em y (difere do enésimo real em y no enésimo decimal). Se deixarmos y igual a w, então obtemos $\delta(w) \notin w$. No entanto, ao mesmo tempo, $\delta(w)$ é definível por uma frase em inglês, então $\delta(w) \in w$, e temos uma contradição. Essa contradição é o paradoxo de Richard.

O paradoxo do mentiroso também se encaixa no esquema de Russell, embora de uma maneira um pouco menos direta:

- $P(x)$ é o predicado "x é verdadeiro."
- $Q(y)$é o predicado "y é definível."
- $\delta(y)$ é a frase "esta frase não pertence ao conjunto y".

Aqui, $w = \{x \mid P(x)\}$ torna-se o conjunto de frases verdadeiras, e $\delta(w)$ torna-se uma versão da sentença do mentiroso: "esta frase não pertence ao conjunto de frases verdadeiras".

Do exposto acima, pode-se concluir que todos, ou pelo menos a maioria, dos paradoxos de autorreferência compartilham a mesma estrutura subjacente - independente de serem semânticos, conjuntistas ou epistêmicos. Priest chama isso de *princípio da solução uniforme*: "mesmo tipo de paradoxo, mesmo tipo de solução." No entanto, se o Esquema de Inclusão pode, de forma geral, ser considerado uma condição necessária e suficiente para a paradoxalidade autorreferencial é discutível (Slater, 2002; Abad, 2008; Badici, 2008; Zhong, 2012, entre outros), de modo que nem todos os autores concordam com o princípio da solução uniforme.

O paradoxo sorites[11] é um paradoxo que, à primeira vista, não envolve autorreferência de forma alguma. No entanto, Priest (2010b, 2013) argumenta que ele ainda se encaixa no Esquema de Inclusão e, portanto, pode ser visto como um paradoxo de autorreferência, ou pelo menos um paradoxo que deveria ter o mesmo tipo de solução que os paradoxos de autorreferência. Isso levou Colyvan (2009), Priest (2010) e Weber (2010b) a propor uma abordagem dialeteica para resolver o paradoxo sorites. Essa abordagem ao paradoxo sorites foi atacada por Beall (2014a, 2014b) e defendida por Weber et al. (2014). Cobreros et

[11] Hyde, Dominic and Diana Raffman, "Sorites Paradox", The Stanford Encyclopedia of Philosophy (Summer 2018 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/sum2018/entries/sorites-paradox/.

al. (2015) investigam a noção de *consequência permissiva* com o objetivo de oferecer um tratamento unificado para os paradoxos da vagueza (como o paradoxo sorites) e os paradoxos de autorreferência. A relação de consequência permissiva é uma versão enfraquecida da relação de consequência clássica no contexto da lógica multivalorada: ela apenas exige que, quando as premissas assumem o valor 1 (são exclusivamente verdadeiras), então a conclusão não deve assumir o valor 0 (não é exclusivamente falsa). Um diagnóstico unificado mais recente dos paradoxos semânticos e soríticos é feito por Bruni & Rossi (2023), que identificam sua origem em uma forma geral de indiscernibilidade.

1.5 Paradoxos sem Negação

A maioria dos paradoxos considerados até agora envolve a negação de maneira essencial, por exemplo, sentenças que afirmam de si mesmas que *não* são verdadeiras ou conhecíveis. O papel central da negação se tornará ainda mais claro quando formalizarmos os paradoxos de autorreferência na Seção 2 abaixo. O *paradoxo de Curry* é um paradoxo de autorreferência semelhante que, no entanto, não envolve diretamente a negação. Uma variante semântica do paradoxo de Curry vem da seguinte sentença de Curry C: "Se esta sentença é verdadeira, então F", onde F pode ser qualquer afirmação, como por exemplo uma que seja obviamente falsa. Suponha que a sentença de Curry C seja verdadeira. Então, ela expressa um fato verdadeiro, isto é, se C é verdadeira, então F. No entanto, já assumimos que C é verdadeiro, então podemos inferir F, usando o Modus Ponens. Provamos que, se assumirmos que C é verdadeira, então F se segue. Isso é exatamente o que a própria sentença de Curry expressa. Em outras palavras, provamos que a própria sentença de Curry é verdadeira! Mas então também temos que F é verdadeira, e isso é um paradoxo, uma vez que F pode ser qualquer afirmação, incluindo coisas que são claramente falsas. Podemos, por exemplo, provar facilmente que o Papai Noel existe, simplesmente deixando F ser a sentença "Papai Noel existe" (Boolos, 1993; Smullyan, 2006). Em um contexto lógico clássico onde a implicação $C \to F$ é equivalente a $\neg C \vee F$, o paradoxo de Curry ainda envolve implicitamente a negação, mas o paradoxo de Curry ainda é independentemente interessante, uma vez que se sustenta com menos suposições sobre a lógica subjacente do que o paradoxo do mentiroso. Veja a entrada sobre o paradoxo de Curry[12] para mais

[12] Shapiro, Lionel and Jc Beall, "Curry's Paradox", The Stanford Encyclopedia of Philosophy (Winter 2021 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/win2021/entries/curry-paradox/.

detalhes.

1.6 Paradoxos sem Autorreferência

A maioria dos paradoxos clássicos de autorreferência envolve autorreferência *direta*, como no paradoxo do mentiroso, onde uma sentença se refere diretamente a si mesma. No entanto, é fácil construir paradoxos que empregam apenas autorreferência *indireta*, ou seja, sentenças que se referem a outras sentenças, que por sua vez se referem a outras, de modo a formar um loop de volta à sentença original. Um exemplo é o *paradoxo do cartão postal*, muitas vezes atribuído a Philip Jourdain (1879–1919), embora, segundo Roy Sorensen (2003, p. 332), o verdadeiro inventor seja G.G. Berry (1867–1928), o bibliotecário de Oxford a quem também é creditado o paradoxo de Berry mencionado anteriormente. No paradoxo do cartão postal, o lado da frente de um cartão postal diz “a sentença no verso é verdadeira”, enquanto o verso diz “a sentença na frente é falsa”. Para que a sentença na frente seja verdadeira, a sentença no verso precisa ser verdadeira, mas para que a sentença no verso seja verdadeira, a sentença na frente precisa ser falsa. Isso gera uma contradição, alcançada de maneira semelhante ao paradoxo do mentiroso. Existem até exemplos muito anteriores de autorreferência indireta na literatura: o Sofisma 9 nos Sophismata de John Buridan, do século 14 (Buridan [SD], Hughes 1982), é estruturalmente equivalente ao paradoxo do cartão postal.

Em 1985, Yablo conseguiu construir um paradoxo semântico que não envolve autorreferência, nem mesmo autorreferência indireta. Em vez disso, ele consiste em uma cadeia infinita de sentenças, cada sentença expressando a falsidade de todas as sentenças subsequentes. Mais precisamente, para cada número natural i , definimos S_i como a sentença “para todo $j > i$, S_j não é verdadeira”. Podemos, então, derivar uma contradição da seguinte maneira:

> Primeiro, provamos que nenhuma das sentenças S_i pode ser verdadeira. Suponha, para obter uma contradição, que S_i seja verdadeira para algum i. Então, é verdade que “para todo $j > i$, S_j não é verdadeira”. Assim, nenhuma das sentenças S_j para $j > i$ é verdadeira. Em particular, S_{i+1} não é verdadeira. S_{i+1} é a sentença “para todo $j > i + 1$, S_j não é verdadeira”. Como essa sentença não é verdadeira, deve haver algum $k > i + 1$ para o qual S_k seja verdadeira. No entanto, isso contradiz o fato de que nenhuma

das sentenças S_j com $j > i$ é verdadeira.

> Agora provamos que nenhuma das sentenças S_i é verdadeira. Então, em particular, temos que para todo $j > 0$, S_j não é verdadeira. Isso é exatamente o que é expresso por S_0, então S_0 deve ser verdadeira. Novamente, temos uma contradição.

Yablo chama este paradoxo de ω-mentiroso, mas outros geralmente se referem a ele como o paradoxo de Yablo. Note que nenhuma das sentenças S_i se refere a si mesma (nem mesmo indiretamente), mas apenas às sentenças que ocorrem posteriormente na sequência. O paradoxo de Yablo é semântico, mas como mostrado por Yablo (2006), paradoxos similares de teoria dos conjuntos, que não envolvem autorreferência, podem ser formulados em certas teorias dos conjuntos.

O paradoxo de Yablo demonstra que podemos ter paradoxos lógicos sem autorreferência — apenas um certo tipo de estrutura não bem-fundada é necessário para obter uma contradição. Obviamente, há diferenças estruturais entre os paradoxos ordinários de autorreferência e o paradoxo de Yablo: os paradoxos ordinários de autorreferência envolvem uma estrutura cíclica de referência, enquanto o paradoxo de Yablo envolve uma estrutura acíclica, mas não bem-fundada, de referência. Mais precisamente, podemos pensar a estrutura referencial subjacente a um paradoxo como um grafo direcionado. Os vértices desse grafo são sentenças, e há uma aresta direcionada da sentença S para a sentença T se S se refere diretamente a T. A estrutura referencial do mentiroso é, então, um grafo com um único laço reflexivo. A estrutura referencial do paradoxo do cartão-postal é um grafo cíclico com 2 vértices, cada um tendo uma aresta direcionada para o outro vértice. Todos os paradoxos de autorreferência direta ou indireta possuem estruturas cíclicas de referência (seus grafos subjacentes são cíclicos). Isso é diferente no paradoxo de Yablo. A estrutura referencial no paradoxo de Yablo é isomórfica à ordenação usual de menor que nos números naturais, que é uma ordem total estrita (não contém ciclos). Mesmo havendo essa diferença, o paradoxo de Yablo ainda compartilha a maioria das propriedades com os paradoxos ordinários de autorreferência. Quando resolvemos paradoxos, podemos então optar por considerá-los todos em um só e nos referirmos a ele como *paradoxos de não bem-fundamentação*. No entanto, a seguir, vamos nos ater ao termo *paradoxos de autorreferência*, embora grande parte do que dissermos também se aplique ao paradoxo de Yablo e a paradoxos relacionados de não bem-fundamentação.

Dado o insight de que não apenas as estruturas cíclicas de referência podem levar a paradoxos, mas também certos tipos de estruturas não bem-fundadas, torna-se interessante estudar mais a fundo essas estruturas de referência e seu potencial para caracterizar as condições necessárias e suficientes para a paradoxicalidade. Essa linha de trabalho foi iniciada por Gaifman (1988, 1992, 2000) e mais tarde continuada por Cook (2004), Walicki (2009) e outros.

Uma quantidade significativa de trabalhos mais mais recentes sobre autorreferência têm se concentrado em tentar fazer uma caracterização completa, do ponto de vista da teoria dos grafos, de quais estruturas de referência admitem paradoxos, incluindo Rabern e Macauley (2013), Cook (2014), e Dyrkolbotn e Walicki (2014). Uma caracterização completa ainda é um problema em aberto (Rabern, Rabern e Macauley, 2013), mas parece ser uma conjectura relativamente difundida que todos os grafos paradoxais de referência são ou cíclicos ou contêm uma estrutura do tipo Yablo. A conjectura foi confirmada para certas subclasses de grafos infinitos (Walicki, 2019), mas ainda está em aberto se ela vale para grafos arbitrários. Se a conjectura for de fato verdadeira, isso significa que, em termos de estruturas de referência, todos os paradoxos de referência são ou do tipo mentiroso ou do tipo Yablo. O que exatamente significa para uma estrutura (grafo) ser do tipo Yablo pode ser definido de várias maneiras diferentes, mas equivalentes, incluindo: 1) o grafo contém o grafo de referência do paradoxo de Yablo, $(\omega, <)$, como um menor finitário (Beringer & Schindler, 2017); 2) o grafo contém um raio (um caminho infinito) no qual há infinitos vértices, cada um com infinitos caminhos disjuntos para infinitos outros vértices no raio (Walicki, 2019).

Embora a estrutura de referência envolvida no paradoxo de Yablo não contenha ciclos (cada sentença refere-se apenas a sentenças posteriores na sequência), ainda é debatido se o paradoxo é autorreferencial ou não (Cook, 2014; Halbach e Zhang, 2017). O próprio Yablo (1993) argumenta que não é autorreferencial, enquanto Priest (1997) defende que é autorreferencial. Butler (2017) afirma que, mesmo que Priest esteja correto, haverá outros paradoxos semelhantes ao de Yablo que não são autorreferenciais no sentido de Priest. Na análise do paradoxo de Yablo, é essencial notar que ele envolve uma sequência *infinita* de sentenças, em que cada sentença se refere a *infinitas* outras sentenças. Para formalizá-lo em um contexto de lógica proposicional, é necessário usar lógica proposicional infinitária (ver o verbete sobre lógica infinitária)[13]. Qualquer variante finita da sequência de Yablo—onde

[13] Bell, John L., "Infinitary Logic", The Stanford Encyclopedia of Philosophy (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/fall2023/entries/logic-infinitary/.

cada sentença se refere a apenas um número finito de sentenças posteriores—deve ser necessariamente consistente (não paradoxal) devido ao teorema da compacidade na lógica proposicional (cada subconjunto finito de sentenças na sequência induz uma relação de referência bem-fundada, e as sentenças podem, assim, ser atribuídas de maneira consistente com valores de verdade de baixo para cima). Na aritmética de primeira e segunda ordem finita, pode-se tentar formalizar o paradoxo de Yablo por meio de um predicado unário $S(x)$, onde, para cada número natural i, $S(i)$ representa a formalização da i-ésima sentença S_i da sequência de Yablo (onde i é o numeral que representa i). Como e se o paradoxo de Yablo pode ser representado de forma verdadeira dessa maneira, e como isso se relaciona à compacidade da lógica subjacente, foi investigado por Picollo (2013).

O paradoxo de Yablo também inspirou a criação de paradoxos semelhantes envolvendo estruturas de referência acíclicas e não-bem-fundadas em outras áreas que não a verdade, por exemplo, a variante "yablosca" do paradoxo de Brandenburger-Keisler na teoria dos jogos epistêmicos por Başkent (2016), uma variante relacionada à provabilidade por Cieśliński e Urbaniak (2013), e uma variante no contexto dos teoremas da incompletude de Gödel por Leach-Krouse (2014).

2. Por que os Paradoxos Importam

Depois de apresentarmos vários paradoxos de autorreferência e discutirmos algumas de suas semelhanças subjacentes, passaremos agora a uma discussão de seu significado. A relevância de um paradoxo está na indicação de uma falha ou deficiência em nossa compreensão dos conceitos centrais envolvidos. No caso dos paradoxos semânticos, parece que nossa compreensão de conceitos semânticos fundamentais, como a verdade (no paradoxo do mentiroso e no paradoxo de Grelling) e a definibilidade (nos paradoxos de Berry e de Richard), é deficiente. No caso dos paradoxos da teoria dos conjuntos, o que está em questão é nossa compreensão do conceito de conjunto. Se compreendêssemos completamente esses conceitos, deveríamos ser capazes de lidar com eles sem sermos levados a contradições.

Para ilustrar isso, considere o paradoxo clássico de Zenão sobre *Aquiles e a tartaruga* (ver a entrada Paradoxos de Zenão para mais detalhes[14]). Nesse paradoxo, parece que po-

[14] Huggett, Nick, "Zeno's Paradoxes", The Stanford Encyclopedia of Philosophy (Fall 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/fall2024/entries/paradox-zeno/.

demos provar que a tartaruga pode vencer uma corrida contra Aquiles, que é 10 vezes mais rápido, se lhe for dado uma pequena vantagem inicial. Zenão usou esse paradoxo como um argumento contra a possibilidade de movimento. Posteriormente, descobriu-se que o paradoxo repousa sobre uma compreensão inadequada do conceito de infinito. Mais precisamente, baseia-se na suposição implícita de que qualquer série infinita de números reais positivos deve ter uma soma infinita. Os desenvolvimentos posteriores da matemática das séries infinitas mostraram que essa suposição é inválida, e assim o paradoxo se dissolve. A aceitação original do argumento de Zenão como um paradoxo era um sintoma de que o conceito de infinito não era suficientemente bem compreendido na época. Por analogia, parece razoável esperar que a existência de paradoxos semânticos e paradoxos da teoria dos conjuntos seja um sintoma de que os conceitos semânticos e de teoria dos conjuntos envolvidos ainda não são suficientemente bem compreendidos. Ou, pelo menos, que precisamos revisar nossa visão sobre o que consideramos "pressupostos naturais". O paradoxo de Russell se baseia na suposição de que qualquer predicado sobre objetos matemáticos determina um conjunto que consiste exatamente nos objetos que satisfazem o predicado, e o paradoxo do mentiroso se baseia na suposição de que é possível que uma linguagem contenha seu próprio predicado de verdade. Como resposta aos paradoxos, esses pressupostos aparentemente sensatos foram revisados, como veremos na Seção 3 a seguir.

Outra resposta possível poderia ser que é a nossa compreensão do próprio conceito de "contradição" que é falha. O raciocínio envolvido nos paradoxos de autorreferência termina em alguma contradição, uma sentença concluída como sendo ao mesmo tempo verdadeira e falsa. Consideramos isso uma impossibilidade, daí o paradoxo, mas talvez não necessário? O *Dialeteísmo* é a visão de que podem existir "contradições verdadeiras", o que significa que não é impossível que uma sentença seja ao mesmo tempo verdadeira e falsa. Se adotarmos o ponto de vista do dialetismo, todos os paradoxos de autorreferência se dissolvem e passam a ser, em vez disso, provas de existência de certas "dialeteias": sentenças que são tanto verdadeiras quanto falsas. Priest (1987) é um forte defensor do dialeteísmo e utiliza seu princípio de solução uniforme (ver a Seção 1.4 acima) para defender a solução dialeteísta.

Para mais informações, veja as entradas sobre dialeteísmo[15] e lógica paraconsistente[16].

Atualmente, não existe uma solução consensual para os paradoxos de autorreferência. Eles continuam a apresentar problemas fundamentais na semântica e na teoria dos conjuntos. Não se pode afirmar que há uma base sólida para esses temas até que uma solução satisfatória para os paradoxos seja fornecida. Problemas surgem quando se tenta formalizar a semântica (o conceito de verdade) e a teoria dos conjuntos. Se a formalização seguir a compreensão intuitiva e "ingênua" desses temas, sistemas inconsistentes que se prologam como os paradoxos serão formalizáveis nesses sistemas.

2.1 Consequências dos Paradoxos Semânticos

O paradoxo do mentiroso é uma barreira significativa para a construção de teorias formais da verdade, pois gera inconsistências nessas teorias potenciais. Grande parte da pesquisa sobre autorreferência se concentra em teorias formais da verdade e em maneiras de contornar o paradoxo do mentiroso. Dois artigos influenciaram os trabalhos sobre teorias formais da verdade e o paradoxo do mentiroso mais do que qualquer outro: "O Conceito de Verdade em Linguagens Formalizadas" (1935), de Alfred Tarski, e "Esboço de uma Teoria da Verdade" (1975), de Saul Kripke. Abaixo, introduziremos algumas ideias e resultados do artigo de Tarski. O artigo de Kripke será discutido na Seção 3.2.

Tarski apresenta uma série de condições que, segundo ele, qualquer *definição adequada de verdade* deve satisfazer. A mais central dessas condições é o que hoje se conhece como *Esquema T* (ou *T-esquema*, ou *Convenção T*, ou *bicondicionais de Tarski*):

(Esquema T) $\phi \leftrightarrow T\langle\phi\rangle$ para todas as sentenças ϕ.

Aqui, T é o predicado destinado a expressar a verdade, e $\langle\phi\rangle$ é o *nome* para a sentença ϕ. Aplicar o predicado T ao nome $\langle\phi\rangle$ gera a expressão $T\langle\phi\rangle$, que representa a frase "ϕ é verdadeira". O Esquema T expressa, portanto, que para toda sentença ϕ, ϕ é válida se, e somente se, a sentença "ϕ é verdadeira" for válida. O esquema T costuma ser considerado

[15] Priest, Graham, Francesco Berto, and Zach Weber, "Dialetheism", The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2024/entries/dialetheism/.

[16] Priest, Graham, Koji Tanaka, and Zach Weber, "Paraconsistent Logic", The Stanford Encyclopedia of Philosophy (Spring 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/spr2025/entries/logic-paraconsistent/.

um conjunto de sentenças dentro de uma teoria formal. É comum utilizar a aritmética de primeira ordem, ou seja, a lógica de predicados de primeira ordem estendida com um conjunto de axiomas padrão para a aritmética, como a PA (Aritmética de Peano) ou Q de Robinson. O que está sendo dito a seguir aplica-se a qualquer formalização de primeira ordem da aritmética. Nesse contexto, $\langle\phi\rangle$ denota o *código de Gödel* de ϕ, e $T\langle\phi\rangle$ é uma abreviação de $T(\langle\phi\rangle)$. O leitor não familiarizado com as codificações de Gödel (também conhecidas como numerações de Gödel) pode simplesmente pensar no mapeamento $\langle\cdot\rangle$ como um mecanismo de nomeação ou citação para fórmulas — tal como as aspas em linguagem natural. Variantes notacionais frequentemente usadas para $\langle\phi\rangle$ são $\ulcorner\phi\urcorner$ e 'ϕ'.

Tarski mostrou que o paradoxo do mentiroso é formalizável em qualquer teoria formal que contenha seu Esquema T, e, portanto, qualquer teoria desse tipo deve ser inconsistente. Esse resultado é frequentemente referido como o *teorema de Tarski sobre a indefinibilidade da verdade*. O resultado é basicamente uma formalização do paradoxo do mentiroso dentro da aritmética de primeira ordem estendida com o Esquema T. Para construir tal formalização, é necessário ser capaz de formular sentenças autorreferenciais (como a sentença do mentiroso) dentro da aritmética de primeira ordem. Essa capacidade é fornecida pelo lema diagonal.

Lema Diagonal

Seja S uma teoria que estende a aritmética de primeira ordem. Para toda fórmula $\phi(x)$, existe uma sentença Ψ tal que $S \vdash \psi \leftrightarrow \phi(\langle\psi\rangle)$.

Aqui, a notação $S \vdash \alpha$ significa que α é demonstrável na teoria S, e $\phi(\langle\psi\rangle)$ é uma abreviação de $\phi(\langle\psi\rangle)$. Suponha a fórmula $\phi(x)$ seja dada com o objetivo de expressar alguma propriedade de sentenças — por exemplo, verdade. O lema diagonal fornece a existência de uma sentença ψ que satisfaz a bi-implicação $\psi \leftrightarrow \phi\langle\psi\rangle$. A sentença $\phi\langle\psi\rangle$ pode ser vista como uma expressão de que a sentença ψ tem a propriedade expressa por $\phi(x)$. A bi-implicação, portanto, expressa que ψ é equivalente à sentença que expressa que ψ tem a propriedade ϕ. Pode-se pensar, então, que ψ é uma sentença que afirma sobre si mesma que possui a propriedade ϕ. No caso da verdade, seria uma sentença que afirma sobre si mesma que é verdadeira. A sentença ψ não é autorreferencial em um sentido estrito, mas, matematicamente, ela se comporta como uma. Portanto, é possível utilizar sentenças geradas pelo lema diagonal para formalizar paradoxos baseados em sentenças autorreferenciais, como o paradoxo do mentiroso. O lema diagonal é, às vezes, chamado de lema do ponto fixo, uma vez que a equivalência $\psi \leftrightarrow \phi\langle\psi\rangle$ pode ser vista como uma expressão de que ψ é um ponto fixo de $\phi(x)$.

Uma teoria em lógica de predicados de primeira ordem é chamada de *inconsistente* se uma contradição lógica puder ser demonstrada nela. O teorema de Tarski (sobre a indefinibilidade da verdade) pode agora ser enunciado e demonstrado.

Teorema de Tarski.

Qualquer teoria que estenda a aritmética de primeira ordem e contenha o esquema T é inconsistente.

Prova. Assuma a existência de uma teoria formal consistente S que estende a aritmética de primeira ordem e contenha o esquema T. Precisamos mostrar que essa suposição leva a uma contradição. A prova imita o paradoxo do mentiroso. Aplicando o lema diagonal, obtemos uma sentença λ que satisfaz $\lambda \leftrightarrow \neg T\langle\lambda\rangle$ em S. A sentença λ expressa sobre si mesma que não é verdadeira, assim λ corresponde à sentença do mentiroso. Instanciando o esquema T com a sentença λ, temos $\lambda \leftrightarrow T\langle\lambda\rangle$. Agora temos que tanto $\lambda \leftrightarrow \neg T\langle\lambda\rangle$ e $\lambda \leftrightarrow T\langle\lambda\rangle$ são válidas em S (prováveis em S) e, portanto, $T\langle\lambda\rangle \leftrightarrow \neg T\langle\lambda\rangle$ deve ser válida em S. Isso contradiz a consistência de S.

Observe que a contradição $T\langle\lambda\rangle \leftrightarrow \neg T\langle\lambda\rangle$ expressa: a sentença do mentiroso é verdadeira se e somente se não for. Compare isso com o paradoxo do mentiroso informal apresentado no início do artigo. O teorema de Tarski mostra que, no contexto da aritmética de primeira ordem, não é possível fornecer o que Tarski considera uma "teoria adequada da verdade". A questão central, então, torna-se: como a configuração formal ou os requisitos para uma teoria adequada da verdade podem ser modificados para recuperar a consistência — ou seja, para evitar que o paradoxo do mentiroso trivialize o sistema? Existem muitas respostas diferentes para essa pergunta, assim como há muitas maneiras diferentes de recuperar a consistência. Na Seção 3, revisaremos as abordagens mais influentes.

2.2 Consequências dos Paradoxos da Teoria dos Conjuntos

Os paradoxos da teoria dos conjuntos constituem um desafio significativo para os fundamentos da matemática. Eles mostram que é impossível ter um conceito de conjunto que satisfaça o *princípio da compreensão irrestrita* (também chamado de *compreensão plena* ou *abstração irrestrita*):

Compreensão irrestrita:

$\forall u(u \in \{x \mid \phi(x)\} \leftrightarrow \phi(u))$, para todas as fórmulas $\phi(x)$.

Em um contexto informal, as fórmulas $\phi(x)$ poderiam ser predicados arbitrários. Em um contexto mais formal, elas seriam fórmulas de, por exemplo, uma linguagem de primeira ordem adequada. O princípio da compreensão irrestrita afirma que, para qualquer propriedade (expressa por ϕ), existe o *conjunto* daquelas entidades que satisfazem a propriedade. Isso soa como um princípio muito razoável e captura, mais ou menos, o conceito intuitivo de conjunto. De fato, é o conceito de conjunto originalmente proposto pelo pai da teoria dos conjuntos, Georg Cantor (1895). Infelizmente, o princípio não é consistente, pois dá origem ao paradoxo de Russell. Considere a propriedade da não-auto-pertinência. Ela pode ser expressa pela fórmula $x \notin x$. Se deixarmos $\phi(x)$ ser a fórmula $x \notin x$, então o conjunto $\{x \mid \phi(x)\}$ torna-se o conjunto de Russell R, e obtemos a seguinte instância do princípio da compreensão irrestrita:

$$\forall u(u \in R \leftrightarrow u \notin u).$$

Análogo ao argumento no paradoxo de Russell, uma contradição é obtida ao instanciar u com R:

$$R \in R \leftrightarrow R \notin R.$$

Essa contradição expressa que o conjunto de Russell é membro de si mesmo se e somente se não é. O que se provou aqui é o seguinte.

Teorema (Inconsistência da Teoria Ingênua dos Conjuntos)

Qualquer teoria que contenha o princípio da compreensão irrestrita é inconsistente.

Compare este teorema com o teorema de Tarski. O teorema de Tarski expressa que, se formalizarmos o princípio intuitivamente mais óbrio sobre a verdade, acabamos com uma teoria inconsistente. O teorema acima expressa que a mesma coisa ocorre quando formalizamos o princípio intuitivamente mais óbrio sobre a existência e a pertinência dos conjuntos.

Dada a inconsistência da compreensão irrestrita, o objetivo passa a ser encontrar uma forma de restringir o princípio da compreensão em si ou os princípios lógicos subjacentes para obter uma teoria consistente, isto é, uma teoria dos conjuntos que não seja trivializada pelo paradoxo de Russell. Muitas teorias alternativas de conjuntos que excluem o princípio da compreensão irrestrita foram desenvolvidas ao longo do último século, entre elas a teoria dos tipos de Russell e Whitehead, a Teoria dos Tipos Simples (ST), a Teoria dos Conjuntos de Gödel-Bernays (GB), a Teoria dos Conjuntos de Zermelo-Fraenkel (ZF) e os Novos Fundamentos de Quine (NF). Todas essas são consideradas consistentes, embora não se conheçam provas simples de sua consistência. Pelo menos, todas elas escapam dos paradoxos conhecidos de autorreferência. Retornaremos a uma discussão sobre isso na Seção 3.

2.3 Consequências dos Paradoxos Epistêmicos

Os paradoxos epistêmicos representam uma ameaça à construção de teorias formais do conhecimento, uma vez que os paradoxos podem ser formalizados em muitas dessas teorias. Suponha que desejemos construir uma teoria formal da *cognoscibilidade* dentro de uma extensão da aritmética de primeira ordem. A razão para escolher formalizar a cognoscibilidade, em vez do conhecimento, é que o conhecimento é sempre relativo a um determinado agente em um certo momento, enquanto a cognoscibilidade é um conceito universal, como a verdade. Poderíamos ter escolhido trabalhar diretamente com o conhecimento, mas isso exigiria mais trabalho e tornaria a apresentação desnecessariamente complicada. Para formalizar a cognoscibilidade, introduzimos um predicado especial K, e usamos sentenças da forma $K\langle\phi\rangle$ para expressar que ϕ é cognoscível. De maneira análoga aos casos de verdade e pertinência de conjunto, deve haver certos princípios lógicos que K precisa satisfazer para que nossa teoria formal se qualifique como uma teoria adequada de cognoscibilidade. Primeiramente, todas as sentenças cognoscíveis devem ser verdadeiras. Essa propriedade pode ser formalizada pelo seguinte princípio lógico:

A1. $K\langle\phi\rangle \to \phi$, para todas as sentenças ϕ.

Claro, esse princípio deve ser cognoscível, ou seja, obtemos o seguinte princípio lógico:

A2. $K\langle K\langle\phi\rangle \to \phi\rangle$, para todas as sentenças ϕ.

Além disso, todos os teoremas da aritmética de primeira ordem devem ser cognoscíveis:

A3. $K\langle\phi\rangle$ para todas as sentenças ϕ da aritmética de primeira ordem.

Além disso, a cognoscibilidade deve ser fechada sob consequências lógicas:

A4. $K\langle\phi \to \psi\rangle \to (K\langle\phi\rangle \to K\langle\psi\rangle)$, para todas as sentenças ϕ.

Agora, os princípios A1–A4 são tudo o que é necessário para formalizar o paradoxo do conhecedor. Mais precisamente, temos o seguinte teorema, de Montague (1963), cuja prova pode ser apresentada na KS forma (veja Bolander 2004).

Teorema de Montague.

Qualquer teoria formal que estende a aritmética de primeira ordem e contém os esquemas axiomáticos A1–A4 é inconsistente.

Prova. Suponha a existência de uma teoria formal consistente S que estende a aritmética de primeira ordem e contém os esquemas axiomáticos A1–A4. Precisamos mostrar que essa suposição leva a uma contradição. A prova imita o paradoxo do conhecedor. Aplique o lema diagonal para obter uma sentença λ que satisfaz $\lambda \leftrightarrow \neg K\langle\lambda\rangle$ em S. A sentença λ expressa de si mesma que não é cognoscível, então λ corresponde aproximadamente à sentença do conhecedor, KS. A primeira parte da argumentação utilizada no paradoxo do conhecedor levou à conclusão de que é de fato verdadeira. Este raciocínio é imitado pelo seguinte argumento formal dentro de S:

1. $\lambda \to \neg K\langle\lambda\rangle$ pela escolha de λ
2. $\neg K\langle\lambda\rangle \to \lambda$ pela escolha de λ
3. $K\langle\lambda\rangle \to \lambda$ axioma A1
4. $(K\langle\lambda\rangle \to \lambda) \to ((\lambda \to \neg K\langle\lambda\rangle) \to \neg K\langle\lambda\rangle)$ tautologia proposicional
5. $(\lambda \to \neg K\langle\lambda\rangle) \to \neg K\langle\lambda\rangle$ modus ponens em 4, 3

6. $\neg K\langle\lambda\rangle$ modus ponens em 5, 1
7. λ modus ponens em 2, 6

Esta prova mostra que λ, nossa versão formal de KS, é demonstrável em S. A prova corresponde ao argumento informal de que KS é verdadeiro. Como argumentado no paradoxo do conhecedor, qualquer agente com capacidades de raciocínio suficientes será capaz de provar a verdade de KS e, assim, chegar a saber que KS é válido. Portanto, KS deve ser conhecível. O que isso significa no presente quadro formal é que também podemos provar a conhecibilidade de λ em S:

8. $K\langle\lambda \to \neg K\langle\lambda\rangle\rangle$ por A3 e escolha de λ
9. $K\langle\neg K\langle\lambda\rangle \to \lambda\rangle$ por A3 e escolha de λ
10. $K\langle K\langle\lambda\rangle \to \lambda\rangle$ axioma A2
11. $K\langle(K\langle\lambda\rangle \to \lambda) \to ((\lambda \to \neg K\langle\lambda\rangle) \to \neg K\langle\lambda\rangle)\rangle$ axioma A3 sobre tautologia proposicional
12. $K\langle(\lambda \to \neg K\langle\lambda\rangle) \to \neg K\langle\lambda\rangle\rangle$ axioma A4 sobre 11, 10
13. $K\langle\neg K\langle\lambda\rangle\rangle$ axioma A4 sobre 12, 8
14. $K\langle\lambda\rangle$ axioma A4 sobre 9, 13

Isso completa a prova da conhecibilidade de λ, correspondendo ao argumento informal de que KS é conhecido por algum agente. Note a semelhança entre as duas provas nas linhas 1–7 e 8–14. A única diferença é que, nas últimas, todas as fórmulas são precedidas por um K extra. Isso se deve ao fato de que as linhas 8–14 expressam o mesmo raciocínio das linhas 1–7, com a única diferença de que as últimas são uma prova da *conhecibilidade* de λ, em vez da *verdade* de λ. Tendo concluído que λ é tanto verdadeira quanto conhecível, agora obtemos imediatamente uma contradição, assim como na parábola do conhecedor:

15. $\neg K\langle\lambda\rangle$ modus ponens em 1, 7
16. $K\langle\lambda\rangle \land \neg K\langle\lambda\rangle$ conjunção de 14 e 15

Isso completa a prova do teorema de Montague.

A prova acima imita diretamente o raciocínio subjacente ao paradoxo do conhecedor. O teorema de Montague mostra que, no contexto da aritmética de primeira ordem, não podemos ter uma teoria do conhecimento ou da conhecibilidade que satisfaça mesmo os princípios básicos A1–A4. O teorema de Montague é uma generalização do teorema de Tarski. Se um símbolo de predicado K satisfaz o esquema T de Tarski, então é fácil ver que também satisfará as esquemas axiomáticos A1–A4. Assim, os esquemas axiomáticos A1–A4 constituem um enfraquecimento do esquema T, e o teorema de Montague mostra que mesmo essa versão muito mais fraca do esquema T é suficiente para produzir inconsistência. Uma possível resposta poderia ser que mesmo os esquemas axiomáticos A1–A4 são muito fortes e deveriam ser enfraquecidos ainda mais. No entanto, como nas formalizações anteriores de paradoxos, não está claro como enfraquecer ainda mais as suposições, uma vez que todas as suposições parecem sensatas e naturais para o conceito que estamos formalizando (conhecibilidade, neste caso). Outra possível saída para o resultado de inconsistência poderia ser preservar os princípios em sua forma atual, mas aplicá-los apenas a um subconjunto das sentenças disponíveis (significando que A1–A4 são exigidas apenas para sentenças $\phi \in S$ para algum subconjunto S). Os princípios deveriam ser válidos para todas as sentenças "normais", mas poderíamos não querer insistir que eles sejam válidos para certas sentenças patológicas que expressam declarações autorreferenciais. No entanto, não está claro que podemos indicar sensatamente exatamente quais sentenças são normais e quais são patológicas. Mas ainda poderíamos ser capazes de encontrar *alguns* subconjuntos sensatos para os quais podemos instanciar nossos princípios e ainda ter uma teoria consistente. Revières e Levesque (1988) mostraram que os princípios permanecem consistentes quando instanciados apenas sobre as chamadas *sentenças regulares*, e esse resultado foi mais tarde generalizado para a classe mais inclusiva das chamadas sentenças RPQ (Morreau e Kraus, 1998). A maioria das soluções para os resultados de inconsistência induzidos por paradoxos é, na verdade, desse tipo: estratificar ou limitar a aplicabilidade dos princípios que levam à inconsistência, sejam os princípios de verdade, existência de conjuntos, conhecibilidade ou algo quarto. Vamos explorar isso muito mais na Seção 3.

Formalizar o conhecimento como um predicado em uma lógica de primeira ordem é referido como o *tratamento sintático* do conhecimento. Alternativamente, pode-se optar por formalizar o conhecimento como um operador modal em uma lógica modal adequada. Isso é conhecido como o *tratamento semântico* do conhecimento (consulte a entrada sobre lógica epistêmica[17]). No tratamento semântico do conhecimento, geralmente se evitam problemas de autorreferência e, portanto, inconsistência, mas isso ocorre à custa do poder expressivo do formalismo (os problemas de autorreferência são evitados pela lógica modal proposicional, que não admite nada equivalente ao lema diagonal para a construção de fórmulas autorreferenciais). Na verdade, as sentenças regulares mencionadas acima são alcançadas exatamente por uma tradução da lógica modal de primeira ordem para a lógica de predicados, portanto, a única maneira de instanciar os princípios apenas sobre as sentenças regulares é uma forma de seguir o tratamento sintático do conhecimento, mas ainda garantir consistência ao limitar o operador de conhecimento à expressividade que possui no tratamento semântico.

2.4 Consequências sobre provabilidade e Computabilidade

O argumento central apresentado na prova do teorema de Tarski está intimamente relacionado ao argumento central no primeiro teorema da incompletude de Gödel (Gödel, 1931). O teorema de Gödel pode receber a seguinte formulação.

> **Primeiro teorema da incompletude de Gödel.**
>
> Se a aritmética de primeira ordem é ω-consistente, então ela é incompleta.

Uma teoria é chamada de ω-*consistente* se a seguinte condição for satisfeita para toda fórmula $\phi(x)$ contendo x como sua única variável livre: Se $\vdash \neg\phi(n)$ (significando que $\neg\phi(n)$ é demonstrável) para cada número natural n, então $\nvdash \exists x\phi(x)$ (significando que $\exists x\phi(x)$ não é demonstrável). A ω-consistência é uma condição mais forte do que a consistência ordinária, de modo que qualquer teoria ω-consistente também será consistente. Uma teoria é *incompleta* se contiver uma fórmula que não pode ser nem provada nem reprovada.

> *Esboço de prova do primeiro teorema da incompletude de Gödel*. Suponha que a aritmética de primeira ordem seja ao mesmo tempo ω-consistente e completa. Precisamos mostrar que isso leva a uma contradição. Gödel constrói

[17] Rendsvig, Rasmus, John Symons, and Yanjing Wang, "Epistemic Logic", The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2024/entries/logic-epistemic/.

uma fórmula Bew (para “Beweis”) na aritmética formal que satisfaz, para todos ϕ e todos n,

(1) $\vdash$ Bew$(n, \langle\phi\rangle)$ se e somente se n é o código de Gödel de uma prova de ϕ.

Assumindo que a teoria seja ω-consistente e completa, podemos provar que, para todas as sentenças ϕ,

(2) $\vdash \exists x$Bew$(x, \langle\phi\rangle)$ se e somente se $\vdash \phi$.

A prova de (2) segue assim. Primeiro, provamos a implicação da esquerda para a direita. Se $\vdash \exists x$Bew$(x, \langle\phi\rangle)$, então há algum n tal que $\nvdash \neg$Bew$(n, \langle\phi\rangle)$, pela ω-consistência. Pela completude, obtemos $\vdash$ Bew$(n, \langle\phi\rangle)$ para esse n. Por (1) acima, obtemos que n denota uma prova de ϕ. Isto é, ϕ é demonstrável, portanto temos $\vdash \phi$. Para provar a implicação da direita para a esquerda, note que se $\vdash \phi$, então deve haver um n tal que $\vdash$ Bew$(n, \langle\phi\rangle)$, por (1). A partir disso, obtemos $\vdash \exists x$Bew$(x, \langle\phi\rangle)$, conforme requerido. Isso conclui a prova de (2). Agora, quando em uma teoria completa temos (2), devemos também ter:

(3) $\vdash \exists x$Bew$(x, \langle\varphi\rangle) \leftrightarrow \phi$, para todas as sentenças ϕ.

Se deixarmos $\exists x$Bew$(x, \langle\phi\rangle)$ abreviada como $T\langle\phi\rangle$, então (3) torna-se:

$\vdash T\langle\phi\rangle \leftrightarrow \varphi$ para todas as sentenças ϕ.

Este é o esquema T! Assim, se assumirmos que a aritmética de primeira ordem é ω-consistente e completa, o esquema T pode ser interpretado nela. Agora, o teorema de Tarski acima mostra que não existe tal teoria consistente, e, portanto, temos uma contradição.

Na prova acima, reduzimos o teorema da incompletude de Gödel a uma aplicação do teorema de Tarski para mostrar a estreita ligação entre os dois (essa versão da prova é devida a Bolander, 2002). Gödel estava bem ciente dessa ligação, e, de fato, parece que Gödel provou o teorema de Tarski antes mesmo de Tarski (Feferman, 1984). O teorema de Gödel pode ser interpretado como uma demonstração de uma limitação no que pode ser alcançado por procedimentos puramente formais. Ele diz que, se a aritmética de primeira ordem é ω-consistente (como se acredita ser), então deve haver sentenças aritméticas que não podem ser provadas nem refutadas pelos procedimentos formais da aritmética de primeira ordem. Pode-se, a princípio, esperar que essa limitação possa ser resolvida pela inclusão de axiomas adicionais, mas Gödel mostrou que o resultado da incompletude ainda vale quando a aritmética de primeira ordem é estendida com um conjunto finito arbitrário de esquemas de axiomas (ou, mais geralmente, um conjunto recursivo arbitrário de axiomas). Assim, obtemos um resultado de limitação geral, dizendo que não pode existir um procedimento formal de prova pelo qual qualquer sentença aritmética possa ser provada como verdadeira ou falsa. Para mais detalhes sobre o teorema da incompletude de Gödel, consulte o verbete sobre Kurt Gödel[18].

O resultado de limitação do teorema de Gödel está intimamente relacionado a outro resultado de limitação conhecido como a *indecidibilidade do problema da parada*. Esse resultado afirma que há limitações no que pode ser computado. Vamos apresentar esse resultado a seguir. O resultado é baseado na noção de uma *máquina de Turing*, que é um modelo genérico de um programa de computador rodando em um computador com memória ilimitada. Assim, qualquer programa rodando em qualquer computador pode ser pensado como uma máquina de Turing (veja o verbete sobre Máquinas de Turing[19] para mais detalhes). Quando rodamos uma máquina de Turing, ela ou termina após um número finito de passos de computação, ou continua rodando para sempre. No caso de ela terminar após um número finito de passos, dizemos que ela *para*. O *problema da parada* é o problema de encontrar uma máquina de Turing que possa decidir se outras máquinas de Turing param ou não. Dizemos que uma máquina de Turing H *decide o problema da parada* se o seguinte for verdadeiro:

> H recebe como entrada um par $(\langle A\rangle, x)$, consistindo no código de Gödel $\langle A\rangle$ de uma máquina de Turing A e uma string arbitrária x. H retorna a resposta "sim" se a máquina de Turing A parar quando dada a entrada x, e "não" caso contrário.

Assim, se uma máquina de Turing H decide o problema da parada, podemos usá-la para determinar, para uma máquina de Turing A arbitrária e uma entrada x arbitrária, se A

[18]Kennedy, Juliette, "Kurt Gödel", The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/win2020/entries/goedel/.

[19]De Mol, Liesbeth, "Turing Machines", The Stanford Encyclopedia of Philosophy (Winter 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/win2024/entries/turing-machine/.

irá parar com a entrada x ou não. A *indecidibilidade do problema da parada* é o seguinte resultado, devido a Turing (1937), afirmando que nenhuma máquina desse tipo pode existir:

Teorema (Indecidibilidade do Problema da Parada)

Não existe nenhuma máquina de Turing que decida o problema da parada.

Prova. Suponha a existência de uma máquina de Turing H que decide o problema da parada. Precisamos mostrar que essa suposição leva a uma contradição. A prova imita o paradoxo de Grelling. Chamamos uma máquina de Turing A de *heterológica* se A não parar com a entrada $\langle A \rangle$, ou seja, se A não parar quando dada seu próprio código de Gödel como entrada. Usando H, podemos construir uma máquina de Turing G que para se, e somente se, ela for dada como entrada o código de Gödel de uma máquina de Turing heterológica. G funciona da seguinte maneira:

G recebe como entrada o código de Gödel de uma máquina de Turing A. Então, ela executa H com a entrada $(\langle A \rangle, \langle A \rangle)$. Se H com a entrada $(\langle A \rangle, \langle A \rangle)$ retorna "não", G é interrompida. Por outro lado, se H com a entrada $(\langle A \rangle, \langle A \rangle)$ retorna "sim" então G é forçada a entrar em um loop infinito (ou seja, forçada a nunca parar).

Analogamente ao paradoxo de Grelling, agora podemos perguntar se G é uma máquina de Turing heterológica ou não. Isso leva à seguinte sequência de equivalências:

$$
\begin{aligned}
G \text{ é heterológica} &\Leftrightarrow G \text{ não para com a entrada } \langle G \rangle \text{ (pela definição de heterológica)} \\
&\Leftrightarrow H \text{ retorna "não" com a entrada } (\langle G \rangle, \langle G \rangle) \text{ (já que } H \text{ decide o proble} \\
&\Leftrightarrow G \text{ para com a entrada } \langle G \rangle \text{ (pela construção de } G\text{)} \\
&\Leftrightarrow G \text{ não é heterológica (pela definição de heterológica)}
\end{aligned}
$$

Isso dá a contradição necessária.

A partir dos dois teoremas acima, vemos que nas áreas de provabilidade e computabilidade, os paradoxos da auto-referência se transformam em resultados de limitação: existem limites para o que pode ser provado e para o que pode ser computado. Isso é bastante semelhante ao que aconteceu nas áreas de semântica, teoria dos conjuntos e epistemologia: os paradoxos da auto-referência se transformaram em teoremas que mostram que há limites para as propriedades que podemos consistentemente assumir para um predicado de verdade (teorema de Tarski), para uma teoria dos conjuntos (inconsistência da teoria dos conjuntos ingênua), e para um predicado de conhecimento (teorema de Montague). É difícil aceitar esses resultados de limitação porque a maioria deles entra em conflito com nossas intuições e expectativas. O papel central desempenhado pela auto-referência em todos eles potencialmente os torna ainda mais difíceis de aceitar, ou, pelo menos, definitivamente os torna mais intrigantes. No entanto, somos forçados a aceitá-los e forçados a aceitar o fato de que, nessas áreas, não podemos ter tudo o que poderíamos (de outra forma) razoavelmente pedir.

3. Resolvendo os Paradoxos

Nesta seção, vamos explorar como resolver — ou melhor, contornar — os paradoxos. Para resolver ou contornar um paradoxo, é necessário enfraquecer algumas das suposições que levam à contradição. É muito difícil escolher quais suposições enfraquecer, já que cada uma das suposições explicitamente declaradas que sustentam um paradoxo parece ser "obviamente verdadeira" — caso contrário, não seria qualificada como um paradoxo. Abaixo, veremos as abordagens mais influentes para resolver os paradoxos.

Até agora, a apresentação foi estruturada de acordo com o tipo de paradoxo: os paradoxos semânticos, conjuntistas e epistêmicos foram tratados separadamente. No entanto, também foi demonstrado que esses três tipos de paradoxos são semelhantes em sua estrutura subjacente, e argumentou-se que uma solução para um deve ser uma solução para todos (o princípio da solução uniforme). Portanto, a seguir, a apresentação será estruturada não de acordo com o tipo de paradoxo, mas de acordo com o tipo de solução. Cada tipo de solução considerado a seguir pode ser aplicado a qualquer um dos paradoxos de autorreferência, embora, na maioria dos casos, as construções envolvidas tenham sido originalmente desenvolvidas com apenas um tipo de paradoxo em mente.

3.1 Construindo Hierarquias Explícitas

Construir hierarquias é um método para contornar os paradoxos semânticos, conjuntistas e epistêmicos. A solução original de Russell para seu paradoxo — assim como a solução original de Tarski para seu problema da *indefinibilidade da verdade* — foi construir hierarquias. No caso de Russell, isso levou à *teoria dos tipos*. No caso de Tarski, resultou no que é conhecido como a *hierarquia de linguagens de Tarski*. Em ambos os casos, a ideia é estratificar o universo do discurso (conjuntos, sentenças) em níveis. Na teoria dos tipos, esses níveis são chamados de tipos. A ideia fundamental da teoria dos tipos é introduzir a restrição de que qualquer conjunto de um determinado tipo pode conter apenas elementos de tipos inferiores (isto é, pode conter apenas conjuntos que estão localizados mais abaixo na estratificação). Isso efetivamente bloqueia o paradoxo de Russell, pois nenhum conjunto pode ser membro de si mesmo.

No caso de Tarski, a estratificação é obtida da seguinte maneira. Suponha que se queira equipar uma linguagem L_0 com um predicado de verdade. A partir do teorema de Tarski (Seção 2.1), sabe-se que esse predicado de verdade não pode fazer parte da própria linguagem L_0 — pelo menos, não enquanto desejarmos que o predicado de verdade satisfaça o esquema T. Em vez disso, constrói-se uma hierarquia de linguagens $L_0, L_1, L_2, ...$, onde cada linguagem L_{i+1} possui um predicado de verdade T_{i+1} que se aplica apenas às sentenças de L_j, para $j \leq i$. Nessa hierarquia, L_0 é chamada de *linguagem-objeto* e, para todo i, L_{i+1} é chamada de *meta-linguagem* de L_i. Essa hierarquia bloqueia efetivamente o paradoxo do mentiroso, já que agora uma sentença só pode expressar a verdade ou falsidade de sentenças em níveis inferiores, e, portanto, uma sentença como a do mentiroso, que expressa sua própria falsidade, não pode ser formada.

A teoria dos tipos de Russell pode ser considerada uma solução para o paradoxo de Russell, uma vez que demonstra como “reparar” a teoria dos conjuntos de forma que o paradoxo desapareça. Similarmente, a hierarquia de Tarski pode ser vista como uma solução para o paradoxo do mentiroso. É a mesma ideia de estratificação que está subjacente às soluções de Russell e Tarski. O ponto a ser destacado é que tanto o paradoxo de Russell quanto o paradoxo do mentiroso dependem crucialmente de noções circulares (*auto*-pertencimento e *auto*rreferência). Ao criar uma estratificação na qual um objeto pode conter ou se referir apenas a objetos em níveis inferiores, a circularidade desaparece. No caso dos paradoxos epistêmicos, uma estratificação semelhante poderia ser obtida ao fazer uma distinção explícita entre conhecimento de primeira ordem (conhecimento sobre o mundo externo), conhecimento de segunda ordem (conhecimento sobre o conhecimento de primeira ordem), conhecimento de terceira ordem (conhecimento sobre o conhecimento de segunda ordem), e assim por diante. Essa estratificação, na verdade, surge naturalmente no tratamento semântico do conhecimento, onde o conhecimento é formalizado como um operador modal.

Construir hierarquias explícitas é suficiente para evitar a circularidade e, portanto, suficiente para bloquear os paradoxos padrão de autorreferência. No entanto, também existem paradoxos, como o de Yablo, que não dependem de circularidade e autorreferência. Esses paradoxos também podem ser bloqueados por uma abordagem de hierarquia, mas é necessário exigir que a hierarquia seja bem fundada, ou seja, que tenha um nível mais baixo. Caso contrário, ainda será possível formular paradoxos decorrentes de fundamentação inadequada (não bem-fundada). Por exemplo, o paradoxo de Yablo pode ser formalizado em uma hierarquia descendente de linguagens. Uma hierarquia descendente de linguagens consiste em linguagens $L_0, L_{-1}, L_{-2}, ...$, onde cada linguagem L_{-i} tem um predicado de verdade que se aplica apenas às sentenças das linguagens L_{-j}, $j > i$. Similarmente, um paradoxo da teoria dos conjuntos não bem-fundados pode ser formulado em uma teoria dos tipos que permite tipos negativos. A conclusão é que uma estratificação do universo não é suficiente por si só para evitar todos os paradoxos — a estratificação também precisa ser bem fundada.

Construir uma hierarquia explícita (e bem fundada) para resolver os paradoxos é hoje considerado, por muitos, uma abordagem excessivamente drástica e rígida. As hierarquias introduzem uma série de tecnicalidades complicastes que não estão presentes em um "universo plano", e, embora os paradoxos de fato desapareçam, o mesmo ocorre com todas as ocorrências não paradoxais de autorreferência. Kripke (1975) oferece o seguinte exemplo ilustrativo tirado do discurso cotidiano. Suponha que N seja a seguinte declaração, feita por Nixon:

(N) Todas as declarações de Jones sobre Watergate são verdadeiras,

e seja J a seguinte declaração, feita por Jones,

(J) A maioria das declarações de Nixon sobre Watergate é falsa.

Em uma hierarquia de linguagem tarskiana, a sentença N teria que estar em um nível superior ao de todas as declarações de Jones, e, inversamente, a sentença J teria que estar em um nível superior ao de todas as declarações de Nixon. Já que N é uma declaração de Nixon, e J é uma declaração de Jones, N teria que estar em um nível superior a J, e J em um nível superior a N. Isso é obviamente impossível, entretanto, em uma hierarquia como a tarskiana, essas sentenças não podem nem mesmo ser formuladas. As sentenças N e J são, de fato, ambas indiretamente autorreferenciais, uma vez que N faz referência a uma totalidade que inclui J, e J faz referência a uma totalidade que inclui N. No entanto, na maioria dos casos, N e J são inofensivas e não produzem um paradoxo. Um paradoxo é produzido apenas no caso especial em que todas as declarações de Jones, exceto possivelmente J, são verdadeiras, e exatamente metade das declarações de Nixon são falsas, desconsiderando N. Kripke usa o fato de que N e J são apenas problemáticas em um certo caso especial como um argumento contra uma abordagem que exclui totalmente a possibilidade de formular N e J.

Outro argumento contra a abordagem da hierarquia é que a estratificação explícita não faz parte do discurso comum, e, portanto, pode ser considerado um tanto *ad hoc* introduzi-la em contextos formais com o único propósito de contornar os paradoxos. Não ter uma estratificação explícita no discurso cotidiano não implica, obviamente, a não existência de uma estratificação implícita subjacente, mas pelo menos não está explicitamente representada em nossa sintaxe.

Os argumentos apresentados acima estão entre as razões pelas quais o trabalho de Russell e Tarski não foi considerado como a solução definitiva para os paradoxos. Muitas soluções alternativas foram propostas. Pode-se, por exemplo, tentar buscar hierarquias *implícitas* em vez de hierarquias *explícitas*. Uma hierarquia implícita é uma hierarquia não refletida explicitamente na sintaxe da linguagem. Na seção seguinte, consideraremos algumas das soluções para os paradoxos obtidas por meio de tais estratificações implícitas.

3.2 Building Implicit Hierarchies

A abordagem da hierarquia de Tarski para os paradoxos semânticos dominou o campo até 1975, quando Kripke publicou seu famoso e influente artigo, “Outline of a Theory of Truth”. Este artigo moldou significativamente a maioria das abordagens posteriores às teorias da verdade e aos paradoxos semânticos. Deve-se notar, no entanto, que ideias bastante semelhantes às de Kripke foram desenvolvidas simultaneamente e de forma independente por Martin e Woodruff (1975), e que uma abordagem paralela em um contexto de teoria dos conjuntos foi desenvolvida independentemente por Gilmore (1974).

3.2.1 A Teoria da Verdade de Kripke

As ideias de Kripke são baseadas em uma análise dos problemas envolvidos na abordagem da hierarquia de Tarski. Kripke lista uma série de argumentos contra ter uma hierarquia de linguagem na qual cada sentença reside em um nível fixo, determinado por sua forma sintática. Ele propõe uma solução alternativa que ainda utiliza a ideia de ter níveis, mas onde os níveis não se tornam uma parte explícita da sintaxe. Em vez disso, os níveis se tornam estágios em uma construção iterativa de um predicado de verdade. Para explicar a construção de Kripke, algum maquinário técnico adicional é necessário.

Em cada estágio na construção de Kripke, o predicado de verdade é apenas *parcialmente definido*, ou seja, ele se aplica apenas a algumas das sentenças da linguagem. Para lidar com tais predicados parcialmente definidos, é empregada *uma lógica de três valores*, ou seja, uma lógica que opera com um terceiro valor, indefinido, além dos valores de verdade verdadeiro e falso. Muitas vezes, o terceiro valor é denotado por “u” ou “$\perp$” (bottom). Um predicado parcialmente definido recebe apenas um dos valores de verdade clássicos, *verdadeiro* ou *falso*, quando é aplicado a um dos termos para os quais o predicado foi definido, e, caso contrário, recebe o valor *indefinido*. Existem várias lógicas de três valores disponíveis, diferindo em como tratam o terceiro valor. Aqui, apenas uma delas é descrita brevemente, chamada de *lógica forte de três valores de Kleene*. Informações mais detalhadas sobre essa lógica e lógicas relacionadas podem ser encontradas na entrada sobre lógica multivalorada[20].

Na lógica forte de três valores de Kleene, o valor $\perp$ (*indefinido*) pode ser interpretado como “ainda não definido”. Assim, pode-se pensar em fórmulas com o valor $\perp$ como fórmulas que na verdade têm um valor de verdade clássico (*verdadeiro* ou *falso*), mas que simplesmente ainda não foi determinado. Essa interpretação de *indefinido* é refletida nas tabelas de verdade para a lógica, apresentadas abaixo. A tabela de verdade superior é para a disjunção, a inferior para a negação:

$\vee$	verdadeiro	falso	$\perp$
verdadeiro	verdadeiro	verdadeiro	verdadeiro
falso	verdadeiro	falso	$\perp$
$\perp$	verdadeiro	$\perp$	$\perp$

$\neg$	
verdadeiro	falso
falso	verdadeiro
$\perp$	$\perp$

[20]Gottwald, Siegfried, “Many-Valued Logic”, The Stanford Encyclopedia of Philosophy (Summer 2022 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/sum2022/entries/logic-manyvalued/.

Essas tabelas de verdade definem a lógica de três valores completamente, uma vez que $\vee$ e $\neg$ são consideradas para formar um conjunto adequado de conectivos e a quantificação existencial e universal é tratada como disjunção e conjunção infinitas, respectivamente.

Para lidar com predicados de verdade parcialmente definidos, é necessário introduzir a noção de modelos parciais. Em um *modelo parcial* para uma linguagem de primeira ordem, cada símbolo de predicado n-place P é interpretado por um par (U, V) de relações n-place disjuntas no domínio do modelo. U é chamado de *extensão* de P e V de *anti-extensão*. No modelo, P é verdadeiro para os objetos em U, falso para os objetos em V e indefinido caso contrário. Dessa forma, qualquer sentença atômica recebe um dos valores de verdade verdadeiro, falso ou indefinido no modelo. As fórmulas não atômicas recebem valores de verdade no modelo usando a lógica forte de três valores de Kleene para avaliar os conectivos.

Dada a definição de um modelo parcial, uma *linguagem parcialmente interpretada* é um par (L, M) onde L é uma linguagem de primeira ordem e M é um modelo parcial de L. Kripke define recursivamente uma sequência de linguagens parcialmente interpretadas $L_0, L_1, L_2, ...$, que diferem apenas em sua interpretação do predicado de verdade T. A primeira linguagem, L_0, é considerada uma linguagem arbitrária na qual tanto a extensão quanto a anti-extensão de T são o conjunto vazio. Assim, em L_0, o predicado de verdade está completamente indefinido. Para qualquer α, a linguagem $L_{\alpha+1}$ é como L_α, exceto que T é interpretado pelo par extensão/anti-extensão (U, V) dado por:

- U é o conjunto de códigos de Gödel $\langle\phi\rangle$ de sentenças ϕ verdadeiras em L_α.
- V é o conjunto de códigos de Gödel $\langle\phi\rangle$ de sentenças ϕ falsas em L_α.

Essa definição imediatamente indica que para todo α,

$$(4) \quad \phi \text{ é verdadeira (falsa) em } L_\alpha \iff T\langle\phi\rangle \text{ é verdadeira (falsa) em } L_{\alpha+1}.$$

O que foi construído é uma sequência $L_0, L_1, L_2, ...$ de linguagens parcialmente interpretadas tal que T é interpretado em $L_{\alpha+1}$ como o predicado de verdade para L_α. Isso é semelhante à hierarquia de linguagens de Tarski, exceto que aqui não há uma distinção sintática entre as diferentes linguagens e seus predicados de verdade.

A sequência $L_0, L_1, L_2, ...$ possui uma propriedade importante: Para cada α, a interpretação de T em $L_{\alpha+1}$ estende a interpretação de T em L_α, ou seja, tanto a extensão quanto a anti-extensão de T aumentam (ou permanecem as mesmas) ao se mover de L_α para $L_{\alpha+1}$. Isso significa que se pode definir uma nova linguagem parcialmente interpretada L_ω deixando a extensão de T ser a união de todas as extensões de T em $L_0, L_1, L_2, ...$; e de maneira similar para a anti-extensão. Assim, em L_ω, a interpretação de T estende a interpretação que T recebe em todas as linguagens anteriores. Isso fornece uma estratégia para continuar a construção iterativa de um predicado de verdade até o transfinito: Para cada ordinal sucessor $\alpha + 1$, defina $L_{\alpha+1}$ a partir de L_α exatamente como no caso finito acima; e para cada ordinal limite σ, defina L_σ a partir das linguagens precedentes $(L_i)_{i<\sigma}$ da mesma forma que L_ω foi definido (para uma explicação detalhada dos números ordinais e seu uso neste contexto, consulte a entrada sobre a teoria da revisão da verdade[21]). Uma consideração simples de cardinalidade agora mostra que essa sequência transfinita de linguagens acabará por *estabilizar*: Existe um ordinal γ tal que $L_\gamma = L_{\gamma+1}$. Assim, a seguinte instância de (4) é obtida:

$$(5) \quad \phi \text{ é verdadeira (falsa) em } L_\gamma \iff T\langle\phi\rangle \text{ é verdadeira (falsa) em } L_\gamma.$$

Isso mostra que L_γ é, na verdade, uma linguagem que contém seu próprio predicado de verdade: qualquer sentença ϕ é verdadeira (falsa) se e somente se a sentença que expressa sua verdade, $T\langle\phi\rangle$, é verdadeira (falsa). A equivalência (5) não é nada mais do que um análogo semântico do esquema T de Tarski em um contexto de três valores. A construção da linguagem L_γ foi uma das principais contribuições de Kripke (1975). Ela mostra que, em um ambiente lógico de três valores, é realmente possível que uma linguagem contenha seu próprio predicado de verdade. É fácil ver que o terceiro valor, *indefinido*, é essencial para que as coisas funcionem: se L_γ fosse uma linguagem totalmente interpretada (ou seja, uma linguagem sem sentenças indefinidas), então L_γ satisfaria o esquema T, conforme (5) acima. No entanto, isso contradiz imediatamente o teorema de Tarski de que tal linguagem totalmente interpretada pode existir.

Entre as sentenças que recebem o valor *indefinido* em L_γ está a sentença do mentiroso. A solução para o paradoxo do mentiroso implícita na teoria de Kripke é a seguinte: uma vez que tanto assumir que a sentença do mentiroso é verdadeira quanto assumir que é falsa

[21]Kremer, Philip and Edoardo Rivello, "The Revision Theory of Truth", The Stanford Encyclopedia of Philosophy (Summer 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2023/entries/truth-revision/.

levam a uma contradição, deve ser nenhuma das duas; ela é *indefinida*. A sentença do mentiroso é dita sofrer de uma lacuna de *valor-verdade*. A ideia de evitar o paradoxo do mentiroso permitindo lacunas de valor-verdade de fato apareceu várias vezes na literatura antes do artigo de Kripke, mas Kripke foi um dos primeiros a torná-la uma parte integral de uma teoria genuína.

Assim como a solução da hierarquia para o paradoxo do mentiroso, a solução da lacuna de valor-verdade é considerada problemática por muitos. A principal crítica é que, ao usar uma semântica de três valores, obtém-se uma linguagem interpretada que é expressivamente fraca. Não se pode, por exemplo, em nenhuma das linguagens de Kripke ter um predicado que expresse a propriedade de ser *indefinido*. Isso é, de fato, notado pelo próprio Kripke. Se uma linguagem parcialmente interpretada contivesse tal predicado, a seguinte sentença do mentiroso reforçada dentro da linguagem poderia ser formulada: "Esta sentença é falsa ou indefinida". A sentença do mentiroso reforçada é verdadeira se e somente se é falsa ou indefinida, então temos um novo paradoxo, chamado *paradoxo do mentiroso reforçado*. O problema com o paradoxo do mentiroso reforçado é conhecido como um *problema de vingança*: dada qualquer solução para o mentiroso, parece que podemos gerar um novo paradoxo reforçado, análogo ao mentiroso, que permanece sem solução. A ideia é que, qualquer que seja o status semântico que a solução suposta afirme que a sentença do mentiroso tenha, se nos for permitido referir-nos livremente a esse status semântico na linguagem objeto, podemos gerar um novo paradoxo.

A incapacidade da linguagem kripkeana de expressar seu próprio predicado *indefinido* também significa que não podemos, na linguagem objeto kripkeana, expressar uma afirmação como: "A sentença do mentiroso é indefinida". Na verdade, na linguagem de Kripke L_γ, a sentença do mentiroso *é* indefinida, portanto, a sentença anterior expressa uma verdade sobre L_γ que não pode ser expressa dentro de L_γ (daí a linguagem é expressivamente incompleta). Para expressar a verdadeira afirmação "A sentença do mentiroso é indefinida", somos forçados a ascender a uma meta-linguagem de L_γ. Como Kripke (1975) coloca: "O fantasma da hierarquia de Tarski ainda está conosco."

3.2.2 Extensões e Alternativas à Teoria da Verdade de Kripke

Sucessores do trabalho de Kripke, muitas tentativas foram feitas para construir linguagens que contêm seu próprio predicado de verdade e que não sofrem com o problema de vingança dos mentirosos reforçados. Muitas dessas tentativas se concentraram em modificar ou estender a lógica forte de três valores subjacente, por exemplo, modificando a semântica do condicional (Field, 2003, 2008) ou permitindo um número ilimitado de valores de verdade (Cook, 2007; Schlenker, 2010; Tourville e Cook, 2016).

A teoria de Kripke contorna o paradoxo do mentiroso atribuindo a ele o valor *indefinido*. Uma maneira alternativa de contornar o paradoxo do mentiroso seria atribuir a ele o valor *tanto verdadeiro quanto falso* em uma lógica paraconsistente adequada. Esta seria a solução correta de acordo com a visão dialeteísta, cf. Seção 2. Uma das lógicas paraconsistentes mais simples é a LP, que é uma lógica de três valores com as mesmas tabelas de verdade da lógica forte de três valores de Kleene apresentadas acima— a única diferença é que o terceiro valor de verdade é interpretado como *tanto verdadeiro quanto falso*, em vez de *indefinido*. Uma razão para preferir uma lógica paraconsistente a uma lógica parcial é que sentenças paradoxais, como a do mentiroso, podem então ser modeladas como *contradições verdadeiras* (dialetéia), em vez de lacunas de valor-verdade. Referimo-nos novamente às entradas sobre dialeteísmo[22] e lógica paraconsistente[23].

A escolha é entre lacunas de *valor-verdade* e *glut de valor-verdade*: uma lacuna de valor-verdade é uma declaração sem valor-verdade, nem verdadeira nem falsa (como *indefinido* na lógica forte de três valores de Kleene), e um glut de valor-verdade é uma declaração com vários valores de verdade, por exemplo, tanto verdadeira quanto falsa (como na lógica paraconsistente LP). Também existem argumentos a favor de permitir tanto lacunas quanto excessos, por exemplo, permitindo que o conjunto de valores de verdade forme um bilattice (Fitting, 2006; Odintsov e Wansing, 2015). O bilattice não trivial mais simples tem exatamente quatro valores, que no contexto de valores de verdade são interpretados como: *verdadeiro*, *falso*, $\bot$ (nem verdadeiro nem falso) e $\top$ (tanto verdadeiro quanto falso).

Para uma discussão mais extensa sobre a teoria de Kripke, seus sucessores e rivais, consulte a entrada sobre o paradoxo do mentiroso[24].

---

[22] Priest, Graham, Francesco Berto, and Zach Weber, “Dialetheism”, The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2024/entries/dialetheism/.

[23] Priest, Graham, Koji Tanaka, and Zach Weber, “Paraconsistent Logic”, The Stanford Encyclopedia of Philosophy (Spring 2025 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/spr2025/entries/logic-paraconsistent/.

[24] Beall, Jc, Michael Glanzberg, and David Ripley, “Liar Paradox”, The Stanford Encyclopedia of Philosophy (Winter 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/win2023/entries/liar-paradox/.

3.2.3 Hierarquias Implícitas em Teorias de Conjuntos

Construir hierarquias implícitas, em vez de explícitas, é também uma ideia que tem sido empregada na teoria dos conjuntos. New Foundations (NF) de Quine (1937) é uma modificação da teoria de tipos simples, onde a estratificação em tipos sintáticos foi substituída por uma estratificação no princípio de compreensão:

Compreensão NF:
$\forall u\,(u \in \{x \mid \phi(x)\} \leftrightarrow \phi(u))$ para todas as fórmulas estratificadas $\phi(x)$.

Uma fórmula ϕ é estratificada se existe um mapeamento σ (uma *estratificação*) das variáveis de ϕ para os números naturais, de modo que se $u \in v$ é uma subfórmula de ϕ, então $\sigma(v) = \sigma(u) + 1$ e se $u = v$ é uma subfórmula de ϕ, então $\sigma(v) = \sigma(u)$. Obviamente, a fórmula $x \notin x$ não é estratificada, e assim o princípio de compreensão NF não pode ser usado para formular o paradoxo de Russell na teoria. As *New Foundations* de Quine são essencialmente obtidas a partir da teoria de tipos, ocultando os tipos da sintaxe. Assim, a teoria ainda faz uso de uma abordagem hierárquica para evitar os paradoxos, mas a hierarquia é tornada implícita ao não ser representada na sintaxe das fórmulas. Cantini (2015) investigou a possibilidade de imitar essa abordagem de hierarquia implícita no contexto das teorias da verdade (conseguindo uma hierarquia de verdade tarskiana implicitamente representada).

A teoria dos conjuntos de Zermelo-Fraenkel (ZF) é outra teoria que se baseia na ideia de uma hierarquia implícita para contornar os paradoxos. No entanto, ela o faz de uma maneira muito menos direta do que NF. Em ZF, os conjuntos são construídos de baixo para cima, começando com o conjunto vazio e iterando uma construção de conjuntos cada vez maiores usando as operações de união e conjunto das partes. Isso produz uma hierarquia com o conjunto vazio no nível mais baixo, nível 0, e com a operação de conjunto das partes produzindo um conjunto de nível $\alpha + 1$ a partir de um conjunto de nível α. Análogo à construção iterativa de Kripke, o procedimento é continuado para o transfinito usando o operador de união nos níveis ordinais limites. A hierarquia obtida é chamada de *hierarquia cumulativa*. Um dos axiomas de ZF, o axioma da fundação, afirma que todo conjunto de ZF reside em um determinado nível nessa hierarquia cumulativa. Em outras palavras, o axioma da fundação afirma que não existem conjuntos em ZF além daqueles que podem ser construídos de baixo para cima pelo procedimento iterativo descrito. Como em uma hierarquia cumulativa, não pode haver conjuntos que contenham a si mesmos, nenhum conjunto universal e nenhum conjunto não bem-fundado, nenhum dos paradoxos conhecidos pode ser imediatamente formulado na teoria. Isso, evidentemente, não garante por si só a consistência de ZF, mas pelo menos ilustra como a ideia de uma hierarquia de conjuntos desempenha um papel significativo em ZF também. ZF possui um status privilegiado entre as teorias de conjuntos, uma vez que é hoje o candidato mais amplamente reconhecido para uma fundamentação formal da matemática.

3.3 Abordagens Gerais de Ponto Fixo

A construção iterativa de Kripke de um predicado de verdade apresentado acima pode ser vista como uma instância de uma *abordagem de ponto fixo* mais geral para a construção de teorias formais da verdade. As abordagens de ponto fixo tornaram-se centrais para as teorias formais contemporâneas da verdade. A ideia principal é ter um *operador de revisão da verdade* e, em seguida, buscar os pontos fixos desse operador. O cerne dessas abordagens de ponto fixo é algum *teorema de ponto fixo adequado*, que garanta a existência de pontos fixos para certos tipos de operadores. Existem vários teoremas diferentes de ponto fixo disponíveis. Considere agora um dos mais simples.

Teorema do Ponto Fixo.

Seja τ um operador monótono em uma ordem parcial completa por cadeias (doravante ccpo). Então τ tem um menor ponto fixo, ou seja, existe um menor f tal que $\tau(f) = f$.

Uma *ccpo* é uma ordem parcial $(D, <)$ na qual todo subconjunto totalmente ordenado de D tem um limite superior mínimo. Os subconjuntos totalmente ordenados de D são chamados de cadeias em D. Um *operador monótono* em $(D, <)$ é um aplicação $\tau : D \to D$ que satisfaz:

$$d_1 \leq d_2 \Rightarrow \tau(d_1) \leq \tau(d_2), \text{ para todos } d_1, d_2 \in D.$$

A construção de Kripke se encaixa no teorema de ponto fixo acima da seguinte forma. Primeiro, note que o conjunto de linguagens parcialmente interpretadas que diferem apenas na interpretação de T forma uma ccpo: simplesmente definimos a ordenação nessas linguagens por $L_1 \leq L_2$ se e somente se a interpretação de T em L_2 estender a interpretação de T em L_1 (isto é, a extensão e a anti-extensão de T em L_1 estão incluídas nas de

L_2). Então, definimos um operador de revisão da verdade τ nessas linguagens da seguinte maneira:

(6) $\tau(L) = L'$, onde $T\langle\phi\rangle$ é verdadeiro (falso) em L' se e somente se ϕ é verdadeiro (falso)

Note que se L_α é uma das linguagens na construção de Kripke, então $L_{\alpha+1} = \tau(L_\alpha)$. A ideia desse operador de revisão da verdade τ é que, se $\tau(L) = L'$, então L' será uma linguagem na qual T é interpretado como o predicado de verdade para L. Portanto, se $\tau(L) = L$ para algum L, isto é, se L é um ponto fixo de τ, então L será uma linguagem contendo seu próprio predicado de verdade. Isso motiva a busca por pontos fixos de τ. Como é fácil perceber que τ é monótono, pelo teorema do ponto fixo, ele tem um menor ponto fixo. Não é difícil ver que esse ponto fixo é exatamente a linguagem L_γ construída na teoria da verdade de Kripke. Assim, a construção de Kripke é recapturada no contexto de pontos fixos para operadores monótonos.

O objetivo de introduzir a maquinaria adicional não foi apenas redescobrir a linguagem L_γ. O objetivo foi, ao contrário, fornecer uma estrutura muito mais geral e abstrata que pode levar a novas teorias da verdade e oferecer novos insights sobre os paradoxos semânticos. Acontece que o operador de revisão da verdade τ, definido acima, tem muitos pontos fixos interessantes além de L_γ. Também é possível obter novas teorias da verdade considerando maneiras alternativas de transformar o conjunto de linguagens interpretadas em uma ccpo. Por exemplo, pode-se adicionar um valor de verdade adicional e considerar a situação em uma lógica de quatro valores, como considerado por Fitting (1997); ou pode-se remover o terceiro valor de verdade *indefinido* e construir uma ccpo em um contexto completamente clássico. No contexto clássico, a atenção é restringida às linguagens totalmente interpretadas (linguagens nas quais toda sentença é ou verdadeira ou falsa), e uma ordenação sobre elas é definida da seguinte por: $L_1 \leq L_2$ se e somente se a extensão do predicado de verdade em L_1 está incluída na extensão do predicado de verdade em L_2, isto é, se e somente se L_2 aponta pelo menos tantas sentenças como verdadeiras quanto L_1. Isso dá uma ccpo. Usando o teorema do ponto fixo neste contexto com um operador de revisão adequadamente definido, é relativamente fácil provar a existência de uma linguagem totalmente interpretada contendo uma *definição positiva de verdade*. Com isso, queremos dizer que a linguagem interpretada tem um predicado T que satisfaz a seguinte versão restrita do esquema T:

$$(7) \quad \phi \leftrightarrow T(\langle\phi\rangle) \text{ para todas as sentenças positivas } \phi,$$

onde as sentenças positivas são aquelas construídas sem o uso de negação $(\neg)$. Dado que (7) é satisfiável em uma linguagem totalmente interpretada, a teoria de primeira ordem contendo as sentenças de (7) como axiomas deve ser consistente. Isso deve ser contrastado com o teorema de Tarski, que afirma que o esquema T *irrestrito* é inconsistente. Se o princípio de compreensão irrestrito for similarmente restrito às fórmulas positivas, também obtemos uma teoria consistente. Isso foi originalmente demonstrado por Gilmore (1974).

A abordagem do ponto fixo também é o ponto de partida da *teoria de revisão da verdade* desenvolvida por Belnap e Gupta (1993). A teoria de revisão da verdade é a teoria mais influente da verdade e os paradoxos semânticos que foi desenvolvida desde a teoria de Kripke. A teoria de revisão considera o operador de revisão da verdade padrão τ, definido por (6), como um operador sobre as linguagens *totalmente* interpretadas. Nessas linguagens, τ não tem um ponto fixo: Se houvesse tal ponto fixo L, então L seria uma linguagem totalmente interpretada satisfazendo o esquema completo T, o que contradiz diretamente o teorema de Tarski. Como τ não tem um ponto fixo nas linguagens totalmente interpretadas, a teoria de revisão considera, em vez disso, sequências transfinitas $L_1, L_2, \dots, L_\omega, L_{\omega+1}, \dots$ de linguagens totalmente interpretadas que satisfazem:

- Para qualquer ordinal sucessor $\alpha + 1$, $L_{\alpha+1} = \tau(L_\alpha)$.
- Para qualquer ordinal limite σ e qualquer sentença ϕ, se ϕ estabiliza no valor verdadeiro (ou falso) na sequência $(L_\alpha)_{\alpha<\sigma}$, então ϕ é verdadeira (ou falsa) em L_σ.

Em tal sequência, cada sentença ϕ ou eventualmente se estabiliza em um valor de verdade clássico (verdadeiro ou falso), ou nunca se estabiliza. Um exemplo de sentença que nunca se estabiliza é a sentença do mentiroso: se a sentença do mentiroso é verdadeira em uma das linguagens L_α, ela será falsa em $L_{\alpha+1}$, e vice-versa. A teoria de revisão, assim, fornece uma explicação da verdade que modela corretamente o comportamento da sentença do mentiroso como uma que nunca se estabiliza em um valor de verdade. Na teoria de revisão, argumenta-se que isso oferece uma explicação mais precisa da verdade e da autorreferência do que a teoria de Kripke, na qual a sentença do mentiroso é simplesmente atribuída ao valor indefinido. Tanto a teoria de revisão da verdade quanto as teorias de ponto fixo ao estilo de Kripke continuam sendo ativamente pesquisadas (Gupta e Standefer, 2017;

Hsiung, 2017; Schindler, 2017). Um relato completo da teoria de revisão pode ser encontrado na entrada sobre a teoria de revisão da verdade[25].

Estudar fenômenos autorreferenciais como pontos fixos não se limita às teorias da verdade. Por exemplo, no contexto de paradoxos epistêmicos, o paradoxo de Brandenburger-Keisler foi interpretado como um resultado de ponto fixo por Abramsky e Zvesper (2015).

4. Desenvolvimentos Recentes

Murzi e Massimiliano (2015) fornecem uma visão geral dos desenvolvimentos recentes nas abordagens para resolver os paradoxos: paracompletude (permitindo lacunas de valores de verdade), paraconsistência (permitindo excessos de valores de verdade), lógicas subestruturais (enfraquecendo os princípios lógicos da lógica clássica) e os problemas de vingança que essas abordagens irão ou podem causar. Desenvolvimentos recentes em lógicas subestruturais como solução para os paradoxos incluem French (2016) (eliminando a reflexividade), Caret, Colin e Weber (2015), Shapiro e Lionel (2015), Mares e Paoli (2014) (eliminando a contração), e Cobreros, Égré, Ripley e van Rooij (2014) (eliminando a transitividade). Mais recentemente, um volume especial da Synthese foi dedicado às abordagens subestruturais para o paradoxo (edição de dezembro de 2021, com uma introdução de Zardini, 2021). O volume de Achourioti et al. (eds., 2015) contém vários artigos sobre autorreferência e como evitar paradoxos no contexto das teorias da verdade.

Volker Halbach e Albert Visser (2014a, 2014b) realizaram um estudo muito detalhado sobre autorreferência na aritmética, estudando o que significa para uma sentença da aritmética atribuir a si mesma uma propriedade e como isso depende da codificação escolhida, dos detalhes da construção do ponto fixo, etc. Albert Visser (2019) também está entre os autores que estudaram o quão pouca autorreferência podemos usar ao provar teoremas clássicos como o segundo teorema da incompletude de Gödel.

Ainda há trabalhos que tentam caracterizar exatamente o que significa algo ser um paradoxo (Hsiung, 2022). Isso está de certa forma no espírito dos Esquemas de Inclusão de Priest e das classificações gráfico-teóricas de paradoxicidade mencionadas acima. Encontrar os ingredientes exatos necessários e suficientes para um paradoxo ainda é um problema

[25]Kremer, Philip and Edoardo Rivello, "The Revision Theory of Truth", The Stanford Encyclopedia of Philosophy (Summer 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = https://plato.stanford.edu/archives/sum2023/entries/truth-revision/.

em aberto. Também houve mais trabalho em definir exatamente o que significa algo ser autorreferencial (Picollo, 2018, 2020). Grabmayr, Halbach e Ye (2023) distinguem entre autorreferência genuína e acidental, e assim tentam fornecer uma análise mais detalhada da autorreferencialidade. Uma análise mais detalhada da autorreferencialidade também pode ser feita observando mais de perto o tipo de numeração de Gödel que usamos (Grabmayr e Visser, 2023; Kripke, 2023).

Referências Bibliográficas

Samson Abramsky and Jonathan Zvesper. From lawvere to brandenburger-keisler: Interactive forms of diagonalization and self-reference. *Journal of Computer and System Sciences*, 81(5):799–812, 2015.

Theodora Achourioti, Kentaro Fujimoto, Hannes Galinon, and José Martínez Fernández, editors. *Unifying the Philosophy of Truth*, volume 36 of *Logic, Epistemology, and the Unity of Science*. Springer, Dordrecht, 2015.

Can Başkent. A yabloesque paradox in epistemic game theory. *Synthese*, pages 1–24, 2016.

Emil Badici. The liar paradox and the inclosure schema 1. *Australasian Journal of Philosophy*, 86(4):583–596, 2008.

Steven J. Bartlett, editor. *Reflexivity — A Source-Book in Self-Reference*. North-Holland Publishing Co., Amsterdam, 1992.

Steven J. Bartlett and Peter Suber, editors. *Self-Reference — Reflections on Reflexivity*. Martinus Nijhoff Publishers, Dordrecht, 1987.

Jon Barwise and John Etchemendy. *The Liar — An Essay on Truth and Circularity*. Oxford University Press, New York, 1987.

Jon Barwise and Lawrence Moss. *Vicious Circles — On the Mathematics of Non-Wellfounded Phenomena*. CSLI Publications, Stanford, 1996.

J. C. Beall, editor. *Liars and Heaps: New Essays on Paradox*. Clarendon Press, Oxford, 2003.

J. C. Beall, editor. *Revenge of the Liar: New Essays on the Paradox*. Oxford University Press, Oxford, 2007.

J. C. Beall. *Spandrels of Truth*. Oxford University Press, Oxford, 2009.

J. C. Beall. Finding tolerance without gluts. *Mind*, 123(491):791–811, 2014a.

J. C. Beall. End of inclosure. *Mind*, 123(491):829–849, 2014b.

Tobias Beringer and Thomas Schindler. A graph-theoretic analysis of the semantic paradoxes. *The Bulletin of Symbolic Logic*, 24(4):442–892, 2017.

Claudio Bernardi. Fixed points and unfounded chains. *Annals of Pure and Applied Logic*, 109(3):163–178, 2001.

Thomas Bolander. Self-reference and logic. *Phi News*, 1:9–44, 2002.

Thomas Bolander. *Logical Theories for Agent Introspection*. PhD thesis, Technical University of Denmark, Denmark, 2004.

Thomas Bolander, Vincent F. Hendricks, and Stig Andur Pedersen, editors. *Self-Reference*. CSLI Publications, Stanford, 2006.

George Boolos. *The Logic of Provability*. Cambridge University Press, Cambridge, 1993.

Adam Brandenburger and H. Jerome Keisler. An impossibility theorem on beliefs in games. *Studia Logica*, 84(2):211–240, 2006.

Riccardo Bruni and Lorenzo Rossi. Truth meets vagueness. unifying the semantic and the soritical paradoxes. *Journal of Philosophical Logic*, 52:1637–1671, 2023.

John Buridan. *Summulae de Dialectica — An annotated translation*. Yale University Press, New Haven and London, 2001. With a philosophical introduction by Gyula Klima.

Jesse M. Butler. An entirely non-self-referential yabloesque paradox. *Synthese*, pages 1–13, 2017.

Andrea Cantini. *Logical Frameworks for Truth and Abstraction — An Axiomatic Study*. North-Holland Publishing Co., Amsterdam, 1996.

Andrea Cantini. Paradoxes, self-reference and truth in the 20th century. In Dov M. Gabbay and John Woods, editors, *Handbook of the History of Logic*, volume 5, pages 875–1013. Elsevier, Amsterdam, 2009.

Andrea Cantini. On stratified truth. In Theodora Achourioti, Kentaro Fujimoto, Hannes Galinon, and José Martínez Fernández, editors, *Unifying the Philosophy of Truth*. Springer, Dordrecht, 2015.

Georg Cantor. Über eine elementare frage der mannigfaltigkeitslehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 1:75–78, 1891.

Georg Cantor. Beiträge zur begründung der transfiniten mengenlehre. *Mathematische Annalen*, 46(4):481–512, 1895.

Colin R. Caret and Zach Weber. A note on contraction-free logic for validity. *Topoi*, 34(1): 63–74, 2015.

Cezary Cieśliński and Radosław Urbaniak. Gödelizing the yablo sequence. *Journal of Philo-*

sophical Logic, 42(5):679–695, 2013.

Pablo Cobreros, Paul Égré, David Ripley, and Robert van Rooij. Reaching transparent truth. *Mind*, 122(488):841–866, 2013.

Pablo Cobreros, Paul Égré, David Ripley, and Robert van Rooij. Vagueness, truth and permissive consequence. In Theodora Achourioti, Kentaro Fujimoto, Hannes Galinon, and José Martínez Fernández, editors, *Unifying the Philosophy of Truth*, pages 409–430. Springer, Dordrecht, 2015.

Mark Colyvan. Vagueness and truth. In Heather Dyke, editor, *From Truth to Reality: New Essays in Logic and Metaphysics*. Routledge, Oxford, 2009.

Roy Cook. Patterns of paradox. *The Journal of Symbolic Logic*, 69(3):767–774, 2004.

Roy Cook. Embracing revenge: on the indefinite extensibility of language. In J. C. Beall, editor, *Revenge of the Liar: New Essays on the Paradox*. Oxford University Press, Oxford, 2007.

Roy Cook. What is a truth value and how many are there? *Studia Logica*, 92(2):183–201, 2009.

Roy Cook. *Paradoxes*. Polity Press, Cambridge, 2013.

Roy Cook. *The Yablo Paradox: An Essay on Circularity*. Oxford University Press, Oxford, 2014.

Jean de Revières and Hector J. Levesque. The consistency of syntactical treatments of knowledge. *Computational Intelligence*, 4:31–41, 1988.

Sjur Dyrkolbotn and Michał Walicki. Propositional discourse logic. *Synthese*, 191(5):863–899, 2014.

G. W. Erickson and J. A. Fossa. *Dictionary of Paradox*. University Press of America, Lanham, 1998.

Solomon Feferman. Kurt gödel: conviction and caution. *Philosophia Naturalis*, 21:546–562, 1984a.

Solomon Feferman. Toward useful type-free theories i. *The Journal of Symbolic Logic*, 49: 75–111, 1984b.

Solomon Feferman. Reflecting on incompleteness. *The Journal of Symbolic Logic*, 56:1–49, 1991.

Hartry Field. A revenge-immune solution to the semantic paradoxes. *Journal of Philosophical Logic*, 32(2):139–177, 2003.

Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008.

Melvin Fitting. Notes on the mathematical aspects of kripke's theory of truth. *Notre Dame Journal of Formal Logic*, 27(1):75–88, 1986.

Melvin Fitting. A theory of truth that prefers falsehood. *The Journal of Philosophical Logic*, 26(5):477–500, 1997.

Melvin Fitting. Bilattices are nice things. In Thomas Bolander, Vincent F. Hendricks, and Stig Andur Pedersen, editors, *Self-Reference*. CSLI Publications, Stanford, 2006.

Thomas E. Forster. *Set Theory with a Universal Set*. The Clarendon Press, Oxford, 1995.

Rohan French. Structural reflexivity and the paradoxes of self-reference. *Ergo*, 3, 2016.

Haim Gaifman. Operational pointer semantics: Solution to self-referential puzzles i. In *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 43–59. Morgan Kaufmann Publishers Inc., 1988.

Haim Gaifman. Pointers to truth. *The Journal of Philosophy*, 89(5):223–261, 1992.

Haim Gaifman. Pointers to propositions. In Alexis Chapuis and Anil Gupta, editors, *Circularity, definition and truth*, pages 79–121. Indian Council of Philosophical Research, 2000.

Paul C. Gilmore. The consistency of partial set theory without extensionality. In *Axiomatic Set Theory (Proc. Sympos. Pure Math., Vol. XIII, Part II)*, pages 147–153. American Mathematical Society, Providence, RI, 1974.

Kurt Gödel. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931.

Balthasar Grabmayr and Albert Visser. Self-reference upfront: A study of self-referential gödel numberings. *The Review of Symbolic Logic*, 16(2):385–424, 2023.

Balthasar Grabmayr, Volker Halbach, and Lingyuan Ye. Varieties of self-reference in metamathematics. *Journal of Philosophical Logic*, 52:1005–1052, 2023.

Anil Gupta and Nuel Belnap. *The Revision Theory of Truth*. MIT Press, Cambridge, MA, 1993.

Anil Gupta and Shawn Standefer. Conditionals in theories of truth. *Journal of Philosophical Logic*, 46(1):27–63, 2017.

Volker Halbach. On a side effect of solving fitch's paradox by typing knowledge. *Analysis*, 68 (298):114–120, 2008.

Volker Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, Cambridge, 2011.

Volker Halbach and Albert Visser. Self-reference in arithmetic i. *The Review of Symbolic Logic*, 7(4):671–691, 2014a.

Volker Halbach and Albert Visser. Self-reference in arithmetic ii. *The Review of Symbolic*

Logic, 7(4):692–712, 2014b.
Volker Halbach and Shuoying Zhang. Yablo without gödel. *Analysis*, 77(1):53–59, 2017.
Leon Horsten. *The Tarskian Turn: Deflationism and Axiomatic Truth*. MIT Press, Cambridge, MA, 2011.
Ming Hsiung. Boolean paradoxes and revision periods. *Studia Logica*, 105(5):881–914, 2017.
Ming Hsiung. From paradoxicality to paradox. *Erkenntnis*, 2022. First online 06 December 2022. doi:10.1007/s10670-022-00640-9.
G. E. Hughes. *John Buridan on self-reference — Chapter Eight of Buridan's "Sophismata", with a Translation, an Introduction, and a Philosophical Commentary*. Cambridge University Press, Cambridge, 1982.
David Kaplan and Richard Montague. A paradox regained. *Notre Dame Journal of Formal Logic*, 1(3):79–90, 1960.
Saul Kripke. Outline of a theory of truth. *The Journal of Philosophy*, 72:690–716, 1975.
Saul Kripke. Gödel's theorem and direct self-reference. *The Review of Symbolic Logic*, 16 (2):650–654, 2023.
Graham Leach-Krouse. Yablifying the rosser sentence. *Journal of Philosophical Logic*, 43 (5):827–834, 2014.
Hannes Leitgeb. On the probabilistic convention t. *The Review of Symbolic Logic*, 1(2): 218–224, 2008.
Edwin Mares and Francesco Paoli. Logical consequence and the paradoxes. *Journal of Philosophical Logic*, 43(2-3):439–469, 2014.
Robert L. Martin. *Recent Essays on Truth and the Liar Paradox*. Oxford University Press, Oxford, 1984.
Robert L. Martin and P. W. Woodruff. On representing 'true-in-l' in l. *Philosophia*, 5:213–217, 1975.
Vann McGee. *Truth, Vagueness, and Paradox — An Essay on the Logic of Truth*. Hackett Publishing Co., Indianapolis, 1990.
Vann McGee. Maximal consistent sets of instances of tarski's schema (t). *The Journal of Philosophical Logic*, 21(3):235–241, 1992.
Richard Montague. Syntactical treatment of modality, with corollaries on reflection principles and finite axiomatizability. *Acta Philosophica Fennica*, 16:153–167, 1963.
Michael Morreau and Sarit Kraus. Syntactical treatments of propositional attitudes. *Artificial*

Intelligence, 106(1):161–177, 1998.

Julien Murzi and Massimiliano Carrara. Paradox and logical revision. a short introduction. *Topoi*, 34(1):7–14, 2015.

Sergei P. Odintsov and Heinrich Wansing. The logic of generalized truth values and the logic of bilattices. *Studia Logica*, 103(1):91–112, 2015.

Eric Pacuit. Understanding the brandenburger-keisler paradox. *Studia Logica*, 86(3):435–454, 2007.

Charlie Pelling. A self-referential paradox for the truth account of assertion. *Analysis*, 71(4): 688–688, 2011.

Donald Perlis and V. S. Subrahmanian. Meta-languages, reflection principles and self-reference. In Dov M. Gabbay, Christopher J. Hogger, and John A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 2, pages 323–358. Oxford University Press, 1994.

Lavinia Picollo. Yablo's paradox in second-order languages: Consistency and unsatisfiability. *Studia Logica*, 101(3):601–617, 2013.

Lavinia Picollo. Reference in arithmetic. *The Review of Symbolic Logic*, 11(3):573–603, 2018.

Lavinia Picollo. Alethic reference. *Journal of Philosophical Logic*, 49(3):417–438, 2020.

Graham Priest. *In Contradiction — A Study of the Transconsistent*. Martinus Nijhoff Publishers, Dordrecht, 1987.

Graham Priest. The structure of the paradoxes of self-reference. *Mind*, 103:25–34, 1994.

Graham Priest. *Beyond the Limits of Thought*. Cambridge University Press, Cambridge, 1995.

Graham Priest. Yablo's paradox. *Analysis*, 57(4):236–242, 1997.

Graham Priest. Badici on inclosures and the liar paradox. *Australasian Journal of Philosophy*, 88(2):359–366, 2010a.

Graham Priest. Inclosures, vagueness, and self-reference. *Notre Dame Journal of Formal Logic*, 51(1):69–84, 2010b.

Graham Priest. Vague inclosures. In *Paraconsistency: Logic and Applications*. Springer, Dordrecht, 2013.

Willard Van Orman Quine. New foundations for mathematical logic. *American Mathematical Monthly*, 44:70–80, 1937.

Willard Van Orman Quine. *The Ways of Paradox and Other Essays*. Harvard University Press, Cambridge, Massachusetts, 1976.

Landon Rabern, Brian Rabern, and Matthew Macauley. Dangerous reference graphs and semantic paradoxes. *Journal of Philosophical Logic*, 42(5):727–765, 2013.

Guillermo E. Rosado Haddock. Recent truth theories: A case study. *Axiomathes*, 12(1): 87–115, 2001.

Bertrand Russell. On some difficulties in the theory of transfinite numbers and order types. *Proceedings of the London Mathematical Society*, 4:29–53, 1905.

Thomas Schindler. Some notes on truths and comprehension. *Journal of Philosophical Logic*, pages 1–31, 2017.

Philippe Schlenker. Super liars. *The Review of Symbolic Logic*, 1(1):1–41, 2010.

Lionel Shapiro. Naive structure, contraction and paradox. *Topoi*, 34(1):75–87, 2015.

Michael Sheard. A guide to truth predicates in the modern era. *The Journal of Symbolic Logic*, 59(3):1032–1054, 1994.

Keith Simmons. *Universality and the Liar — An Essay on Truth and the Diagonal Argument*. Cambridge University Press, Cambridge, 1993.

Hartley Slater. The fallacy in russell's schema. *Russell: the Journal of Bertrand Russell Studies*, 22(2):4, 2002.

Hartley Slater. What priest (amongst many others) has been missing. *Ratio*, 23(2):184–198, 2010.

Raymond M. Smullyan. *Gödel's Incompleteness Theorems*. Oxford University Press, Oxford, 1992.

Raymond M. Smullyan. *Diagonalization and self-reference*. Oxford University Press, Oxford, 1994.

Raymond M. Smullyan. Self-reference in all its glory! In Thomas Bolander, Vincent F. Hendricks, and Stig Andur Pedersen, editors, *Self-Reference*. CSLI Publications, Stanford, 2006.

Jeff Snapper. The liar paradox in new clothes. *Analysis*, 72(2):319–322, 2012.

Roy A. Sorensen. *A Brief History of the Paradox: Philosophy and the Labyrinths of the Mind*. Oxford University Press, New York, 2003.

Alfred Tarski. Der wahrheitsbegriff in den formalisierten sprachen. *Studia Philosophica*, 1: 261–405, 1935.

Alfred Tarski. *Undecidable Theories*. North-Holland Publishing Co., Amsterdam, 1968.

Richmond Thomason. A note of syntactical treatments of modality. *Synthese*, 44:391–395, 1980.

Nicholas Tourville and Roy T. Cook. Embracing the technicalities: Expressive completeness and revenge. *The Review of Symbolic Logic*, 9(2):325–358, 2016.

Dustin Tucker and Richmond H. Thomason. Paradoxes of intensionality. *The Review of Symbolic Logic*, 4(3):394–411, 2011.

Alan M. Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(2):230–265, 1936. Correction: ibid., 43: 544–546 (1937).

Rafał Urbaniak. Leitgeb, 'about,' yablo. *Logique & Analyse*, 207:239–254, 2009.

Jordi Valor Abad. The inclosure scheme and the solution to the paradoxes of self-reference. *Synthese*, 160(2):183–202, 2008.

Jean van Heijenoort. *From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931*. Harvard University Press, Cambridge, MA, 1967.

Albert Visser. Semantics and the liar paradox. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 4, pages 617–706. Kluwer, Dordrecht, 1989.

Albert Visser. From tarski to gödel – or how to derive the second incompleteness theorem from the undefinability of truth without self-reference. *Journal of Logic and Computation*, 29(5):595–604, 2019.

Michał Walicki. Reference, paradoxes and truth. *Synthese*, 171(1):195–226, 2009.

Michał Walicki. Kernels of digraphs with finitely many ends. *Discrete Mathematics*, 342: 473–486, 2019.

Zach Weber. Explanation and solution in the inclosure argument. *Australasian Journal of Philosophy*, 88(2):353–357, 2010a.

Zach Weber. A paraconsistent model of vagueness. *Mind*, 119:1026–1045, 2010b.

Zach Weber et al. Tolerating gluts. *Mind*, 123(491):813–828, 2014.

Stefan Wintein. Assertoric semantics and the computational power of self-referential truth. *Journal of Philosophical Logic*, 41(2):317–345, 2012.

Stephen Yablo. Truth and reflection. *Journal of Philosophical Logic*, 14(3):297–349, 1985.

Stephen Yablo. Paradox without self-reference. *Analysis*, 53:251–252, 1993.

Stephen Yablo. Circularity and paradox. In Thomas Bolander, Vincent F. Hendricks, and Stig Andur Pedersen, editors, *Self-Reference*, pages 165–184. CSLI Publications, Stanford, 2006.

Elia Zardini. Truth without contra(di)ction. *Review of Symbolic Logic*, 4(4):498–535, 2011.

Elia Zardini. Substructural approaches to paradox: an introduction to the special issue. *Synthese*, 199(3):493–525, 2021.

Haixia Zhong. Definability and the structure of logical paradoxes. *Australasian Journal of Philosophy*, 90(4):779–788, 2012.

William S. Zwicker. Playing games with games: The hypergame paradox. *The American Mathematical Monthly*, 94(6):507–514, 1987.

(III) Paradoxo do Mentiroso[1]

Título Original: Liar Paradox
Autores: BEALL, Jc. GLANZBERG, Michael. RIPLEY, Ellie.
Tradução: Annelyze Reis e Kherian Gracher
Revisão: Caio Cezar Silva

A primeira sentença deste ensaio é uma mentira. Há algo paradoxal em tal afirmação, como já se reconhece desde a Antiguidade. Para entender o porquê, lembre-se que toda proposição falsa é, por definição, não verdadeira. A sentença inicial é verdadeira? Se sim, então o que ela afirma é uma mentira, e, portanto, não pode ser verdadeira. Inversamente, suponha que não seja verdadeira. Nós (a saber, os autores) a enunciamos, e, em geral, os enunciados são proferidos com a intenção de serem tomados como verdadeiros. Declarar algo dessa maneira, quando o que é dito é falso, é uma mentira. Mas, dado o que afirma a sentença, ela é verdadeira afinal!

O fato de que há algum tipo de enigma associado a sentenças como a primeira deste ensaio é algo que tem sido frequentemente acentuado ao longo da história da filosofia. Tal questão foi discutida na Antiguidade clássica, notadamente pelos Megáricos, mas também foi mencionada por Aristóteles e por Cícero. Como um dos *insolubilia*, foi objeto de extensa investigação por parte dos lógicos medievais, como Buridan. Mais recentemente, o estudo

[1]BEALL, Jc; GLANZBERG, Michael; RIPLEY, Ellie, “Liar Paradox”, In: ZALTA, E. N.; NODELMAN, U. (eds.). The Stanford Encyclopedia of Philosophy (Fall 2025 Edition). Stanford, CA: The Metaphysics Research Lab, 2025. Disponível em: https://plato.stanford.edu/archives/fall2025/entries/liar-paradox/.
A seguir está a tradução da entrada sobre o Paradoxo do Mentiroso de Jc Beall, Michael Glanzberg e Ellie Ripley na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/fall2025/entries/liar-paradox/. Esta versão traduzida pode diferir da versão atual da entrada, que pode ter sido atualizada desde o momento desta tradução. A versão atual está localizada em https://plato.stanford.edu/entries/liar-paradox/. Agradecemos aos editores Edward N. Zalta e Uri Nodelman pela permissão para traduzir e publicar esta entrada.

desse problema tem desempenhado um papel central no desenvolvimento da lógica matemática moderna, tornando-se, inclusive, um campo de pesquisa autônomo e amplamente explorado. O paradoxo é por vezes denominado ”paradoxo de Epimênides”, dado que a tradição atribui uma sentença análoga à primeira deste ensaio a Epimênides de Creta, a quem se imputa a afirmação de que todos os cretenses são sempre mentirosos. O fato de que tal afirmação tenha sido proferida por um cretense encontra-se, nada menos, do que em uma das epístolas do Novo Testamento![2]

Mentir é um fenômeno complexo, mas o que torna sentenças como a primeira deste ensaio enigmáticas não está, em essência, vinculado a intenções, normas sociais ou quaisquer fatores similares. Antes, parece tratar-se de algo relacionado à verdade, ou, ao menos, a alguma noção semântica relacionada à verdade. O enigma é comumente denominado “o paradoxo do Mentiroso”, embora tal designação se refira, na realidade, a uma família de paradoxos associada a sentenças desse tipo. A designação como família de paradoxos é adequada, pois elas parecem conduzir a conclusões incoerentes, como, por exemplo: “tudo é verdadeiro”. De fato, o paradoxo do mentiroso parece permitir que se chegue a tais conclusões com base na lógica, a partir de alguns princípios bastante evidentes, que por vezes foram considerados como princípios da própria lógica. Assim, deparamo-nos com a situação surpreendente de que algo próximo — ou análogo — à própria lógica possa conduzir à incoerência. Este é, talvez, o mais impetuoso dos tipos de paradoxo, e enfrentá-lo tem sido uma tarefa central da lógica desde os seus primórdios.

Neste ensaio, examinaremos os membros mais relevantes da família dos paradoxos do mentiroso, bem como algumas das principais propostas sobre como tais paradoxos poderiam ser resolvidos. Os últimos milênios produziram uma grande variedade de soluções propostas, mas não será possível analisá-las todas; ao invés disso, nos concentraremos em algumas que, nas discussões recentes, mostraram-se particularmente relevantes.

[2] Assim, um paradoxo quase ocorre no Novo Testamento. Para uma discussão instigante sobre esse tema, veja Anderson (1970). Para uma análise mais aprofundada sobre a história do paradoxo do Mentiroso, consulte Sorensen (2003). Há trabalhos recentes de grande importância sobre as teorias medievais do paradoxo do mentiroso e sua relevância para as abordagens contemporâneas. Veja os artigos reunidos em Rahman, Tulenheimo e Genot (2008) e, por exemplo, Read (2002, 2006); Restall (2008); e Simmons (1993).

1. O Paradoxo e o Fenômeno Mais Amplo

1.1 Mentiroso de Falsidade Simples

Considere uma sentença denominada ‘FMentiroso’, que afirma de si mesma (isto é, afirma a respeito de *FMentiroso*) que é falsa:

FMentiroso: *FMentiroso é falsa.*

Tal formulação parece conduzir a uma contradição da seguinte maneira. Se a sentença “FMentiroso é falsa” é verdadeira, então, dado o que ela afirma, FMentiroso é falsa. Contudo, FMentiroso é precisamente a sentença “FMentiroso é falsa”, de modo que se pode concluir que, se FMentiroso é verdadeira, então FMentiroso é falsa. Por outro lado, se FMentiroso é falsa, então a sentença “FMentiroso é falsa” é verdadeira. Novamente, dado que FMentiroso é idêntica à sentença “FMentiroso é falsa”, conclui-se que, se FMentiroso é falsa, então FMentiroso é verdadeira. Assim, demonstramos que FMentiroso é falsa se, e somente se, FMentiroso é verdadeira. No entanto, se toda sentença é ou verdadeira ou falsa, então a sentença FMentiroso é ou verdadeira ou falsa, mas nesse caso — segundo o raciocínio acima — ela é ambos: verdadeira e falsa. Isso constitui uma contradição. E contradições, de acordo com muitas teorias lógicas (por exemplo, a lógica clássica, a lógica intuicionista, entre outras), implicam trivialidade, isto é, que toda sentença é verdadeira.

Uma resposta óbvia consiste em negar que toda sentença seja verdadeira ou falsa, ou seja, negar o princípio da bivalência. Como discutiremos na seção 4, alguns desdobramentos dessa ideia permanecem relevantes na pesquisa contemporânea sobre o paradoxo do mentiroso. Ainda assim, uma variante simples da sentença do mentiroso mostra que tal resposta imediata não esgota o problema.

1.2 Mentiroso Não-Verdadeiro Simples

Ao invés de operar com a noção de falsidade, podemos construir uma sentença do Mentiroso utilizando o predicado composto “não-verdadeiro”.[3] Considere uma sentença denomi-

[3]Note-se que a terminologia não é uniforme. Van Fraassen (1968) introduziu o termo ’Mentiroso Fortalecido’ para denominar o que estamos chamando de ’Mentiroso Não-Verdadeiro Simples’. Porém, o termo cunhado por Van Fraassen é frequentemente utilizado para uma forma de paradoxo da revanche [ou vingança] ao modo de C. Parsons (1974), como discutido na seção 4.1.3.

nada ‘NMentiroso’ (N para ‘não-verdadeiro’), que afirma de si mesma que não é verdadeira:

NMentiroso: *NMentiroso não é verdadeira.*

O argumento que conduz à contradição é similar ao caso de FMentiroso. Em resumo: se NMentiroso é verdadeira, então, conforme o que afirma, ela não é verdadeira; e se NMentiroso não é verdadeira, então ela é verdadeira. No entanto, se toda sentença é ou verdadeira ou não verdadeira, então NMentiroso é verdadeira ou não verdadeira e, nesse caso, ela é ambos: verdadeira e não verdadeira. Isso constitui uma contradição. E, de acordo com muitas teorias lógicas, uma contradição implica trivialidade.

As duas formas do paradoxo do mentiroso que examinamos até o momento dependem de alguma autorreferência explícita — sentenças que falam diretamente de si mesmas. Contudo, tal autorreferência explícita pode ser evitada, como será demonstrado pela próxima família de paradoxos do mentiroso.

1.3 Ciclos do Mentiroso

Considere um diálogo muito conciso (isto é, com uma sentença cada) entre os irmãos Max e Agnes:

Max: *A afirmação de Agnes é verdadeira.*

Agnes: *A afirmação de Max não é verdadeira.*

O que Max disse é verdadeiro se, e somente se, o que Agnes disse é verdadeiro. Contudo, o que Agnes disse (a saber: ”*A afirmação de Max não é verdadeira.*”) é verdadeiro se, e somente se, o que Max disse não é verdadeiro. Portanto, o que Max disse é verdadeiro se, e somente se, o que Max disse não é verdadeiro. Entretanto, se o que Max disse é ou verdadeiro ou não verdadeiro, então é ambos: verdadeiro e não verdadeiro. E isso constitui, assim como nos casos de FMentiroso e NMentiroso, uma contradição, a qual implica, segundo muitas teorias lógicas, uma trivialidade.

Paradoxos do mentiroso também podem ser formulados por meio de estruturas sentenciais mais complexas, e não apenas por modos de referência complexos. Um exemplo particularmente relevante envolve compostos booleanos.

1.4 Compostos Booleanos

Compostos booleanos podem ser incorporados a sentenças do mentiroso de diversas maneiras. Um exemplo relativamente simples é o seguinte. Considere a sentença denominada DMentiroso (D para disjuntivo):

DMentiroso: *Ou DMentiroso não é verdadeira, ou 1=0.*

Primeiramente, observe que, se DMentiroso não é verdadeira, então ela deve ser verdadeira. Se DMentiroso não é verdadeira, então, por raciocínio análogo ao que vimos anteriormente, a primeira disjunta de DMentiroso é verdadeira. Uma disjunção é verdadeira se ao menos uma de suas disjuntas o for, logo, DMentiroso é verdadeira. Assim, se DMentiroso não é verdadeira, ela é ambos: verdadeira e não verdadeira, o que constitui uma contradição. Por *reductio ad absurdum*, conclui-se, então, que DMentiroso é verdadeira; portanto, uma de suas disjuntas deve ser verdadeira. Se for a primeira, incorremos em contradição; logo, deve ser a segunda. Concluímos, assim, que 1=0. Dessa forma, demonstramos que 1=0. Além disso, a sentença "1=0" não desempenhou nenhum papel substancial no raciocínio acima. Poderíamos substituí-la por qualquer outra sentença e obter, desse modo, uma prova dessa sentença.

Fizemos uma pausa para mencionar DMentiroso, porque ela está relacionada a outro paradoxo importante: o paradoxo de Curry, o qual envolve condicionais que afirmam de si mesmas apenas se forem verdadeiras (isto é, se a própria condicional for verdadeira), então alguma absurdidade também o será (por exemplo: "*se esta sentença é verdadeira, então 1=0*", ou "*se esta sentença é verdadeira, então tudo é verdadeiro*", e assim por diante). Pelo menos em linguagens nas quais o condicional é o condicional material, de modo que $A \supset B$ é equivalente a $\neg A \vee B$, DMentiroso é equivalente à sentença de Curry "DMentiroso é verdadeira $\supset$ 1=0". Embora essa equivalência estabeleça certas conexões entre o paradoxo do Mentiroso e o paradoxo de Curry, convém assinalar uma diferença fundamental. Para o paradoxo de Curry é mais relevante onde está o condicional do que o condicional material (ou alguma variante modalizada deste). Nesses contextos, o paradoxo de Curry não ostenta a negação de forma explícita, como ocorre com DMentiroso. Para maiores esclarecimentos, consulte o verbete sobre o paradoxo de Curry.[4]

[4] N.T.: https://plato.stanford.edu/entries/curry-paradox/

1.5 Sequências Infinitas

A questão de saber se o paradoxo do mentiroso exige, de fato, algum tipo de circularidade tem sido objeto de amplo debate. Os ciclos do Mentiroso (como, por exemplo, o diálogo entre Max e Agnes) demonstram que a autorreferência explícita não é necessária; contudo, é evidente que tais ciclos envolvem uma referência circular. Yablo (1993b) argumentou que um paradoxo mais complexo, composto por uma sequência de múltiplas sentenças, pode gerar um paradoxo do Mentiroso sem circularidade.

O paradoxo de Yablo fundamenta-se em uma sequência infinita de afirmações A_0, A_1, A_2, ..., em que cada A_i afirma que todas as A_k maiores (isto é, todas aquelas para as quais $k > i$) são não verdadeiras. Em outras palavras, cada sentença declara que todas as sentenças subsequentes são não verdadeiras. Dado que estamos diante de uma sequência infinita, esta versão do paradoxo do Mentiroso parece evitar o tipo de circularidade evidente nos exemplos anteriores. No entanto, a contradição ainda parece emergir. Se A_0 é verdadeira, então todas as sentenças maiores A_k são não verdadeiras, e *a fortiori* A_1 é não verdadeira. Mas isso implica que há ao menos uma sentença A_k (com $k > 1$) que é verdadeira, o que contradiz A_0. Por outro lado, se A_0 é não verdadeira, então existe ao menos uma sentença verdadeira A_k maior que A_0. Supondo que A_m seja tal sentença (ou seja, uma sentença verdadeira maior que A_0), temos que A_{m+1} é não verdadeira, e nesse caso, há alguma afirmativa verdadeira maior que A_{m+1}. Contudo, isso contradiz A_m. Em vista disso, se A_0 (a primeira sentença da sequência infinita) é ou verdadeira ou não verdadeira, então ela é ambas: verdadeira e não verdadeira. E isso constitui, como nos casos anteriores, uma contradição.[5]

2. Elementos Básicos

Já observamos um tipo característico de raciocínio associado ao paradoxo do Mentiroso. Também identificamos certa estrutura comum presente em todos os exemplos analisados, como a ocorrência de predicados de verdade e de algo análogo à negação. Faremos aqui uma pausa na exposição para discutir esses elementos constituintes do paradoxo, com foco nos paradoxos do Mentiroso mais básicos. O que exatamente dá origem ao paradoxo do

[5] A questão sobre se o paradoxo de Yablo evita de fato a autorreferência é objeto de intenso debate. Consulte, por exemplo, Barrio (2012), Beall (2001), Cook (2006, 2014), Ojea (2012), Picollo (2012), Priest (1997), Sorensen (1998) e Teijeiro (2012).

Mentiroso, e quais, entre os enigmas que acabamos de examinar, podem ser considerados verdadeiramente "básicos", são objetos de controvérsia; uma vez que diferentes abordagens para a solução do paradoxo adotam visões distintas sobre esses pontos. Assim, nosso objetivo, nesta seção, é esclarecer certos temas recorrentes entre as diversas variantes do paradoxo, e não oferecer um diagnóstico exaustivo de sua fonte paradoxal.

Destacamos três aspectos do paradoxo do Mentiroso: o papel desempenhado pelos predicados de verdade, os tipos de princípios utilizados no raciocínio acerca da verdade que são necessários, e a maneira pela qual um paradoxo pode ser derivado a partir desses recursos.

2.1 Predicado de Verdade

O primeiro elemento necessário para a construção de um paradoxo do mentiroso é um predicado de verdade, que denotaremos aqui por Tr. Seguimos o costume usual na lógica de usar Tr como um predicado de sentenças. No entanto, especialmente ao considerarmos certas abordagens para a resolução do paradoxo, convém lembrar que tal tratamento deve ser visto mais como uma conveniência expositiva do que como um compromisso sério acerca dos portadores da verdade.

Assumimos que, juntamente com o predicado de verdade, dispomos de nomes adequados para sentenças. Para uma dada sentença A, suponha que $\ulcorner A \urcorner$ seja um nome dessa sentença. A predicação da verdade a A então assume a forma $Tr(\ulcorner A \urcorner)$.

Digamos que um predicado $Tr(x)$ é um predicado de verdade para uma linguagem $\mathcal{L}$ somente se $Tr(\ulcorner A \urcorner)$ for bem-formada para toda sentença A de $\mathcal{L}$. Tipicamente, espera-se que Tr obedeça a certos princípios que regem seu comportamento sobre as sentenças de uma dada linguagem. É a esses princípios que nos voltaremos a seguir.[6]

[6]Uma nota sobre a terminologia: ao falarmos de uma linguagem, entendemos uma *linguagem interpretada*, que inclui uma sintaxe, uma interpretação ou modelo (contendo um domínio de objetos e interpretações para o vocabulário não lógico), bem como um esquema de valoração que determina a verdade no modelo para fórmulas complexas. Para fins expositivos, adotamos, por vezes, uma linguagem mais simplificada e não rigorosa, esmaecendo a distinção entre uma lógica e uma linguagem (concebendo, de maneira não estritamente precisa, o esquema de valoração como o aspecto relevante da "consequência lógica" para a linguagem). Ademais, em certos momentos, é uma *teoria*, e não propriamente uma linguagem, que está em questão. Assinalaremos quando tais distinções forem relevantes. Dada a ampla gama de ideias lógicas que abordamos de forma relativamente breve neste ensaio, deixamos muitas dessas sutilezas para exposições mais extensas.

2.2 Princípios da Verdade

A tradição, que remonta a Tarski (1935), sustenta que o comportamento do predicado de verdade Tr é descrito pela seguinte bicondicional:

$$Tr(\ulcorner A \urcorner) \leftrightarrow A.$$

De fato, Tarski considerava que a bicondicional em questão era a bicondicional material da lógica clássica. Essa bicondicional é usualmente denominada *esquema-T* [*T-schema*][7]. Para uma análise mais aprofundada do esquema-T e das concepções de verdade em Tarski, veja as entradas dedicadas a Alfred Tarski e às definições de verdade segundo Tarski.[8]

O paradoxo do Mentiroso tem sido um ponto central na reflexão sobre lógicas não clássicas (como já tivemos uma amostra, por exemplo, na ideia de que o princípio da bivalência poderia ser rejeitado como parte de uma solução para o paradoxo do Mentiroso). Assim, devemos deter-nos para considerar quais princípios devem reger o predicado de verdade Tr no caso de a lógica clássica não ser mantida.

A principal ideia quanto ao que poderia substituir o esquema-T aponta para dois tipos de 'regras' (isto é, para dois tipos de 'regras de inferência', em certo sentido) ou princípios característicos do predicado de verdade. Se se tem uma sentença A, pode-se inferir $Tr(\ulcorner A \urcorner)$, isto é, pode-se 'capturar' A com o predicado de verdade. Inversamente, se se tem $Tr(\ulcorner A \urcorner)$, pode-se inferir A, isto é, pode-se 'liberar' A do predicado de verdade. Em algumas lógicas, captura e liberação acabam sendo equivalentes ao esquema-T, mas frequentemente é útil examiná-las separadamente:

Captura: A implica $Tr(\ulcorner A \urcorner)$. (Que também pode ser escrito como $A \vdash Tr\ (\ulcorner A \urcorner)$.)

Liberação: $Tr(\ulcorner A \urcorner)$ implica A. (Que também pode ser escrito como $Tr(\ulcorner A \urcorner) \vdash A$.)

O termo 'implica' aqui é empregado em um sentido lógico, embora o exato significado, assim como as alternativas disponíveis, dependa da lógica de fundo assumida. Para os fins desta discussão, consideramos tal implicação sob a forma da regra: o argumento de A para B é válido, o que registramos (como acima) por meio do símbolo de derivabilidade, $\vdash$. Em

[7]N.T.: O uso de colchetes, com exceção dos usados nos raciocínios lógicos, indica interferência do tradutor no texto com intuito de elucidar algum conceito ou informar o termo no seu original em inglês.

[8]N.T.: https://plato.stanford.edu/entries/tarski/
https://plato.stanford.edu/entries/tarski-truth/

certos contextos lógicos (por exemplo, na lógica clássica, na qual vale o chamado ‘teorema de dedução’), isso é equivalente à demonstrabilidade de uma condicional; contudo, em outros contextos, tal equivalência não se mantém. Em qualquer dos casos, captura e liberação, tomadas em conjunto, tornam A e $Tr(\ulcorner A \urcorner)$ logicamente equivalentes, no sentido de serem inter deriváveis. Em formas mais fortes, captura e liberação podem conduzir à plena intersubstitutividade de A e $Tr(\ulcorner A \urcorner)$ em contextos extensionais. Como discutiremos com mais profundidade na seção 4.1, isso é relevante para certas concepções acerca da natureza da verdade. Assim, o símbolo $\vdash$ está sendo usado aqui como um marcador esquemático para uma gama de noções lógicas distintas, cada qual oferecendo um conceito de inferência válida em uma dada teoria lógica.

(Há aqui uma série de sutilezas lógicas que não exploraremos, especialmente sobre como se dá a formulação de regras e à questão de quais regras são consistentes. Diferentes formulações dessas regras variam significativamente quanto à sua força lógica[9]. Para uma análise mais aprofundada sobre como formular, em lógica clássica, formas consistentes das regras de captura e liberação, veja a entrada sobre teorias axiomáticas da verdade.[10] Na terminologia de Friedman e Sheard (1987), as formas de regra de captura e liberação são denominadas, respectivamente, ’T-Intro’ e ’T-Elim’, ao passo que suas formas condicionais são chamadas ’T-In’ e ’T-Out’. Preferimos aqui uma terminologia mais abrangente, pois ela evidencia um padrão geral de comportamento comum a uma grande variedade de predicados e operadores, por exemplo: o *conhecimento* admite liberação, mas não captura; a *possibilidade* admite captura, mas não liberação; e assim por diante. A verdade, por sua vez, é especial por admitir ambas.)

2.3 O Mentiroso em Resumo

O paradoxo do Mentiroso tem início em uma linguagem que contém um predicado de verdade, o qual obedece a alguma forma das regras de captura e liberação. Passamos agora a examinar com maior precisão como um paradoxo resulta dessas suposições.

[9]Deste modo, nossa terminologia de “formas de regra” talvez não seja a mais adequada. Para os fins limitados a que nos propomos aqui, utilizamo-la meramente para assinalar a diferença entre um argumento válido, registrado como uma “forma de regra”, e a demonstrabilidade de uma condicional, apresentada como uma “forma condicional”. Embora esta não seja a única distinção lógica possível, é aquela que assumirá maior importância no desenvolvimento de nossa análise.

[10]N.T.: https://plato.stanford.edu/entries/truth-axiomatic/

2.3.1 Existência de Sentenças Similares ao Mentiroso

Deixando de lado os paradoxos do tipo Yablo, o paradoxo do Mentiroso depende de alguma forma de autorreferência, seja direta, como nos casos simples do Mentiroso já apresentados, seja indireta, como nos ciclos do Mentiroso. A maioria das línguas naturais não encontra grandes dificuldades em gerar autorreferência. A primeira frase deste ensaio é um exemplo disso. A autorreferência pode ser acidental, como no caso em que alguém escreve: "A única sentença no quadro da sala 101 não é verdadeira", tendo-a escrito, por acaso, na própria sala 101 (como observou C. Parsons, 1974).

Em linguagens formais, a autorreferência também é fácil de encontrar. Qualquer linguagem capaz de expressar alguma sintaxe básica pode gerar sentenças autorreferenciais por meio da chamada diagonalização (ou, mais propriamente, qualquer linguagem dotada de uma teoria adequada da sintaxe ou da aritmética).[11] Uma linguagem que contenha um predicado de verdade e essa sintaxe básica terá, então, uma sentença M tal que M implica $\neg Tr(\ulcorner M \urcorner)$, e vice-versa:

$$M \dashv\vdash \neg Tr(\ulcorner M \urcorner).$$

Este é um 'ponto fixo' do (predicado composto) $\neg Tr$, e é, na prática, o nosso Mentiroso não-verdadeiro simples.

(Tecnicamente, é mais simples expressar a propriedade do 'ponto fixo' em termos de implicações, como fizemos acima. Mas, intuitivamente, a ideia é que, de certo modo, M 'é apenas' $\neg Tr(\ulcorner M \urcorner)$. Isso pode ser formalizado com mais precisão se pensarmos na sentença do Mentiroso M como originando-se de um *nome* c que denota a sentença $\neg Tr(c)$. Dessa forma, podemos compreender a existência do Mentiroso como refletida na identidade

[11] A situação no tocante às linguagens formais é, na verdade, um pouco mais sutil do que nossa breve exposição dá a entender. Na maioria dos casos, as citações de canto [*corner quotes* ou também chamadas citações de Quine], indicam termos formais correspondentes aos números de Gödel das sentenças, e não são aspas genuínas no sentido habitual (isto é, não denotam diretamente a expressão contida entre elas). Por conseguinte, o sentido em que tais linguagens fazem referência a sentenças é delicado. No entanto, com mínimos recursos, é possível representar a sintaxe e construir sentenças diagonais. Assim, há um sentido, ainda que sutil, no qual tais linguagens conseguem expressar autorreferência. Sobre este tema, consulte-se as entradas relativas à lógica da demonstrabilidade [https://plato.stanford.edu/entries/logic-provability/] e Gödel [https://plato.stanford.edu/entries/goedel/] (especialmente a seção sobre os teoremas da incompletude), bem como Heck (2007).

$c = \ulcorner \neg Tr(c) \urcorner$. Para mais detalhes sobre essa abordagem, ver Heck 2012.)

2.3.2 Outras "Leis" Lógicas

Outros elementos evidentes nos paradoxos do Mentiroso mais comuns dizem respeito ao comportamento lógico dos conectivos básicos ou às características da implicação. Alguns dos princípios relevantes são os seguintes:

- Princípio do Terceiro Excluído (PTE): $\vdash A \vee \neg A$.

  [Para qualquer proposição A, ou A é verdadeira ou sua negação $\neg A$ é.]

- Explosão (EFQ – *Ex Falso Quodlibet*)[12]: $A, \neg A \vdash B$.

  [A partir de uma contradição, isto é, A e $\neg A$, qualquer proposição B pode ser deduzida.]

- Princípio da Disjunção (PD)[13]: Se $A \vdash C$ e $B \vdash C$, então $A \vee B \vdash C$.

  [Se C pode ser deduzida tanto de A quanto de B, então pode ser deduzida da disjunção $A \vee B$.]

- Adjunção: Se $A \vdash B$ e $A \vdash C$, então $A \vdash B \wedge C$.

  [Se B e C podem ser deduzidas a partir de A, então também se pode deduzir sua conjunção $B \wedge C$ a partir de A.]

(Isto não quer dizer que essas sejam as únicas características lógicas envolvidas nos paradoxos do Mentiroso mais comuns, mas pode-se argumentar que são as mais importantes entre as mais evidentes.)

[12]É este princípio ou regra que reiteramos anteriormente ao afirmar que, segundo os critérios de muitas teorias lógicas, uma contradição arbitrária implica absurdo ou trivialidade, no sentido de que dela se seguem todas as sentenças. Este princípio é frequentemente designado por seu título clássico *ex falso quodlibet*, razão pela qual sua abreviação costumeira é EFQ, apesar do nome alternativo "explosão".

[13]Este princípio é frequentemente denominado '$\vee$-Out', '$\vee$-Elim' ou, de maneira mais sugestiva, '*raciocínio por casos*'.

2.3.3 O Paradoxo do Mentiroso em Abstrato

Dado os elementos discutidos anteriormente, podemos apresentar uma forma ligeiramente mais abstrata do paradoxo. (Nosso objetivo é utilizar essa forma abstrata para destacar diferentes respostas ao paradoxo.) Suponhamos que temos uma linguagem $\mathscr{L}$ com um predicado de verdade Tr, e que $\mathscr{L}$ possui sintaxe suficiente para construir uma sentença M tal que $M \dashv\vdash \neg Tr(\ulcorner M \urcorner)$. Suponhamos também que a lógica de $\mathscr{L}$ admite os princípios do Terceiro Excluído (PTE) e da Explosão (EFQ), além de satisfazer os princípios da Disjunção (PD) e da Adjunção.

Um argumento de que nossa sentença do Mentiroso M, implica uma contradição procede da seguinte forma:

1. $Tr(\ulcorner M \urcorner) \vee \neg Tr(\ulcorner M \urcorner)$ (PTE)
2. Caso Um:
   a. $Tr(\ulcorner M \urcorner)$
   b. M (2a: liberação)
   c. $\neg Tr(\ulcorner M \urcorner)$ (2b: definição de M)
   d. $\neg Tr(\ulcorner M \urcorner) \wedge Tr(\ulcorner M \urcorner)$ (2a, 2c: adjunção)
3. Caso Dois:
   a. $\neg Tr(\ulcorner M \urcorner)$
   b. M (3a: definição de M)
   c. $Tr(\ulcorner M \urcorner)$ (3b: captura)
   d. $\neg Tr(\ulcorner M \urcorner) \wedge Tr(\ulcorner M \urcorner)$ (3a, 3c: adjunção)
4. $\neg Tr(\ulcorner M \urcorner) \wedge Tr(\ulcorner M \urcorner)$ (1–3: PD)

Esta versão do paradoxo do Mentiroso é apenas uma entre muitas. Com um pouco mais de complexidade, por exemplo, pode-se evitar tanto a captura quanto a liberação, substituindo-as por outros pressupostos de fundo. Também existem variantes intuicionistas do paradoxo do Mentiroso, embora não exploraremos aqui a lógica intuicionista.[14]

Até agora, mostramos que, com os elementos fornecidos, nossa sentença do Mentiroso M implica uma contradição (formalizando assim o raciocínio presente em NMentiroso). A partir daí, é um passo curto até o completo absurdo — isso, claro, se a mera contradição já não for suficientemente absurda por si só. Invocamos, então, o princípio da explosão (EFQ) para concluir a demonstração. (Também assumimos que $A \wedge B$ implica A e B, ou seja, que a simplificação é válida em $\mathcal{L}$; mas, essa suposição não é realmente necessária.)

5. B (4: EFQ)

B, aqui, pode ser qualquer — toda e qualquer — sentença que você desejar (ou que lhe desagrade, conforme o caso)! O princípio da explosão afirma que de uma contradição qualquer sentença se segue; ele legitima a passagem de uma única contradição à trivialidade lógica.

Diante de tamanho absurdo (isto é, trivialidade), concluímos que há algo de errado no raciocínio do Mentiroso apresentado acima. A pergunta que se impõe é: o quê? Esta, em última instância, é a questão fundamental que o paradoxo do Mentiroso coloca.

3. Significado

Já vimos que, com alguns pressupostos elementares acerca da verdade e da lógica, ocorre um desastre lógico. Mas qual é o significado mais amplo de tal resultado?

De tempos em tempos, argumenta-se que o paradoxo do Mentiroso revela algo de profundo sobre a própria filosofia. Por exemplo, Grim (1991) defende que o paradoxo mostra que o mundo é, de certo modo, essencialmente "incompleto" e que não pode haver um ser onisciente. McGee (1991) e outros sugerem que o paradoxo revela que a noção de verdade é uma noção vaga. Glanzberg (2001) sustenta que o paradoxo nos mostra algo importante

[14] Novamente para o caso clássico, Friedman e Sheard (1987) fornecem uma lista exaustiva de teorias inconsistentes relativas a uma teoria base relativamente fraca. Eles demonstram que qualquer uma das formas condicionais clássicas de captura ou liberação são inconsistentes por si mesmas, em relação à teoria de base suplementada pela escolha adequada de princípios que asseguram a completude $(Tr(\ulcorner A \urcorner) \vee Tr(\ulcorner \neg A \urcorner))$ ou a consistência $(\neg(Tr(\ulcorner A \urcorner) \wedge Tr(\ulcorner \neg A \urcorner)))$ da verdade.

sobre a natureza da dependência de contexto na linguagem, enquanto Eklund (2002) defende que ele revela algo essencial sobre a competência semântica e as línguas que falamos. Gupta e Belnap (1993) alegam que o paradoxo traz à tona propriedades importantes da noção geral de definição. E há outras lições, bem como variações dessas, que também têm sido extraídas do paradoxo.

A preocupação mais imediata, ao menos para os nossos propósitos aqui, é o que o paradoxo do Mentiroso nos revela acerca dos princípios fundamentais que regem a noção de verdade e acerca da própria lógica. Em tom cético, o próprio Tarski (1935, 1944) parece ter entendido que o paradoxo evidencia a incoerência da noção ordinária de verdade, a qual exigiria substituição por uma noção mais cientificamente respeitável. (Para mais informações sobre Tarski, veja o verbete sobre Tarski[15] e sobre as definições de verdade de Tarski[16]. Para saber sobre suas metas e propósitos, consulte Heck 1997.) A mais comum, e talvez predominante entre as tentativas de solução ao paradoxo, é a ideia de que os princípios básicos que regem a verdade são mais sutis do que aquilo que o esquema-T é capaz de expressar.

O paradoxo do Mentiroso também tem sido o núcleo de diversos argumentos contrários à lógica clássica, dado que são precisamente algumas características fundamentais dessa lógica que permitem que os princípios de captura e liberação conduzam ao absurdo. Entre essas características, destacam-se os argumentos em favor de lógicas paracompletas (como nas propostas de Kripke, 1975; Field, 2008) ou paraconsistentes (como em Asenjo, 1966; Priest, 1984, 2006). Contudo, Ripley (2013b) argumenta que é possível preservar a lógica clássica ao mesmo tempo em que se abandonam essas características problemáticas.

Em muitos casos, inspirados por concepções mais amplas acerca do significado do paradoxo, houve diversas tentativas de, de uma forma ou de outra, resolvê-lo. É a essas soluções propostas que nos voltamos agora.

4. Algumas Famílias de Soluções

Nesta seção, faremos um breve levantamento de algumas abordagens destinadas a resolver o paradoxo do Mentiroso. Agrupamos as soluções propostas em famílias, procurando explicar as ideias fundamentais que as orientam. Em muitos casos, uma exposição completa

[15]N.T.: https://plato.stanford.edu/entries/tarski/
[16]N.T.: https://plato.stanford.edu/entries/tarski-truth/

exigiria uma apresentação técnica extensa, a qual não será desenvolvida aqui. Os leitores interessados são encorajados a consultar as referências que fornecemos para cada ideia básica.

4.1 Lógicas Paracompletas e Paraconsistentes

Uma das principais ideias para a resolução do paradoxo do Mentiroso é que ele nos revela algo sobre a própria lógica, de fato, algo ainda obscuro sobre a lógica. A ideia central é que os princípios de captura e liberação são princípios conceituais fundamentais que regem a noção de verdade e, por isso, não podem ser modificados. Em vez disso, a lógica básica deve ser uma lógica não clássica, a fim de evitar o desastre lógico do tipo que examinamos na seção 2.

Uma forma importante de motivar soluções não clássicas é recorrer a uma forma de deflacionismo acerca da verdade. Tais visões tomam algo análogo ao esquema-T como característica definidora da verdade e, por isso, não suscetível a modificação (ver, por exemplo, Horwich 1990). De maneira mais estrita, as chamadas concepções de verdade transparentes, ou “translúcidas”, ou “disquotacionais puras” (e.g., Field 1994, 2008; Beall 2005) consideram que a propriedade definidora da verdade consiste na intersubstitutividade, em todos os contextos não opacos, a sentença A por $Tr(\ulcorner A \urcorner)$ e vice-versa. Isso faz com que os princípios de captura e liberação, aplicados de forma irrestrita a todas as sentenças de uma linguagem, sejam exigências da própria noção de verdade (ao menos onde temos $A \vdash A$ ou, mais fortemente, $\vdash A \to A$).[17] Consulte o verbete sobre verdade para uma discussão mais aprofundada.[18]

Manter fixos os princípios de captura e liberação, aplicando-os a todas as sentenças sem restrição, conduz à trivialidade, salvo se a lógica for não clássica. Há duas principais subfamílias de teorias da verdade não-clássicas (transparentes): as *paracompletas* e as *paraconsistentes*. A seguir, delineamos as ideias centrais de cada uma dessas teorias.

[17] Beall e Glanzberg (2008) argumentam em prol de uma estreita conexão entre concepções gerais acerca da natureza da verdade e os caminhos possíveis para a resolução do paradoxo.
[18] N.T.: https://plato.stanford.edu/entries/truth/

4.1.1 Paracompleta

Segundo as abordagens paracompletas do paradoxo do Mentiroso, a principal lição do paradoxo é que o princípio do Terceiro Excluído (PTE) 'falha' em algum aspecto. Em outras palavras: o paradoxo do Mentiroso nos ensina que algumas sentenças (notadamente, as sentenças mentirosas) 'nem se sustentam nem deixam de se sustentar' (em certo sentido), ou seja, não são nem verdadeiras nem falsas. Como consequência, a lógica da verdade deve ser não clássica.

Essa ideia talvez pareça mais natural em resposta ao Mentiroso de falsidade simples. Nesse caso, é tentador afirmar que há um status distinto da verdade e da falsidade, e que a sentença mentirosa M possui esse status intermediário. Contudo, tal resposta não basta, por exemplo, para o Mentiroso não-verdadeiro simples, que não afirma nada sobre falsidade. Assim, o raciocínio básico revisado na seção 2.3 deve falhar, e segundo a perspectiva paracompleta, o responsável é o PTE. As instâncias do Mentiroso no PTE 'falham' (em algum aspecto) de acordo com a perspectiva paracompleta; tais sentenças caem no 'vazio' entre a verdade e a falsidade (para empregar uma metáfora comum).

Diversas propostas foram formuladas com o intuito de empregar lógicas não clássicas na resolução do paradoxo do Mentiroso. Um dos primeiros exemplos pode ser encontrado nos trabalhos de van Fraassen (1968, 1970). No entanto, o trabalho mais influente nas últimas décadas tem sido o de Kripke, cuja abordagem impactou não apenas as tentativas de resolução baseadas em lógica não clássica, mas também uma série de outras estratégias que abordaremos na seção 4.2. Dada essa relevância, convém fazer uma breve exposição da estrutura teórica proposta por Kripke.

4.1.2 A Teoria de Kripke

Lógicas nas quais o PTE falha não são particularmente difíceis de encontrar. Entre as várias opções disponíveis, há um número de lógicas trivalentes, que admitem um terceiro valor de verdade além de verdadeiro e falso. Sentenças como as do tipo Mentiroso recebem esse terceiro valor. Uma das lógicas mais amplamente utilizadas nesse contexto é a lógica de Kleene forte (K_3). Não entraremos aqui nos detalhes técnicos da lógica K_3, mas mencionaremos apenas as propriedades de K_3 que precisaremos. (Para uma análise mais aprofundada, ver o verbete sobre lógica multivalorada[19], ou Priest 2008.) Antes de tudo, é

[19]N.T.: https://plato.stanford.edu/entries/logic-manyvalued/

importante notar:

$$\nvdash_{K_3} A \vee \neg A.$$

O PTE falha. De fato, segundo a lógica K_3, não há verdades lógicas (ou sentenças logicamente válidas). (Voltaremos a este ponto ao tratar da questão de um 'condicional adequado', mais adiante.)

O desafio ao se empregar K_3 para fundamentar uma teoria paracompleta consiste em explicar como princípios como captura e liberação (mesmo como forma de regra) poderiam ser preservados. E, caso se siga a linha deflacionista, o desafio é ainda maior: explicar como captura e liberação irrestritas e completas poderiam ser mantidas. Uma maneira de compreender o trabalho fundamental de Kripke (1975) (e os trabalhos relacionados de Martin e Woodruff (1975)) é vê-lo justamente como uma tentativa de alcançar esse objetivo.

Kripke inicia com uma linguagem totalmente clássica $\mathcal{L}_0$, que não contém predicado de verdade (ou, mais geralmente, nenhum termo semântico). (Recordemos que estamos assumindo que uma linguagem vem equipada com um esquema de valoração. No caso de $\mathcal{L}_0$ esse esquema é clássico.) Ele então considera estender essa linguagem a uma nova linguagem $\mathcal{L}_0^+$, que contém um predicado de verdade Tr. O predicado Tr é tomado como aplicável a toda sentença da linguagem expandida $\mathcal{L}_0^+$, inclusive àquelas da linguagem original $\mathcal{L}_0$. Assim, trata-se de um predicado de verdade autoaplicável (como exige a perspectiva inspirada no deflacionismo que mencionamos anteriormente), mesmo que se tenha começado com uma linguagem desprovida de predicado de verdade.

Podemos pensar a linguagem $\mathcal{L}_0^+$ como sendo interpretada por um modelo clássico $\mathcal{M}_0$. Kripke nos mostra como construir uma interpretação $\mathcal{M}_0^+$ para a linguagem expandida. A principal inovação está em conceber o predicado de verdade como parcial. Em vez de possuir apenas uma extensão, ele possui tanto uma extensão (o conjunto das sentenças das quais é verdadeiro) quanto uma antiextensão (o conjunto das sentenças das quais é falso). A extensão e a antiextensão são mutuamente exclusivas, mas não precisam juntas esgotar o domínio de $\mathcal{M}_0$. Sentenças patológicas, como a sentença do Mentiroso M, não pertencem nem à extensão nem à antiextensão de Tr. (De fato, poderíamos ter interpretado a linguagem-base $\mathcal{L}_0$ também por um modelo parcial. No entanto, na aplicação pretendida, a parcialidade surge apenas com predicados semânticos, como Tr.)

Não pertencer nem à extensão nem à antiextensão de Tr funciona como ter um terceiro valor de verdade, e assim podemos interpretar $\mathcal{L}_0^+$ atuando como uma linguagem com um esquema de valoração K_3. Tratando a linguagem desta forma, Kripke mostra como construir uma extensão e uma antiextensão muito plausíveis para Tr, normalmente denotadas por $\mathcal{E}$ e $\mathcal{A}$, respectivamente. A propriedade importante do novo modelo estendido $\langle \mathcal{M}_0, \langle \mathcal{E}, \mathcal{A} \rangle \rangle$ é que o valor de verdade de qualquer sentença A e da sentença $Tr(\ulcorner A \urcorner)$ é exatamente o mesmo. A é verdadeira, falsa, ou nenhuma das duas se $Tr(\ulcorner A \urcorner)$ também o é. Além disso, interpretando a linguagem expandida $\mathcal{M}_0^+$ como uma linguagem K_3, temos como consequência em K_3 que $A \dashv\vdash Tr(\ulcorner A \urcorner)$, exatamente como se desejava.

Kripke demonstra como construir $\mathcal{E}$ e $\mathcal{A}$ por meio de um processo indutivo. Parte-se de uma 'aproximação' inicial da extensão e antiextensão de Tr, e essa aproximação é aperfeiçoada sucessivamente, até que o processo de aprimoramento deixe de ser produtivo — isto é, atinge-se um ponto fixo. De fato, no contexto da solução baseada em K_3, o procedimento mais natural consiste em começar com extensões e antiextensões vazias, e então adicionar as sentenças que se revelam verdadeiras nas etapas subsequentes do processo. A construção de Kripke pode ser aplicada a diversas lógicas distintas, incluindo outras lógicas multivaloradas, como a lógica de Kleene fraca [*Weak Kleene*], bem como lógicas sobrevaloradas [*supervaluation logics*]. Para uma discussão mais aprofundada veja, por exemplo, Burgess (1986) e McGee (1991). As construções ao estilo de Kripke envolvem um grau considerável de sutileza matemática. Para uma exposição acessível de mais detalhes técnicos e conceituais, recomenda-se Soames (1999). Para uma apresentação mais rica em termos matemáticos, veja McGee (1991).

4.1.3 Condicionais adequados

Lógicas como K_3 sofrem com a ausência de um condicional natural ou 'adequado' (em particular, de um que satisfaça as propriedades $A, A \to B \vdash B$ e $\vdash A \to A$). Essa limitação revela uma deficiência importante na abordagem kripkeana ao paradoxo do Mentiroso. A linguagem expandida $\mathcal{L}_0^+$ não consegue expressar, em forma condicional, as propriedades de captura e liberação da verdade (isto é, T-bicondicionais): Tr é transparente nessa formulação, então $Tr(\ulcorner A \urcorner)$ e A são plenamente intersubstituíveis. Contudo, nesta teoria, a disjunção $\neg A \vee A$ não é verdadeira para todas as sentenças A, e, portanto, tampouco $\neg Tr(\ulcorner A \urcorner) \vee A$ é verdadeira para todo A. Porém, $\neg Tr(\ulcorner A \urcorner) \vee A$ é logicamente equivalente, nesta teoria, a $Tr(\ulcorner A \urcorner) \to A$, dado que o condicional usado é o condicional material. Conclui-se, então, que a construção de Kripke aqui considerada não satisfaz todos os T-bicondicionais, os quais seriam os candidatos naturais para expressar, dentro da própria teoria, os traços fundamentais de captura e liberação característicos do conceito de verdade.

Um avanço recente e importante no sentido de suplementar a estrutura de Kripke com um condicional adequado é o trabalho de Field (2008). A teoria de Field representa um progresso significativo, mas é suficientemente complexa para ultrapassar os limites desta introdução (muito elementar). Os leitores interessados devem consultar a própria exposição de Field para obter uma noção de como tal modificação pode ser realizada. Ver Field (2008), e discussões adicionais em Beall (2009).

Um uso importante dos condicionais na lógica é a formalização da quantificação universal restrita, expressando a conexão entre A e B na forma "Todos os As são Bs". Isso tem desempenhado recentemente um papel central em várias discussões sobre condicionais e paradoxos; ver, por exemplo, Beall et al. (2006); Beall (2011); Field (2014); e Ripley (2015).

4.1.4 Paraconsistente

Como mencionamos, duas abordagens importantes para o paradoxo do Mentiroso que se concentram em lógicas não clássicas são as abordagens paracompleta e paraconsistente. Esboçamos acima uma opção paracompleta. Voltamo-nos agora para uma opção paraconsistente. Aqui, a ideia básica é permitir a contradição (por exemplo, até e incluindo o passo 4 da derivação na seção 2.3.3), mas alterar a lógica rejeitando o princípio da explosão — e, assim, evitar o absurdo envolvido no passo 5.

Assim como a abordagem paracompleta que acabamos de examinar, as abordagens paraconsistentes ao paradoxo do Mentiroso encontram uma motivação fácil e natural em concepções de verdade transparentes ou de outra forma adequadamente 'minimalistas', que exigem a plena intersubstitutividade entre A e $Tr(\ulcorner A \urcorner)$ e, portanto, não podem restringir os princípios de captura e liberação. No entanto, as abordagens paraconsistentes também encontraram motivação em uma concepção antideflacionista inspirada em Dummett, que leva a sério o papel da verdade como objetivo da asserção (cf. Dummett 1959). Com efeito, Priest (2006) argumenta que essa concepção (não transparente) da verdade motiva tanto o esquema-T quanto o PTE, e que isso implica que a sentença mentirosa M é simultaneamente verdadeira e não verdadeira. Assim, de acordo com qualquer linha dialeteica (segundo a qual ao menos uma sentença é ao mesmo tempo verdadeira e não verdadeira), a única opção é rejeitar o EFQ.

4.1.5 Dialeteísmo

Priest (1984, 2006) tem sido uma das vozes mais influentes na defesa de uma abordagem paraconsistente para a resolução do paradoxo do Mentiroso. Ele propôs uma lógica paraconsistente (e não paracompleta) hoje conhecida como LP [*Logic of Paradox*], a qual preserva o princípio do terceiro excluído, mas rejeita o princípio da explosão.[20] Essa lógica possui a característica distintiva de permitir contradições verdadeiras. É essa a abordagem da verdade que Priest denomina dialeteica. (Para uma discussão mais extensa, ver o verbete sobre dialeteísmo.[21])

Formalmente, a LP pode ser entendida como uma lógica trivalente; mas enquanto K_3 apresenta lacunas de valor de verdade, LP apresenta excessos de valor de verdade [*gluts*]. Assim, em LP, há sentenças que podem ser simultaneamente verdadeiras e falsas. Contudo, como discutiremos mais detalhadamente na seção 4.1.3, a maneira adequada de descrever tanto lacunas quanto excessos é uma questão delicada. Por ora, limitamo-nos à observação preliminar de que, da mesma forma que K_3 pode ser caracterizada como apresentando lacunas — por admitir um terceiro valor de verdade —, LP pode, de modo correspondente, ser caracterizada como contendo excessos.

De modo análogo, técnicas ao estilo de Kripke podem ser aplicadas para produzir uma interpretação para um predicado de verdade, partindo-se de uma linguagem clássica $\mathcal{L}_0$ que não contenha tal predicado. Novamente, são atribuídas uma extensão e uma antiextensão ao predicado Tr. Enquanto a construção original de Kripke previa que a extensão e a antiextensão fossem disjuntas, mas sem necessariamente esgotar o domínio, neste caso admite-se que a extensão e a antiextensão se sobreponham, desde que, em conjunto, esgotem o domínio do modelo. Isso implementa a ideia de excessos, da mesma forma que a versão anterior implementa a ideia de lacunas. Técnicas relacionadas às de Kripke podem, então, ser empregadas para construir uma extensão e uma antiextensão para Tr. O resultado, mais uma vez, é uma interpretação na qual A e $Tr(\ulcorner A \urcorner)$ recebem o mesmo valor de verdade no modelo.

Essa construção não foi apresentada pelo próprio Kripke, mas variantes dela foram desenvolvidas por diversos autores, incluindo Dowden (1984), Leitgeb (1999), Priest (1984, 2006), Visser (1984) e Woodruff (1984).

[20] A lógica foi inicialmente desenvolvida por Asenjo (1966), embora tenha sido popularizada pelos trabalhos de Priest. A denominação 'LP' é devida a Priest, enquanto o termo “dialeteísmo” foi cunhado por Priest e Routley.
[21] N.T.: https://plato.stanford.edu/entries/dialetheism/

4.1.6 Combinando Paracompletude e Paraconsistência

Embora tenhamos identificado as abordagens paracompleta e paraconsistente ao Paradoxo do Mentiroso como duas opções distintas, elas não são incompatíveis. Com efeito, se forem consideradas como teorias da negação (caso se adote essa perspectiva), pode-se sustentar que a negação não é nem exaustiva nem 'explosiva' — isto é, que não satisfaz nem PTE e nem EFQ. Uma abordagem desse tipo é a teoria da verdade (transparente) baseada em *FDE* [*First-Degree Entailment*, ou Implicação de Primeiro Grau], discutida em Dunn (1969) (ver Gupta e Belnap (1993); Leitgeb (1999); Visser (1984); Woodruff (1984); Yablo (1993a); e — em essência — Brady (1989)).

(As teorias baseadas em LP e as teorias baseadas em K_3 são — pelo menos em um nível (padrão de primeira-ordem) — simplesmente lógicas fortalecidas da lógica mais ampla *FDE*. Para uma discussão geral sobre tais estruturas, consulte, por exemplo, Priest 2008.)

4.1.7 Poder Expressivo e 'Revanche'

Trabalhando em lógica clássica, Tarski (1935) concluiu, de forma célebre, a partir do paradoxo do Mentiroso, que uma linguagem não pode definir seu próprio predicado de verdade. De modo mais geral, ele entendeu que a lição do paradoxo do Mentiroso é que as linguagens não podem expressar toda a gama de conceitos semânticos que descrevem seu próprio funcionamento. Um dos principais objetivos das abordagens não clássicas ao paradoxo do Mentiroso que examinamos aqui é evitar essa conclusão, que muitos consideraram excessivamente drástica. Contudo, o quão bem-sucedidas essas abordagens têm sido a esse respeito permanece uma questão altamente controversa.

Em certo sentido, tanto as abordagens paracompletas quanto as paraconsistentes alcançam o resultado desejado: apresentam linguagens que contêm predicados de verdade que se aplicam a sentenças dessa própria linguagem e possuem a característica de que A e $Tr(\ulcorner A \urcorner)$ têm o mesmo valor de verdade. Nesse aspecto, ambas fornecem linguagens que contêm seus próprios predicados de verdade.

No caso paracompleto, a questão de saber se isso é suficiente tem sido amplamente debatida. A visão paracompleta sustenta que a sentença do Mentiroso M não é nem verdadeira nem falsa, e isso é fundamental para a preservação da consistência. Contudo, observe que a abordagem paracompleta que discutimos acima não pode enunciar esse fato, pois não pode resultar verdadeiro que $\neg Tr(\ulcorner M \urcorner)$. Se isso fosse verdadeiro, então M seria verdadeira, e, em seguida, $Tr(\ulcorner M \urcorner)$ também seria verdadeira, levando-nos novamente à contradição.

Segue-se daí um ponto adicional. Como aludimos acima, isso mostra que K_3, com um predicado de verdade, não será capaz de enunciar o caráter lacunar das lacunas, enquanto LP poderá enunciar tanto propriedades de lacunas quanto de excessos de verdade. Assim, como mencionamos, a condição de lacunas e excessos pode ser bastante complexa.

No que diz respeito ao problema da 'revanche'[22], a principal dificuldade é simplesmente que a abordagem paracompleta não consegue enunciar corretamente a sua própria solução ao paradoxo do Mentiroso. O significado exato desse fato tem sido objeto de debate. É certamente o caso que o conjunto de sentenças verdadeiras no tipo de modelo construído por Kripke não inclui $\neg Tr(\ulcorner M \urcorner)$. Por essa razão, alguns autores, como McGee (1991), T. Parsons (1984) e Soames (1999), sustentaram que o fracasso da sentença do Mentiroso em ser verdadeira é um fato adicional que vai além do que o predicado de verdade precisa expressar, e, portanto, é irrelevante para o êxito da solução do paradoxo. (A posição de McGee possui outro aspecto, que discutiremos na seção 4.2.3.)

Apesar disso, parece haver um fato semântico importante acerca da verdade na linguagem paracompleta, intimamente relacionado — se não idêntico — a um fato sobre a verdade *per se*, que a linguagem não consegue expressar. Assim, tem-se argumentado que essa abordagem falha em alcançar uma teoria da verdade plenamente adequada. O próprio Kripke observa que existem certos conceitos semânticos que não podem ser expressos, e esse argumento foi desenvolvido de maneira mais incisiva por C. Parsons (1974).

Uma maneira de explicitar o que falta na linguagem paracompleta é introduzir uma nova noção de *determinabilidade*, de modo que o status do enunciado do Mentiroso seja o de não ser determinadamente verdadeiro. Se assim for, então a linguagem paracompleta de Kripke não pode expressar esse conceito de determinabilidade. Algumas abordagens que incorporam ideias paracompletas procuraram suplementar o método de Kripke acrescentando noções de verdade determinável. McGee (1991) o faz em um contexto basicamente clássico. Em um contexto não clássico e paracompleto, Field (2008) complementa a abordagem paracompleta básica com infinitos operadores de 'determinadamente', cada um definido em termos do 'condicional adequado' de Field, e cada um proporcionando uma noção diferente (e mais forte) de 'verdade'. (Veja também alguns dos artigos em Beall (ed.) 2008.)

Frequentemente se argumenta em favor das abordagens paraconsistentes que estas

[22]N.T.: também podendo ser traduzido como 'vingança'.

não têm dificuldade em ‘caracterizar’ o status dos enunciados do Mentiroso: eles são verdadeiros e falsos (isto é, são verdadeiros e têm a negação verdadeira). As teorias baseadas em LP conseguem expressar essa caracterização. Por outro lado, alguns autores, como Littmann e Simmons (2004) e S. Shapiro (2004), consideraram que há um problema dual: a saber, caracterizar os enunciados ‘normais’, aqueles que não são simultaneamente verdadeiros e falsos. (Alguns formulam esse problema como o de caracterizar ser apenas verdadeiro.) Se isso constitui realmente um problema é algo que deixamos em aberto. (Para discussões sobre o tema, veja Field 2008 e Priest 2006.)

Outro problema que surge neste contexto é o dos chamados ‘paradoxos de revanche’. Podemos ilustrá-lo com o caso do Mentiroso de falsidade simples. Suponha que se tome este paradoxo como o paradigma do paradoxo do Mentiroso, e que se proponha uma solução simples que rejeita a bivalência. Em resposta, apresenta-se o paradoxo do Mentiroso não-verdadeiro simples, que mina a solução proposta. Este é o padrão de ‘revanche’, em que uma solução para o paradoxo é refutada com base numa forma levemente modificada do próprio paradoxo. Frequentemente são propostos paradoxos de revanche contra soluções paracompletas: muitos pontos em que a linguagem paracompleta falha em expressar algum conceito semântico fornecem meios para construir problemas de revanche. A incapacidade de expressar corretamente o status do Mentiroso não-verdadeiro simples é um exemplo. Outro exemplo envolve a noção de determinabilidade. Se tomarmos a via da determinabilidade e atribuirmos ao enunciado do Mentiroso o status de não ser determinadamente verdadeiro, então é possível construir um problema de revanche por meio de uma sentença que afirma de si mesma que não é determinadamente verdadeira.

De maneira semelhante, argumenta-se que as abordagens paraconsistentes enfrentam um tipo de problema de revanche, pois precisam tratar separadamente o paradoxo de Curry, discutido na seção 1.4, e o paradoxo do Mentiroso. Trata-se de uma questão técnica bastante complexa, pois depende da natureza do condicional utilizado na formulação da sentença de Curry. Se esse condicional obedece à propriedade de destacamento, então ele não pode ser um excesso de valor de verdade, como ocorre com o Mentiroso em abordagens paraconsistentes. No entanto, a questão de saber se essa é a abordagem correta para o condicional tem sido objeto de controvérsia. Para uma discussão mais aprofundada, veja Beall (2014, 2015).

Vimos que algumas abordagens (por exemplo, McGee 1991; em certos aspectos, T. Parsons 1984 e Soames 1999) rejeitam o problema da revanche, enquanto outras buscam solucioná-lo por meio da introdução de aparato adicional (por exemplo, Field 2008). Como discutiremos mais adiante na seção 4.3, visões contextualistas, como as de Burge (1979), Glanzberg (2004a) e C. Parsons (1974), tendem a entender a revanche não como um problema separado, mas como o próprio cerne do fenômeno do Mentiroso. Para uma discussão mais aprofundada sobre a revanche e sua natureza, veja os artigos reunidos em Beall (ed.) (2008) e L. Shapiro (2006).

4.2 Lógicas Subestruturais

Há outra maneira de entender o paradoxo, considerando-o como decorrente de pressupostos equivocados incorporados às lógicas padrão. Essa perspectiva não vê o problema como relacionado a qualquer conectivo específico ou a um elemento particular do vocabulário, mas sim a certas *regras estruturais* que regem a relação de consequência em questão. Essas abordagens, baseadas nas chamadas *lógicas subestruturais*, dividem-se em três principais vertentes: a não-contrativa, a não-transitiva e a não-reflexiva. (Há, na realidade, uma diversidade muito maior entre as lógicas subestruturais do que esta divisão sugere; particularmente, muitas não se enquadram em nenhuma dessas categorias ou se encaixam em mais de uma. Todavia, são essas três que parecem mais adequadas para lidar com os paradoxos que nos interessam aqui.)

4.2.1 Lógicas Não-Contrativas

A abordagem subestrutural mais desenvolvida para os paradoxos atua atacando a regra estrutural da contração. A contração é o princípio que nos diz que, sempre que temos $\Gamma, A, A \vdash B$, então também temos $\Gamma, A \vdash B$; isto é, é o princípio que afirma que podemos utilizar premissas repetidamente enquanto as contamos apenas uma vez. Retornando ao argumento apresentado na seção 2, podemos observar que, em dois momentos, uma suposição é usada duas vezes para alcançar uma conclusão: a suposição 2a é usada duas vezes até chegar a 2d, e a suposição 3a é usada duas vezes até chegar a 3d. Na forma como apresentamos o argumento, não destacamos essa característica, mas é exatamente nesse ponto que uma abordagem não-contrativa tem seu foco.

Os detalhes da resposta dependerão de como os conectivos que representamos como '$\vee$' e '$\wedge$' são interpretados; na ausência da contração, cada um dos conectivos de conjunção e disjunção apresenta versões 'aditivas' e 'multiplicativas', e os proponentes das visões não-contrativas divergem quanto a quais dessas versões eles admitem. A diferença entre conjunção aditiva e multiplicativa é a seguinte: uma conjunção aditiva pode desempenhar o papel de qualquer um de suas conjuntas, enquanto uma conjunção multiplicativa pode desempenhar o papel de ambas as conjuntas simultaneamente. Na presença da contração, a conjunção aditiva é suficiente para a multiplicativa: pode ser usada uma vez para suprir o papel da primeira conjunta e novamente para suprir o papel da segunda. A contração permite que essas duas utilizações contem como uma só. Sem a contração, entretanto, a conjunção aditiva não necessariamente basta para a multiplicativa. (A conjunção multiplicativa basta para a aditiva na presença de uma regra estrutural chamada enfraquecimento, não discutida neste artigo.) A situação para a disjunção é dual: na presença de contração, a disjunção aditiva é suficiente para a multiplicativa, mas fora desse contexto isso não se verifica necessariamente.

O duplo uso apontado acima terá maior relevância se esses conectivos forem interpretados multiplicativamente: se 2d realmente tiver de desempenhar o papel de $\neg Tr(\ulcorner M \urcorner)$ e $Tr(\ulcorner M \urcorner)$ conjuntamente, então de fato utilizará duas cópias de 2a, uma para cada conjunto. Em uma leitura aditiva de 2d e 3d, esse aparente duplo uso não precisa ser problemático, uma vez que 2d apenas necessita desempenhar o papel de um de seus conjuntos. Embora possa ser qualquer um deles, qualquer que seja o conjunto escolhido, uma única utilização de 2a será suficiente. Nessa leitura aditiva, são os princípios PTE e EFQ que entram em questão; por exemplo, com $\wedge$ interpretado aditivamente, são necessárias duas ocorrências da mesma contradição para que se derive uma sentença arbitrária (pois ambos os conjuntos precisam ser utilizados), enquanto a derivação acima produz apenas uma. (A situação para 3d e 3a é semelhante, em ambos os casos.) Não consideramos aqui detalhes adicionais; para uma análise mais aprofundada dessas escolhas e das abordagens não-contrativas em geral, veja Beall e Murzi (2013), Grishin (1982), Petersen (2000), Restall (1994), Ripley (2015), L. Shapiro (2011a, 2015) e Zardini (2011, 2013). (Alguns desses trabalhos focam em paradoxos de natureza conjuntista, e não em paradoxos da verdade, mas muitas das questões são paralelas. Veja também o verbete sobre o paradoxo de Russell[23].)

[23]N.T.: https://plato.stanford.edu/entries/russell-paradox/

4.2.2 Lógicas Não-Transitivas

Outra forma de abordagem subestrutural trabalha atacando diversas regras estruturais associadas à transitividade da consequência. A mais conhecida dessas regras é a regra do corte, que nos permite passar de $\Gamma \vdash B$ e $\Delta, B \vdash C$ para $\Delta, \Gamma \vdash C$. Contudo, também pode ser útil considerar outras propriedades relacionadas à transitividade, como aquela denominada transitividade simples em Weir (2015), que procede de $A \vdash B$ e $B \vdash C$ para $A \vdash C$. (Ou seja, a transitividade simples é o caso especial do corte em que Δ é vazio e Γ é um conjunto unitário.)

Algumas abordagens não-transitivas podem ser compreendidas por meio dos mesmos modelos trivalorados utilizados para K_3 e LP (novamente, remetemos ao verbete sobre lógica multivalorada[24] para maiores detalhes). A diferença reside em como a consequência é definida nesses modelos. Em todos os casos, consequência equivale à ausência de um contramodelo, mas existem diferentes compreensões sobre o que um modelo precisa ser para constituir um contramodelo a um argumento. Dependendo da compreensão de contramodelo adotada, os mesmos modelos trivalorados podem dar origem à lógica paracompleta K_3, à lógica paraconsistente LP, a uma lógica simultaneamente paracompleta e paraconsistente, às vezes denominada S_3 ou $FDRM$, ou — nosso tópico atual — a duas lógicas distintas que incluem contraexemplos à regra do corte e que passaram a ser conhecidas como *não-transitivas*.

Uma abordagem sem a regra do corte é desenvolvida e defendida em Weir (2005, 2015) (e, para a teoria ingênua dos conjuntos, em Weir (1998, 1999), sendo ali denominada 'neoclássica'). Nesta abordagem, o terceiro valor nos modelos é considerado como sendo nem verdadeiro nem falso, e um contramodelo a um argumento de Γ para B deve: ou tornar toda sentença em Γ verdadeira e B não verdadeira, ou tornar B falsa e todas as sentenças em Γ verdadeiras, exceto uma, que deve ser tornada não falsa. A ideia motivadora é que argumentos válidos devem preservar a verdade, e também preservar a falsidade de forma retrospectiva, em certo sentido: se um argumento válido tem todas as suas premissas, exceto uma, verdadeiras, e a conclusão falsa, então a premissa restante deve ser falsa. Isso permite contraexemplos à regra do corte, mas não à transitividade simples, e possibilita a manutenção da consistência. A lógica resultante é mais fraca do que a lógica clássica. Na nossa versão do paradoxo do Mentiroso, a dificuldade situa-se em PTE: a abordagem de

[24]N.T.: https://plato.stanford.edu/entries/logic-manyvalued/

Weir permite contraexemplos ao PTE.

Uma abordagem distinta sem a regra do corte é desenvolvida e explorada em Barrio *et al*. (2015); Cobreros *et al*. (2013, 2015); Fjellstad (2016); e Ripley (2013a, 2015). Nesta abordagem, um contramodelo a um argumento não pode atribuir o terceiro valor a qualquer sentença que ocorra no argumento. Ou seja, um contramodelo a um argumento de Γ para B deve agir exatamente como um contramodelo clássico em relação ao argumento. Se atribuir o terceiro valor a alguma sentença, essa sentença não pode pertencer a Γ nem ser B. Essa concepção permite contraexemplos à regra do corte e, diferentemente da abordagem de Weir, também admite contraexemplos à transitividade simples. Apresenta ainda a característica curiosa de que todo argumento válido na lógica clássica permanece válido. Ou seja, todos os contraexemplos à regra do corte e à transitividade simples envolvem o uso do princípio de captura, liberação ou algum outro comportamento especial do predicado de verdade. Apesar dessa nuance clássica, essas abordagens são também dialeteístas; a afirmação de que a sentença do Mentiroso é simultaneamente verdadeira e não verdadeira revela-se como um teorema. Tal afirmação é forçada a assumir o terceiro valor, razão pela qual não pode haver contra modelo para qualquer argumento que a envolva.

Talvez devido à importância da regra do corte na teoria da demonstração, as abordagens não-transitivas são frequentemente estudadas por meio de sistemas de prova, em vez de modelos. O uso essencial de propriedades de transitividade em derivações paradoxais foi observado em Tennant (1982); uma abordagem dos paradoxos que rejeita tanto a regra do corte quanto à transitividade simples, em um contexto geral, pode ser encontrada em Hallnäs (1991); Hallnäs e Schroeder-Heister (1991); e Schroeder-Heister (2004). Há comentários filosóficos esclarecedores sobre a regra do corte em Schroeder-Heister (1992), que também observa algumas relações entre as abordagens não-contrativas e não-transitivas.

4.2.3 Logicas Não-Reflexivas

Uma terceira possibilidade para uma abordagem subestrutural aos paradoxos consiste em atacar a *reflexividade*, o princípio segundo o qual toda sentença implica a si mesma. Há uma estreita analogia entre reflexividade e transitividade, conforme explicado em Frankowski (2004); Girard *et al*. (1989, p. 28); e Ripley (2012), de modo que essa linha de abordagem acaba por apresentar pontos em comum com a família das abordagens não-transitivas. As abordagens não-reflexivas aos paradoxos têm sido, até o momento, menos exploradas, mas parecem constituir uma direção promissora para pesquisas futuras; para mais desenvolvimentos, ver French (2016) e Meadows (2014). Veja também Malinowski (1990) para trabalhos gerais sobre lógicas não-reflexivas.

4.3 Lógica Clássica

Até aqui, examinamos diversas opções para responder ao paradoxo do Mentiroso mediante a reconsideração da lógica básica. Há, contudo, diversas abordagens que mantêm a lógica clássica inalterada e procuram encontrar outras formas de neutralizar o paradoxo.

Uma característica marcante da maioria dessas abordagens é a disposição de, de algum modo, *restringir* o âmbito de aplicação dos princípios de captura e liberação, a fim de bloquear o raciocínio paradoxal. Tal postura é antitética à visão deflacionista da verdade discutida na seção 4.1, mas é compatível com outra concepção de verdade. Essa concepção alternativa entende que a principal característica da verdade é relatar uma propriedade semântica não trivial das sentenças (por exemplo, a correspondência com um fato no mundo, ou a atribuição de um valor em um modelo). Muitas abordagens que permanecem no âmbito da lógica clássica incorporam a ideia de que uma compreensão adequada desta propriedade permite formas restritas de captura e liberação, e que isso, por sua vez, possibilita bloquear o paradoxo sem qualquer abandono da lógica clássica.

Consideraremos a seguir diversas abordagens importantes ao paradoxo dentro da lógica clássica, a maioria das quais incorpora essa concepção de algum modo.

4.3.1 A Hierarquia de Linguagens de Tarski

Tradicionalmente, o principal caminho para resolver o paradoxo dentro da lógica clássica é a hierarquia de linguagens e metalinguagens proposta por Tarski. Tarski concluiu, a partir do paradoxo, que nenhuma linguagem poderia conter seu próprio predicado de verdade (em sua terminologia, nenhuma linguagem pode ser 'semanticamente fechada').

Em vez disso, Tarski propôs que o predicado de verdade para uma linguagem só poderia ser encontrado em uma metalinguagem expandida. Por exemplo, começa-se com uma linguagem interpretada $\mathcal{L}_0$, que não contém nenhum predicado de verdade. Em seguida, realiza-se um "salto"para uma linguagem expandida $\mathcal{L}_1$, a qual contém um predicado de verdade, mas um predicado que se aplica apenas às sentenças de $\mathcal{L}_0$. Com essa restrição, é relativamente simples definir um predicado de verdade que relate de maneira completamente precisa os valores de verdade de cada sentença de $\mathcal{L}_0$, obedecendo aos princípios de captura e liberação, sem incorrer em qualquer paradoxo. Naturalmente, esse processo não se encerra aí. Caso desejemos descrever a verdade em $\mathcal{L}_1$, será necessário subir para $\mathcal{L}_2$ para obter um predicado de verdade para $\mathcal{L}_1$. E assim sucessivamente. O processo continua indefinidamente. Em cada etapa, uma nova linguagem clássica interpretada é produzida, apta a expressar a verdade para as linguagens de nível inferior. (Para mais detalhes sobre a matemática dessa hierarquia de linguagens, ver Halbach (1997).)

Por que não há paradoxo do Mentiroso nesse tipo de hierarquia de linguagens? Porque a restrição de que nenhum predicado de verdade pode se aplicar a sentenças da própria linguagem é imposta de forma sintática. Qualquer sentença M equivalente a $\neg Tr(\ulcorner M \urcorner)$ não é sintaticamente bem-formada. Assim, não há paradoxo do Mentiroso porque não há sequer uma sentença do Mentiroso. Para uma discussão mais aprofundada sobre as concepções de verdade de Tarski, consulte as entradas relativas à Tarski e às definições de verdade de Tarski.[25]

A abordagem hierárquica de Tarski foi alvo de diversas críticas. Uma delas é que, à luz de casos de autorreferência que ocorrem naturalmente, a sua decisão de tornar as sentenças do Mentiroso sintaticamente malformadas parece excessivamente drástica. Embora o próprio Tarski estivesse mais interessado em resolver o paradoxo do Mentiroso para linguagens formais, sua solução mostra-se inverossímil quando aplicada a muitos usos naturais de 'verdade'. Outro problema relevante foi destacado por Kripke (1975). Conforme observa Kripke, qualquer conjunto de níveis fixado sintaticamente torna extremamente difícil — senão impossível — acomodar várias afirmações não-paradoxais dentro da hierarquia. Por exemplo, se Jc declara que "tudo o que Michael diz é verdadeiro", essa afirmação deve ser feita a partir de um nível da hierarquia superior a tudo o que Michael diz. Contudo, se entre as coisas ditas por Michael consta que "tudo o que Jc diz é verdadeiro", então as afirmações de Michael deveriam estar num nível superior às de Jc. Assim, algumas afirmações de Michael estariam em um nível superior a algumas de Jc, e vice-versa. O que é impossível. Além disso, mesmo nos casos em que é possível atribuir coerentemente um nível a uma enunciação, é difícil explicar o que determina que a referência à verdade envolva exatamente um nível específico e não outro. O que faz com que uma afirmação pertença a um nível da hierarquia em vez de a outro?

Outro desafio enfrentado pela hierarquia de Tarski é explicar por que não podemos simplesmente definir a verdade para toda a hierarquia, quantificando sobre os níveis. Assim,

[25] N.T.: Consulte a nota 5

teríamos um predicado como 'verdadeiro em algum nível'. Se tais predicados forem permitidos, o paradoxo retorna, o que obriga os defensores da hierarquia tarskiana a sustentar que tais predicados não são possíveis. Explicar por que não são possíveis é um problema comum a todas as abordagens hierárquicas. (Para uma discussão mais aprofundada, veja Glanzberg (2015).)

A partir desse tipo de dificuldade, muitos concluíram que a hierarquia de linguagens e metalinguagens proposta por Tarski obtém uma solução para o paradoxo do Mentiroso ao custo de uma restrição excessivamente implausível.

4.3.2 A Construção Fechada de Kripke

À luz das críticas dirigidas à teoria de Tarski, diversas abordagens ao paradoxo do mentiroso buscaram conservar a lógica clássica, mas permitindo certo grau de autoaplicação do predicado de verdade. Sabemos, a partir do raciocínio apresentado na seção 2.3, que algumas restrições sobre captura e liberação serão então necessárias. Um dos objetivos tem sido determinar quais restrições são bem fundamentadas e como implementá-las adequadamente.

Uma maneira de proceder nesse sentido foi sugerida pelo próprio Kripke. Em vez de considerar o aparato de Kripke, brevemente revisado na seção 4.1.1, como parte de uma abordagem lógica não-clássica, pode-se vê-lo como um passo intermediário na construção de uma interpretação clássica de um predicado de verdade autoaplicável Tr.

Recordemos que a construção de Kripke inicia-se com uma linguagem clássica $\mathcal{L}_0$ sem predicado de verdade. Passa-se, então, para uma linguagem expandida $\mathcal{L}_0^+$, contudo, diferentemente de uma metalinguagem tarskiana, essa linguagem contém um predicado de verdade Tr que se aplica a todas as sentenças de $\mathcal{L}_0^+$. Kripke demonstra como construir uma interpretação parcial de Tr, fornecendo uma extensão $\mathcal{E}$ e uma antiextensão $\mathcal{A}$. Pode-se, porém, considerar simplesmente o modelo clássico $\langle \mathcal{M}_0, \mathcal{E} \rangle$, utilizando unicamente a extensão. Esta é a chamada construção 'fechada', pois o hiato entre a extensão e a antiextensão é eliminado, classificando-se como falsas, em um modelo clássico, todas as sentenças situadas nesse intervalo.

Sabemos que tal interpretação não pode satisfazer integralmente as propriedades de captura e liberação (nem a intersubstitutividade completa entre A e $Tr(\ulcorner A \urcorner)$). Todavia, ela assegura uma forma restrita dessas propriedades. No modelo fechado, é válido o seguinte:

$$[Tr(\ulcorner A \urcorner) \vee Tr(\ulcorner \neg A \urcorner)] \rightarrow [Tr(\ulcorner A \urcorner) \leftrightarrow A].$$

Isso nos mostra que captura e liberação (na forma do esquema-T) valem para sentenças que são bem-comportadas, no sentido de satisfazerem $Tr(\ulcorner A \urcorner) \vee Tr(\ulcorner \neg A \urcorner)$.

O que acontece com a sentença do Mentiroso nessa abordagem? Assim como no caso trivalorado, o Mentiroso é interpretado como caindo dentro da lacuna. M não pertence nem a $\mathcal{E}$ nem a $\mathcal{A}$. Assim, M está fora do domínio em que Tr é interpretado como bem-comportado. Como a situação é clássica, e $\ulcorner M \urcorner \notin \mathcal{E}$, sabemos que $\neg Tr(\ulcorner M \urcorner)$ é verdadeira no modelo fechado; da mesma forma, também é verdadeira $\neg Tr(\ulcorner \neg M \urcorner)$.

Para sentenças bem-comportadas, temos a propriedade do ponto fixo segundo a qual A e $Tr(\ulcorner A \urcorner)$ possuem o mesmo valor de verdade, e assim a semântica de $\mathcal{L}_0^+$ e a semântica que ela atribui a Tr correspondem exatamente. Para sentenças patológicas como M, essa correspondência não ocorre e, de fato, não pode ocorrer, sob pena de trivialidade.

Foi observado por Feferman (1984), em um ponto relacionado à construção fechada, que se formos cuidadosos com a negação, podemos dispensar totalmente $\mathcal{A}$ na construção de Kripke. Assim, a construção pode ser realizada sem qualquer apelo implícito à lógica multivalorada. Formas relacionadas de pensar a construção de Kripke são discutidas por McGee (1991).

4.3.3 Determinação Revisitada

Na seção 4.1.3, notamos que abordagens paracompletas ao paradoxo podem ser vulneráveis a "paradoxos de revanche" baseados em alguma ideia de verdade indeterminada ou falta de valor de verdade. Questões relacionadas também surgem no caso clássico. Discutiremos algumas delas a seguir.

4.3.4 Grounding

A construção fechada de Kripke pode ajudar a preencher a ideia de um operador de determinação discutido na seção 4.1.3. Em vez de um operador, ela nos permite definir um predicado $D(\ulcorner A \urcorner)$ por $Tr(\ulcorner A \urcorner) \vee Tr(\ulcorner \neg A \urcorner)$. D representa 'determinadamente' no sentido de ser aplicável a sentenças que possuem um valor de verdade de acordo com Tr, ou seja, 'determinado' pelo modelo produzido pela construção de Kripke. Aplica-se também, como observamos, a todas as sentenças que são bem-comportadas no sentido de obedecerem ao esquema-T (ou às regras de captura e liberação).

Formalmente, as sentenças às quais D se aplica no modelo gerado pela construção de Kripke são aquelas que pertencem a $\mathcal{E}$ ou cujas negações pertencem a $\mathcal{E}$ (equivalentemente, pertencem a $\mathcal{A}$). Kripke denominou esse fenômeno de *grounding*.[26]

Tem sido frequentemente observado que existe também uma noção mais informal de determinação ou grounding, à qual a noção formal expressa por D corresponde aproximadamente (cf. Herzberger 1970). A ideia é que as sentenças determinadas são aquelas que possuem propriedades semânticas bem definidas. Onde não temos tais propriedades semânticas bem definidas, não devemos esperar que o predicado de verdade reporte algo bem-comportado, nem devemos esperar que propriedades como captura e liberação se mantenham. A construção de Kripke constrói $\mathcal{E}$ em estágios, começando com sentenças sem termos semânticos e acrescentando complexidade semântica a cada etapa. Alcança-se $\mathcal{E}$ no limite desse processo, o que nos permite pensar em $\mathcal{E}$ como indicando o limite em que valores semânticos são atribuídos por um processo bem definido. Assim, a noção formal de grounding fornecida por D é por vezes sugerida como refletindo a extensão em que as sentenças possuem propriedades semânticas bem definidas.

A noção de grounding gerou sua própria literatura, com Leitgeb (2005) sendo um impulso chave. Ver também Bonnay e van Vugt (2015), Meadows (2013) e Schindler (2014).

4.3.5 McGee sobre a Verdade e a Verdade Definida

Outra visão que faz uso de uma forma de determinabilidade é defendida por McGee (1991). A teoria de McGee, como muitas das que examinamos aqui, é rica em complexidade, de modo que não podemos fazer-lhe justiça em sua totalidade. A teoria possui muitos componentes, incluindo abordagens matematicamente sofisticadas para a verdade relacionadas às ideias kripkeanas que discutimos, dentro de um contexto que preserva a lógica clássica.

McGee baseia-se em duas noções: verdade e verdade definida. A verdade definida é uma forma da ideia que anteriormente caracterizamos como determinabilidade. Contudo,

[26]Uma nota técnica: Kripke definiu o conceito de ser fundamentado [*grounded*] como pertencer ao ponto fixo mínimo. Não discutimos aqui outros pontos fixos, mas há muitos, e uma sentença pode pertencer a alguns deles sem, contudo, pertencer ao ponto fixo mínimo.

McGee descreve essa ideia utilizando técnicas lógicas bastante sofisticadas. Mencionaremos essas técnicas brevemente, para aqueles familiarizados com o pano de fundo técnico. Formalmente, para McGee, a verdade definida é identificada com a demonstrabilidade em uma linguagem parcialmente interpretada, utilizando uma extensão da lógica clássica conhecida como lógica-$\mathcal{A}$, que incorpora fatos acerca da interpretação parcial. Trata-se, portanto, de algo distinto da noção de grounding que acabamos de discutir. McGee trata 'definitivamente' como um *predicado*, em paridade com o predicado de verdade, e não como um operador sobre sentenças, como ocorre em alguns outros desenvolvimentos. Com a noção adequada de verdade definida, McGee demonstra que uma linguagem parcialmente interpretada contendo seu próprio predicado de verdade pode satisfazer formas restritas de captura e liberação, expressas em termos de verdade definida. Onde $\mathcal{D}ef$ é o predicado de definitividade, McGee mostra como vincular verdade e verdade definida, validando a seguinte relação:

$$Def(\ulcorner A \urcorner) \text{ sse } Def(\ulcorner Tr(\ulcorner A \urcorner) \urcorner)$$
$$Def(\ulcorner \neg A \urcorner) \text{ sse } Def(\ulcorner \neg Tr(\ulcorner A \urcorner) \urcorner)$$

De fato, McGee demonstra que essas condições podem ser satisfeitas dentro de uma teoria que incorpora tanto a verdade quanto a verdade definida, na qual a verdade respeita formas apropriadas de captura e liberação, e onde também uma declaração formal da bivalência para a verdade resulta ser definitivamente verdadeira. McGee, assim, apresenta uma teoria que permite uma aplicação fortemente autorreferencial da verdade e da verdade definida, dentro de um quadro clássico.

Embora a verdade possa satisfazer a propriedade formal da bivalência, é crucial para a abordagem de McGee que a verdade definida seja uma noção aberta, que pode ser fortalecida (formalmente, por meio do fortalecimento de uma linguagem parcialmente interpretada). Dessa forma, a verdade definida satisfaz formas mais fracas de captura e liberação do que a própria verdade. (Algumas instâncias de $Def(\ulcorner A \urcorner) \to A$ deixam de ser definitivamente verdadeiras, segundo McGee.) Além disso, McGee sugere que esse comportamento da verdade e da verdade definida torna a verdade um predicado vago. Permanece em debate se a teoria de McGee consegue evitar os tipos de problemas de revanche que afetam outras abordagens kripkeanas.

4.3.6 Outras Abordagens Clássicas

Já examinamos alguns representantes importantes das abordagens que buscam resolver o paradoxo do Mentiroso dentro da lógica clássica. Existem diversas outras, muitas das quais envolvem matemática complexa. Faremos uma pausa para mencionar algumas das mais relevantes, embora, devido à complexidade matemática envolvida, nos limitemos a indicá-las de forma geral.

4.3.7 Teorias Axiomáticas da Verdade

Existe uma vertente importante de trabalho na teoria da prova que busca desenvolver teorias axiomáticas da verdade autoaplicáveis dentro da lógica clássica, incluindo os trabalhos de Cantini (1996), Feferman (1984, 1991), Friedman e Sheard (1987), Halbach (2011) e Horsten (2011). A ideia é encontrar maneiras de expressar regras como captura e liberação que preservem a consistência. Entre as opções, incluem-se maior cuidado na formulação das regras de inferência da teoria da prova e na formulação de regras restritas. As ideias principais são discutidas no verbete sobre 'teorias axiomáticas da verdade'[27], às quais não exploraremos os detalhes.

4.3.8 Verdade e Definições Indutivas

O trabalho de Kripke sobre a verdade foi desenvolvido em conjunto com algumas ideias importantes sobre definições indutivas (como podemos ver, por exemplo, nas últimas partes de Kripke 1975). Essas conexões são exploradas de forma mais aprofundada nos trabalhos de Burgess (1986) e McGee (1991). Cabe também mencionar o trabalho de Aczel (1980), que combina ideias sobre definições indutivas e o cálculo lambda.

4.4 Abordagens Contextualistas

Outra família de propostas para a solução do paradoxo do Mentiroso são as *soluções contextualistas*. Estas também utilizam a lógica clássica, mas fundamentam suas soluções principalmente em algumas ideias oriundas da filosofia da linguagem. Elas interpretam a principal lição do paradoxo do Mentiroso como sendo o fato de que predicados de verdade

[27] N.T.: https://plato.stanford.edu/entries/truth-axiomatic/

exibem uma forma de *dependência do contexto*, mesmo em fragmentos de linguagem que, de outra forma, não apresentariam tal dependência. Tais soluções procuram explicar como isso pode ocorrer e baseiam nessa explicação a resolução dos problemas suscitados pelo paradoxo do Mentiroso.

As teorias contextualistas compartilham com as diversas abordagens que já examinamos, a ideia de que há algo de indeterminado ou semanticamente malformado em nossa sentença do Mentiroso M. No entanto, as visões contextualistas atribuem um papel especial às questões de 'revanche' e à falta de poder expressivo.

4.4.1 Instabilidade e Revanche

Uma maneira de entender por que o predicado de verdade não se comporta adequadamente diante da sentença do Mentiroso é considerar que, na realidade, não há um portador de verdade bem definido fornecido por essa sentença. Para ilustrar esta ideia (como discutido por C. Parsons (1974)), suponha que os portadores de verdade sejam proposições expressas por sentenças em contextos específicos, e que a sentença do Mentiroso falhe em expressar uma proposição. Este é o início de uma explicação de como o Mentiroso acaba por ser não fundado [*ungrounded*] ou, em certo sentido, indeterminado. Pelo menos, não deveríamos esperar que Tr se comporte adequadamente em casos em que as sentenças falham em expressar proposições.

Contudo, trata-se de uma proposta instável. Podemos raciocinar que, se a sentença do Mentiroso falha em expressar uma proposição, ela falha, portanto, em expressar uma proposição verdadeira. À maneira de um paradoxo de revanche, se a nossa sentença do Mentiroso afirmasse originalmente "esta sentença não expressa uma proposição verdadeira", então recuperaríamos a sentença paradoxal do Mentiroso. E, ao fazê-lo, teríamos demonstrado que tal sentença diz algo verdadeiro, e assim expressa uma proposição verdadeira. Portanto, partindo da suposição de que a sentença do Mentiroso é indeterminada ou carece de status semântico, raciocinamos que ela deve possuir, sim, um status semântico apropriado e, de fato, expressar algo verdadeiro. Assim, retornamos ao paradoxo.

Os contextualistas não veem isso como um novo paradoxo de 'revanche', mas sim como o problema básico colocado pelo Mentiroso. Em primeiro lugar, em um contexto onde as sentenças são dependentes do contexto, a formulação natural de uma assertiva verdadeira é sempre em termos de expressar uma proposição verdadeira, ou alguma aplicação semanticamente cuidadosa do predicado de verdade. De modo mais relevante, para o contextualista, a principal questão subjacente ao Mentiroso está incorporada no raciocínio aqui apresentado. Ele envolve dois passos fundamentais. Primeiro, atribui-se ao Mentiroso um status semanticamente defeituoso — isto é, a falha em expressar uma proposição ou ser, de algum modo, indeterminado. Segundo, conclui-se, a partir do primeiro passo, que o Mentiroso deve ser verdadeiro — e, portanto, não indeterminado ou falho em expressar uma proposição. Ambos os passos parecem ser frutos de raciocínios corretos, e, portanto, as conclusões alcançadas em ambos deveriam ser verdadeiras. O principal desafio do paradoxo do Mentiroso, segundo os contextualistas, é explicar como isso pode ocorrer e como o segundo passo pode ser não-paradoxal. (Esse tipo de raciocínio é explorado por Glanzberg (2004c) e C. Parsons (1974). Para uma discussão crítica, ver Gauker (2006).)

Assim, os contextualistas procuram explicar como a sentença do Mentiroso pode possuir um status semântico instável, alternando entre defeituoso e não defeituoso no decorrer desse tipo de inferência. Eles o fazem apelando para o papel do *contexto* na fixação do status semântico das sentenças. Sentenças podem possuir diferentes status semânticos em diferentes contextos. Portanto, para os contextualistas, deve haver algum efeito não trivial do contexto envolvido na sentença do Mentiroso e, de maneira mais geral, na predicação de verdade.

4.4.2 Parâmetros Contextuais nos Predicados de Verdade

Uma abordagem contextualista proeminente, defendida por Burge (1979) e desenvolvida por Koons (1992) e Simmons (1993), parte da ideia de que a própria hierarquia tarskiana oferece uma maneira de ver o predicado de verdade como dependente do contexto. A hierarquia de Tarski postula uma hierarquia de predicados de verdade Tr_i. E se i não fosse meramente um marcador de nível em uma hierarquia, mas um autêntico parâmetro contextual? Se assim for, então a sentença do Mentiroso é de fato dependente do contexto: ela tem a forma $\neg Tr_i(\ulcorner M \urcorner)$, onde i é fixado pelo contexto. O contexto, portanto, determina o nível do predicado de verdade.

Essa ideia pode ser vista como um aprimoramento da abordagem tarskiana original em vários aspectos. Primeiro, uma vez que possuímos um parâmetro contextual, desaparece a necessidade de insistir que sentenças do Mentiroso jamais sejam bem-formadas. Assim, podemos considerar cada Tr_i como incluindo algum alcance limitado de aplicabilidade a sentenças de sua própria linguagem. Utilizando técnicas kripkeanas como a construção fechada que revisamos acima, é possível construir predicados como Tr_i que tenham tanto grau de autoaplicabilidade quanto os do próprio Kripke. (Burge 1979 e o pós-escrito de C. Parsons 1974 consideram brevemente como técnicas kripkeanas poderiam ser aplicadas nesse contexto. Embora opere em um cenário bastante diferente, as ideias de Gaifman (1988, 1992) podem ser interpretadas como mostrando como modos ainda mais sutis de interpretar um predicado de verdade dependente de contexto podem ser desenvolvidos.)

Com os devidos cuidados, outros problemas associados à hierarquia tarskiana também podem ser evitados. Burge propõe que o parâmetro i em Tr_i seja determinado por um processo pragmático de natureza griceana. Em essência, os falantes implicam que i deve ser ajustado para um nível no qual o discurso em que eles estejam possa ser interpretado de maneira coerente (com uma extensão máxima coerente para Tr_i). Assim, a verdade de fato encontra seu próprio nível, e, portanto, a objeção de Kripke sobre como fixar níveis para sentenças não-paradoxais pode ser contornada.

Essa abordagem dá substância à ideia de que a sentença do Mentiroso é dependente do contexto. Qualquer sentença que contenha Tr_i será dependente do contexto, herdando um parâmetro contextual ao longo do caminho. Isso oferece uma maneira de dar sentido aos argumentos sobre a instabilidade do status semântico de M que motivaram o contextualismo. Em um contexto inicial, fixamos um nível i. Este é o nível em que M é interpretado. Chamemos essa interpretação de M_i. M_i afirma $\neg Tr_i(\ulcorner M_i \urcorner)$. Pelo raciocínio usual do Mentiroso, mostramos que M_i deve carecer de status semântico determinado — ou falhar em expressar uma proposição. Como discutimos, em seguida raciocinamos que M deve ser verdadeiro. Segundo a visão contextualista aqui apresentada, isso corresponde à afirmação de que M_i é verdadeiro de acordo com algum outro contexto, no qual um predicado de verdade mais amplo está em jogo. Isso equivale a ser verdadeiro em algum nível superior da hierarquia. Podemos concluir, por exemplo, que a sentença do Mentiroso, tal como foi usada no nível i, é verdadeira segundo um nível mais alto $k > i$. Assim, $Tr_k(\ulcorner M_i \urcorner)$, onde $k > i$.

Essa forma de contextualismo sustenta que, uma vez que reconhecemos o comportamento dependente do contexto de Tr_i, podemos compreender adequadamente a instabilidade de M. Isso pode ser visto como uma melhoria tanto em relação à visão Tarskiana quanto como uma incorporação de algumas das técnicas de lógica clássica que revisamos na seção 4.2. Dependendo de como a visão de Burge for desenvolvida tecnicamente, ela terá ou a plena validade das regras de captura e liberação em cada nível, ou a captura e liberação com as mesmas restrições da construção fechada de Kripke.

A visão que postula parâmetros contextuais no predicado de verdade enfrenta diversas questões. Por exemplo, é legítimo perguntar por que deveríamos considerar que o predicado de verdade possui realmente um parâmetro contextual, especialmente se estivermos nos referindo a um predicado de verdade como o que utilizamos na linguagem natural. Meramente observar que tal parâmetro evitaria o paradoxo não demonstra que ele esteja de fato presente na linguagem natural. Ademais, permanece controverso se é aceitável conceber a verdade como ocorrendo em níveis, sejam eles baseados em contexto ou não. (Nem todos os que defendem parâmetros contextuais para o predicado de verdade concordam quanto ao papel da hierarquia. Em particular, Simmons (1993) defende uma visão que ele denomina ”teoria da singularidade”, a qual, segundo ele, evita estruturas hierárquicas explícitas.) Por fim, o recurso burgeano aos mecanismos griceanos para a fixação dos níveis de verdade também foi contestado. (Por exemplo, Gaifman (1992) questiona se o processo griceano desempenha algum papel substancial na teoria de Burge.)

As abordagens contextualistas apresentam diversas variantes, cada uma fazendo uso de instrumentos conceituais ligeiramente distintos. Nas teorias contextualistas, a escolha entre essas variantes frequentemente depende de questões relativas tanto à filosofia da linguagem quanto à lógica. Já mencionamos anteriormente uma maneira diferente de desenvolver ideias contextualistas proposta por Gaifman (1988, 1992). Faremos agora uma breve revisão de algumas outras alternativas.

4.4.3 Efeitos Contextuais em Domínios dos Quantificadores

Outra abordagem contextualista, oriunda do trabalho de C. Parsons (1974), busca fundamentar a dependência contextual da sentença do Mentiroso — e, em última instância, do predicado de verdade — a partir de componentes mais básicos. O ponto central é perceber que a dependência contextual da sentença do Mentiroso deriva da dependência contextual dos domínios dos quantificadores.

A quantificação entra em cena quando refletimos sobre como justificar a predicação de verdade em sentenças que exibem dependência do contexto. Em tal ambiente, não faz sentido atribuir verdade diretamente a sentenças. Nem todas as sentenças possuirão o tipo adequado de propriedades semânticas determinadas para serem portadoras de valor de verdade; ou, como temos colocado, nem todas as sentenças expressarão proposições. Assim, afirmar que uma sentença S é verdadeira em um contexto c é afirmar que há uma proposição p expressa por S em c, e que essa proposição p é verdadeira.

A atual proposta contextualista parte da observação de que os quantificadores na linguagem natural tipicamente possuem domínios de quantificação dependentes do contexto. Quando dizemos “Todos estão aqui”, não queremos dizer todos no mundo, mas sim todos dentro de algum subdomínio fornecido contextualmente. Segundo esta perspectiva contextualista, a dependência de contexto entra no caso do Mentiroso através dos efeitos contextuais sobre o domínio do quantificador proposicional $\exists p$.

Em particular, esse domínio deve se expandir no decorrer do raciocínio sobre o status semântico do Mentiroso. No contexto inicial, $\exists p$ deve abranger um domínio suficientemente pequeno para que não exista uma proposição que M possa expressar. No contexto subsequente, o domínio se expande para permitir que M expresse alguma proposição verdadeira. Propostas sobre como essa expansão ocorre, e sobre como modelar o predicado de verdade e a relação de expressão de uma proposição na presença do Mentiroso, foram exploradas por Glanzberg (2001, 2004a), com base no trabalho de C. Parsons (1974). Os defensores dessa abordagem argumentam que ela localiza melhor o foco da dependência de contexto do que os parâmetros nos predicados de verdade.

4.4.4 Teoria da Situação

Outra variante da estratégia contextualista para resolver o paradoxo do Mentiroso, desenvolvida por Barwise e Etchemendy (1987) e por Groeneveld (1994), baseia-se na *teoria das situações*, em vez de domínios de quantificação, para fornecer o ponto da dependência de contexto. A teoria das situações é um ramo altamente desenvolvido da filosofia da linguagem, de modo que, mais uma vez, esboçaremos apenas de maneira bastante geral como essa perspectiva funciona.

Uma situação é um estado parcial no qual o mundo pode se encontrar: algo como a sendo F. As situações são classificadas pelo que se denominam ‘tipos de situação’. Uma proposição envolve classificar situação como pertencente a um tipo de situação. Assim, uma proposição $[s; (\sigma)]$ nos informa que a situação s é do tipo σ. A situação s aqui desempenha diversos papéis, incluindo o de fornecer um contexto.

No que diz respeito ao paradoxo do Mentiroso, Barwise e Etchemendy interpretam as proposições do Mentiroso como tendo a forma $fs = [s; (Tr, fs; 0)]$, relativa a uma situação inicial s. Esta é uma proposição fs que diz de si mesma que sua falsidade é um fato que ocorre em s. (Na notação de Barwise e Etchemendy, o ’0’ indica falsidade, de modo que o tipo de situação é que o estado de coisas correspondente à falsidade da proposição se verifica. A proposição afirma que esse fato ocorre em s.). Há um sentido em que esta proposição não pode ser expressa. Em particular, o estado de coisas $(Tr, fs; 0)$ não pode pertencer à situação s. (Barwise e Etchemendy, na verdade, afirmam que a proposição é expressável, mas abrem mão do que denominam o fechamento-F de s. Entretanto, há uma observação central comum entre essas perspectivas, e os detalhes não são essenciais para nossos propósitos aqui.) Forma-se então uma situação distinta $s' = s \cup [(Tr, fs; 0)]$, e a proposição $[s'; (Tr, fs; 0)]$ relativa a esta nova situação — este novo 'contexto' — é verdadeira.

Esta concepção claramente apresenta diversas semelhanças com a abordagem baseada na restrição dos domínios de quantificação. Em especial, ambas as estratégias buscam demonstrar como o domínio dos conteúdos expressáveis em determinados contextos pode expandir-se, a fim de dar conta da instabilidade associada à sentença do Mentiroso. Para uma análise das relações entre as abordagens baseadas na teoria das situações e nos domínios de quantificação, ver Glanzberg (2004a). Barwise e Etchemendy também discutem, em seu trabalho de 1987 (capítulo 11), as relações entre sua abordagem fundamentada em situações e abordagens mais tradicionais. Para uma comparação pormenorizada entre a estrutura teórica de Barwise e Etchemendy e uma estrutura burgeana de predicados de verdade indexados, consulte Koons (1992).

4.4.5 Questões para o Contextualismo

Constitui um desafio central para os contextualistas fornecer um relato completo e bem fundamentado da origem e da natureza da mudança de contexto envolvida no paradoxo do Mentiroso — embora, naturalmente, muitos defensores do contextualismo considerem já ter superado tal desafio. Em favor da abordagem contextualista, destaca-se o fato de que ela trata o fenômeno da revanche como o problema fundamental, mostrando-se, assim, amplamente imune aos tipos de problemas de revanche que afetam outras estratégias que examinamos. Todavia, é possível que haja ainda outra forma de revanche a ser aplicada. Para manter a consistência, os contextualistas devem impor restrições a quantificadores como 'todos os contextos'. Para tanto, parece necessário negar a existência de quantificadores absolutamente irrestritos. Glanzberg (2004b, 2006) defende que essa é, de fato, a conclusão correta, embora tal posição seja altamente controversa. Para uma visão geral das discussões a respeito, ver os artigos reunidos em Rayo e Uzquiano (2006).

4.5 A Teoria da Revisão

Outra abordagem ao paradoxo do Mentiroso, defendida por Gupta (1982), Herzberger (1982), Gupta e Belnap (1993), entre outros, é a *teoria da revisão da verdade*. Essa abordagem partilha certas características com as perspectivas que examinamos na seção 4.3, na medida em que toma a lógica clássica como pressuposto. Contudo, cremos que ela também possui afinidades com as concepções discutidas na seção 4.4, ao repensar aspectos fundamentais da semântica. Todavia, trata-se de uma proposta distinta. Esboçaremos a seguir alguns dos fundamentos desta teoria. Para uma discussão mais aprofundada sobre as bases da teoria da revisão, bem como sobre suas relações com o contextualismo, ver L. Shapiro (2006). Para mais detalhes e referências adicionais, consultar a entrada dedicada à teoria da revisão da verdade.[28]

A teoria da revisão da verdade parte da ideia de que podemos tomar o esquema-T em sua apresentação inicial. De fato, Gupta e Belnap (1993) retomam uma sugestão de Tarski (1944), segundo a qual as instâncias do esquema-T podem ser vistas como definições parciais de verdade; presumivelmente, todas as instâncias reunidas, para a linguagem apropriada ou para uma família de linguagens, constituiriam uma definição completa. Ao mesmo tempo, a teoria da revisão mantém-se fiel à lógica clássica. Assim, como já sabemos, o paradoxo do Mentiroso surge em qualquer linguagem que disponha de recursos expressivos suficientes para gerar sentenças do tipo do Mentiroso.

Em resposta, a teoria da revisão propõe uma maneira distinta de abordar as propriedades semânticas do predicado de verdade. Em consonância com nossas práticas aqui, podemos começar com um modelo clássico $\mathcal{M}_0$ para uma linguagem $\mathcal{L}_0$ sem um predicado de verdade, e considerar o que ocorre ao acrescentarmos um predicado de verdade Tr para formar a linguagem estendida $\mathcal{L}_0^+$. Esta linguagem possui um predicado de verdade plenamente autoaplicável e, assim, é capaz de gerar a sentença do Mentiroso M.

Para construir um modelo clássico para $\mathcal{L}_0^+$, precisamos de uma extensão para Tr. Escolhemos um conjunto: chamemo-lo H, para a hipótese sobre qual poderia ser a extensão de Tr. H pode ser $\emptyset$, pode ser todo o domínio de $\mathcal{M}_0$, ou pode ser qualquer outro conjunto. Não é necessário que seja uma boa aproximação das propriedades semânticas de Tr.

Ainda que não seja, $\langle \mathcal{M}_0, H \rangle$ fornece um modelo clássico, no qual podemos interpretar $\mathcal{L}_0^+$. Com isso, podemos aplicar o esquema-T relativamente à nossa hipótese H e observar

[28] N.T.: https://plato.stanford.edu/entries/truth-revision/

o que obtemos. Mais precisamente, podemos deixar que $\tau(H) = [\ulcorner A \urcorner \mid A$ seja verdadeira em $\langle \mathcal{M}_0, H \rangle]$. $\tau(H)$ é, em geral, uma hipótese melhor sobre o que é verdadeiro em nossa linguagem do que H poderia ter sido. Pelo menos, é evidente que, se H fez suposições errôneas sobre a verdade de sentenças do fragmento sem verdade $\mathcal{L}_0$, elas são corrigidas em $\tau(H)$, que contém todas as sentenças de $\mathcal{L}_0$ verdadeiras em $\mathcal{M}_0$. Assim, $\langle \mathcal{M}_0, \tau(H) \rangle$ é, em geral, um modelo melhor de $\mathcal{L}_0^+$ do que $\langle \mathcal{M}_0, H \rangle$.

Melhor em muitos aspectos. Contudo, quando consideramos sentenças paradoxais como M, observamos algo distinto. Como hipótese inicial, consideremos $H = \emptyset$. Vejamos o que acontece com a verdade de M à medida que aplicamos τ:

n	valor de verdade de $\mathcal{L}$ em $\langle \mathcal{M}_0, \tau^n(\emptyset) \rangle$
0	verdadeiro
1	falso
2	verdadeiro
3	falso
4	verdadeiro
$\vdots$	$\vdots$

A sentença do Mentiroso jamais se estabiliza nesse processo. Ocorre uma alternância de valores de verdade que persiste indefinidamente. Isto demonstra, segundo a teoria da revisão, que a verdade é um conceito circular. Como tal, ela não possui uma extensão no sentido ordinário. Em vez disso, a verdade é caracterizada por uma regra de revisão de extensões, a qual nunca alcança um ponto de estabilização.

Na terminologia da teoria da revisão, τ é uma *regra de revisão*. Ela nos conduz de uma hipótese inicial acerca da interpretação de Tr a uma nova hipótese. As sequências de valores geradas por tais regras de revisão, a partir de uma hipótese inicial, são denominadas *sequências de revisão*. Deixamos para uma exposição mais detalhada a importante questão sobre a maneira correta de definir sequências de revisão. *transfinitas* (veja o verbete sobre teorias da revisão da verdade[29]).

A propriedade característica das sentenças paradoxais, como a sentença do Mentiroso, é sua instabilidade em sequências de revisão: não existe ponto algum na sequência em que elas atinjam um valor de verdade estável. A teoria da revisão, a partir disto, classifica as

[29]N.T.: https://plato.stanford.edu/entries/truth-revision/

sentenças em estavelmente verdadeiras, estavelmente falsas e instáveis. A teoria desenvolve ainda noções de consequência baseadas nessas categorias e em conceitos correlatos. Para uma apresentação mais ampla desta teoria rica e complexa, recomenda-se a leitura do verbete sobre teorias da revisão da verdade.[30]

4.6 Visões de Inconsistência

Conforme observamos na seção 2.3.3, o paradoxo do Mentiroso, na presença da captura e liberação irrestritas e da lógica clássica, conduz à contradição. Enquanto estivermos no âmbito da lógica clássica, a qual admite EFQ, tal contradição resulta em trivialidade. A maioria das soluções que analisamos (com exceção da teoria da revisão) tenta evitar este resultado de alguma maneira, seja restringindo os princípios de captura e liberação, seja abandonando a lógica clássica. Todavia, existe outra proposta que, embora menos frequentemente defendida, sustenta que o paradoxo do Mentiroso revela que as línguas que falamos — línguas que contêm seus próprios predicados de verdade — são *inconsistentes*.

Esta, porém, não é uma posição de fácil formulação. Embora o próprio Tarski pareça ter sugerido algo semelhante (especificamente no que se refere às línguas naturais), Herzberger (1967) argumentou que é impossível haver uma linguagem inconsistente.

Em contraste, Eklund (2002) leva a sério a ideia de que nossas intuições semânticas — expressas, por exemplo, através da aceitação irrestrita dos princípios de captura e liberação, são, de fato, inconsistentes. Eklund reconhece que essa posição não faria sentido caso tais intuições tivessem sua origem apenas em nossa apreensão das condições de verdade das sentenças. No entanto, ele propõe um quadro alternativo de competência semântica que torna essa visão plausível, fortemente relacionado às concepções de significado baseadas no papel conceitual. Segundo Eklund, deveríamos conceber a competência semântica em termos de um conjunto de princípios que os falantes estão dispostos a aceitar em virtude de seu conhecimento de uma língua. Tais princípios podem ser internamente inconsistentes. Ainda assim, eles determinam valores semânticos: estes serão aquilo que mais se aproxima de satisfazer os princípios — isto é, aquilo que os torna maximalmente corretos — mesmo que nenhum valor semântico possa satisfazê-los completamente devido à inconsistência subjacente.

Eklund endossa uma ideia anteriormente sugerida por Chihara (1979). O principal ob-

[30] N.T.: https://plato.stanford.edu/entries/truth-revision/

jetivo de Chihara é oferecer aquilo que denomina um *diagnóstico* do paradoxo, isto é, uma explicação de por que o paradoxo surge e por que ele nos parece tão convincente. No entanto, ao longo de sua argumentação, Chihara também sugere que a origem do paradoxo reside em nossa aceitação do esquema-T (por convenção, segundo ele), apesar de sua inconsistência.

Uma posição relacionada, embora distinta, é defendida por Patterson (2007, 2009). Patterson argumenta que a competência linguística coloca o indivíduo em um estado cognitivo que se relaciona a uma teoria inconsistente — uma teoria que inclui o esquema-T irrestrito e que é regida pela lógica clássica. Ele prossegue explorando como tal estado cognitivo poderia permitir a comunicação bem-sucedida, ainda que nos vincule a uma teoria falsa.

Uma teoria de inconsistência de natureza diferente é defendida por Scharp (2013). Scharp sustenta que o conceito de verdade é inconsistente, de modo análogo ao conceito de massa na física pré-relativística. Sendo assim, o conceito de verdade seria inadequado para a teorização rigorosa. O que se faz necessário, segundo Scharp, é substituir o conceito inconsistente de verdade por uma família de conceitos consistentes, que possam desempenhar melhor as funções teóricas desejadas. Scharp desenvolve precisamente essa família de conceitos e oferece uma teoria correspondente.

5. Considerações Finais

Há ainda muito a ser dito sobre o paradoxo do Mentiroso além do que foi tratado aqui: existem outras abordagens possíveis às variantes do Mentiroso que abordamos, e outros paradoxos relacionados, como os de denotação e propriedades, entre outros. Existem, igualmente, mais resultados técnicos relevantes, bem como implicações e aplicações filosóficas de grande importância. Nosso objetivo, entretanto, se pautou mais por sugerir caminhos do que esgotar o tema, e esperamos ter oferecido ao leitor uma indicação sobre o que é o paradoxo do Mentiroso e sobre quais podem ser suas consequências.

Referências Bibliográficas

Theodore Achourioti, Henri Galinon, José Martinez Fernández, and Kentaro Fujimoto, editors. *Unifying the Philosophy of Truth*. Springer, Berlin, 2015.

Peter Aczel. Frege structures and the notions of proposition, truth and set. In Jon Barwise,

H. J. Keisler, and Kenneth Kunen, editors, *The Kleene Symposium*, pages 31–59. North-Holland, Amsterdam, 1980.

Alan Ross Anderson. St. paul's epistle to titus. In Martin (1970), pages 1–11.

F. G. Asenjo. A calculus of antinomies. *Notre Dame Journal of Formal Logic*, 7(1):103–105, 1966. doi: 10.1305/ndjfl/1093958482.

Eduardo Barrio. The yablo paradox and circularity. *Análisis Filosófico*, 32(1):7–20, 2012.

Eduardo Barrio, Lucas Rosenblatt, and Diego Tajer. The logics of strict-tolerant logic. *Journal of Philosophical Logic*, 44(5):551–571, 2015. doi: 10.1007/s10992-014-9342-6.

Jon Barwise and John Etchemendy. *The Liar*. Oxford University Press, Oxford, 1987.

Jc Beall. Is yablo's paradox non-circular? *Analysis*, 61(3):176–187, 2001. doi: 10.1111/1467-8284.00292.

Jc Beall. Transparent disquotationalism. In Beall and Armour-Garb (2005), pages 7–22.

Jc Beall, editor. *Revenge of the Liar*. Oxford University Press, Oxford, 2008.

Jc Beall. *Spandrels of Truth*. Oxford University Press, Oxford, 2009.

Jc Beall. Adding to relevant restricted quantification. *Australasian Journal of Logic*, 10:36–44, 2011. Available online.

Jc Beall. End of inclosure. *Mind*, 123(491):829–849, 2014. doi: 10.1093/mind/fzu075.

Jc Beall. Free of detachment: logic, rationality, and gluts. *Noûs*, 49:410–423, 2015. doi: 10.1111/nous.12029.

Jc Beall and Bradley Armour-Garb, editors. *Deflationism and Paradox*. Oxford University Press, Oxford, 2005.

Jc Beall and Michael Glanzberg. Where the paths meet: Remarks on truth and paradox. In Peter A. French and Howard K. Wettstein, editors, *Midwest Studies in Philosophy Volume XXXII: Truth and its Deformities*. Wiley-Blackwell, Boston, 2008.

Jc Beall and Julien Murzi. Two flavors of curry's paradox. *Journal of Philosophy*, 110(3):143–165, 2013. doi: 10.5840/jphil2013110336.

Jc Beall, Ross T. Brady, Allen P. Hazen, Graham Priest, and Greg Restall. Relevant restricted quantification. *Journal of Philosophical Logic*, 35(6):587–598, 2006. doi: 10.1007/s10992-005-9008-5.

Denis Bonnay and Floris Tijmen van Vugt. Groundedness, truth, and dependence. In Achourioti et al. (2015), pages 355–368. doi: 10.1007/978-94-017-9673-6_18.

Ross T. Brady. The non-triviality of dialectical set theory. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic: Essays on the Inconsistent*, pages 437–

470. Philosophia Verlag, Munich, 1989.
Tyler Burge. Semantical paradox. *Journal of Philosophy*, 76(4):169–198, 1979. doi: 10.2307/2025724. Reprinted in Martin 1984: 83–117.
John P. Burgess. The truth is never simple. *Journal of Symbolic Logic*, 51(3):663–681, 1986. doi: 10.2307/2274021.
Andrea Cantini. *Logical Frameworks for Truth and Abstraction: An Axiomatic Study*. Elsevier, Amsterdam, 1996.
Charles Chihara. The semantic paradoxes: A diagnostic investigation. *Philosophical Review*, 88(4):590–618, 1979. doi: 10.2307/2184846.
Pablo Cobreros, Paul Egré, David Ripley, and Robert van Rooij. Reaching transparent truth. *Mind*, 122(488):841–866, 2013. doi: 10.1093/mind/fzt110.
Pablo Cobreros, Paul Egré, David Ripley, and Robert van Rooij. Vagueness, truth, and permissive consequence. In Achourioti et al. (2015), pages 409–430. doi: 10.1007/978-94-017-9673-6_21.
Roy Cook. There are non-circular paradoxes (but yablo's isn't one of them). *The Monist*, 89(1):118–149, 2006. doi: 10.5840/monist200689137.
Roy Cook. *The Yablo Paradox*. Oxford University Press, Oxford, 2014.
Bradley H. Dowden. Accepting inconsistencies from the paradoxes. *Journal of Philosophical Logic*, 13:125–130, 1984. doi: 10.1007/BF00453017.
Michael Dummett. Truth. *Proceedings of the Aristotelian Society*, 59(1):141–162, 1959. doi: 10.1093/aristotelian/59.1.141. Reprinted in Dummett 1978, Truth and Other Enigmas, Cambridge: Harvard University Press, 1–24.
Matti Eklund. Inconsistent languages. *Philosophy and Phenomenological Research*, 64(2):251–275, 2002. doi: 10.1111/j.1933-1592.2002.tb00001.x.
Solomon Feferman. Toward useful type-free theories, i. *Journal of Symbolic Logic*, 49(1):75–111, 1984. doi: 10.2307/2274093. Reprinted in Martin 1984: 237–287.
Solomon Feferman. Reflecting on incompleteness. *Journal of Symbolic Logic*, 56(1):1–49, 1991. doi: 10.2307/2274902.
Hartry Field. Deflationist views of meaning and content. *Mind*, 103(411):249–285, 1994. doi: 10.1093/mind/103.411.249.
Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008.
Hartry Field. Naive truth and restricted quantification: Saving truth a whole lot better. *Review of Symbolic Logic*, 7(1):147–191, 2014. doi: 10.1017/S1755020313000312.

Andreas Fjellstad. Naive modus ponens and failures of transitivity. *Journal of Philosophical Logic*, 45(1):65–72, 2016. doi: 10.1007/s10992-015-9351-0.

Szymon Frankowski. Formalization of a plausible inference. *Bulletin of the Section of Logic*, 33(1):41–52, 2004.

Rohan French. Structural reflexivity and the paradoxes of self-reference. *Ergo*, 3(5):113–131, 2016. doi: 10.3998/ergo.12405314.0003.005.

Harvey Friedman and Michael Sheard. An axiomatic approach to self-referential truth. *Annals of Pure and Applied Logic*, 33:1–21, 1987. doi: 10.1016/0168-0072(87)90073-X.

Haim Gaifman. Operational pointer semantics: Solution to self-referential puzzles i. In Moshe Y. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 43–59, Los Altos, 1988. Morgan Kaufmann.

Haim Gaifman. Pointers to truth. *Journal of Philosophy*, 89(5):223–261, 1992. doi: 10.2307/2027167.

Christopher Gauker. Against stepping back: A critique of contextualist approaches to the semantic paradoxes. *Journal of Philosophical Logic*, 35(4):393–422, 2006. doi: 10.1007/s10992-006-9026-y.

Jean-Yves Girard, Paul Taylor, and Yves Lafont. *Proofs and Types*. Cambridge University Press, Cambridge, 1989.

Michael Glanzberg. The liar in context. *Philosophical Studies*, 103(3):217–251, 2001. doi: 10.1023/A:1010314719817.

Michael Glanzberg. A contextual-hierarchical approach to truth and the liar paradox. *Journal of Philosophical Logic*, 33(1):27–88, 2004a. doi: 10.1023/B:LOGI.0000019227.09236.f5.

Michael Glanzberg. Quantification and realism. *Philosophy and Phenomenological Research*, 69(3):541–572, 2004b. doi: 10.1111/j.1933-1592.2004.tb00518.x.

Michael Glanzberg. Truth, reflection, and hierarchies. *Synthese*, 142(3):289–315, 2004c. doi: 10.1007/s11229-005-3718-7.

Michael Glanzberg. Context and unrestricted quantification. In Rayo and Uzquiano (2006), pages 45–74.

Michael Glanzberg. Complexity and hierarchy in truth predicates. In Achourioti et al. (2015), pages 211–243. doi: 10.1007/978-94-017-9673-6_10.

Patrick Grim. *The Incomplete Universe*. MIT Press, Cambridge, 1991.

Viacheslav Nikolaevich Grishin. Predicate and set-theoretic calculi based on logic without contractions. *Mathematics of the USSR–Izvestiya*, 18(1):41–59, 1982. doi: 10.1070/

IM1982v018n01ABEH001382.

Willem Groeneveld. Dynamic semantics and circular propositions. *Journal of Philosophical Logic*, 23(3):267–306, 1994. doi: 10.1007/BF01048483.

Anil Gupta. Truth and paradox. *Journal of Philosophical Logic*, 11(1):1–60, 1982. doi: 10.1007/BF00302338. Reprinted in Martin 1984: 175–235.

Anil Gupta and Nuel Belnap. *The Revision Theory of Truth*. MIT Press, Cambridge, 1993.

Volker Halbach. Tarskian and kripkean truth. *Journal of Philosophical Logic*, 26(1):69–80, 1997. doi: 10.1023/A:1017977304199.

Volker Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, Cambridge, 2011.

Lars Hallnäs. Partial inductive definitions. *Theoretical Computer Science*, 87(1):115–142, 1991. doi: 10.1016/S0304-3975(06)80007-1.

Lars Hallnäs and Peter Schroeder-Heister. A proof-theoretic approach to logic programming ii: Programs as definitions. *Journal of Logic and Computation*, 1(5):635–660, 1991. doi: 10.1093/logcom/1.5.635.

Hans G. Herzberger. The truth-conditional consistency of natural language. *Journal of Philosophy*, 64(2):29–35, 1967. doi: 10.2307/2023768.

Hans G. Herzberger. Paradoxes of grounding in semantics. *Journal of Philosophy*, 67(6): 146–167, 1970. doi: 10.2307/2023885.

Hans G. Herzberger. Notes on naive semantics. *Journal of Philosophical Logic*, 11(1):61–102, 1982. doi: 10.1007/BF00302339. Reprinted in Martin 1984: 133–174.

Leon Horsten. *The Tarskian Turn: Deflationism and Axiomatic Truth*. MIT Press, Cambridge, 2011.

Paul Horwich. *Truth*. Basil Blackwell, Oxford, 1990.

Robert C. Koons. *Paradoxes of Belief and Strategic Rationality*. Cambridge University Press, Cambridge, 1992.

Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72(19):690–716, 1975. doi: 10.2307/2023885. Reprinted in Martin 1984: 54–81.

Hannes Leitgeb. Truth and the liar in de morgan-valued models. *Notre Dame Journal of Formal Logic*, 40(4):496–514, 1999. doi: 10.1305/ndjfl/1012429715.

Hannes Leitgeb. What truth depends on. *Journal of Philosophical Logic*, 34(2):155–192, 2005. doi: 10.1007/s10992-004-3758-3.

Greg Littmann and Keith Simmons. A critique of dialetheism. In Priest et al. (2004), pages 314–335.

Grzegorz Malinowski. Q-consequence operation. *Reports on Mathematical Logic*, 24:49–59, 1990.

Robert L. Martin, editor. *The Paradox of the Liar*. Ridgeview, Atascadero, 1970.

Robert L. Martin, editor. *Recent Essays on Truth and the Liar Paradox*. Oxford University Press, Oxford, 1984.

Robert L. Martin and Peter W. Woodruff. On representing 'true-in-l' in l. *Philosophia*, 5: 213–217, 1975. Reprinted in Martin 1984: 47–51.

Vann McGee. *Truth, Vagueness, and Paradox*. Hackett, Indianapolis, 1991.

Toby Meadows. Truth, dependence and supervaluation: Living with the ghost. *Journal of Philosophical Logic*, 42(2):221–240, 2013. doi: 10.1007/s10992-011-9219-x.

Toby Meadows. Fixed points for consequence relations. *Logique et Analyse*, 57(227):333–357, 2014.

Ignacio Ojea. The structural collapse approach reconsidered. *Análisis Filosófico*, 32(1): 61–68, 2012.

Charles Parsons. The liar paradox. *Journal of Philosophical Logic*, 3(4):381–412, 1974. doi: 10.1007/BF00257482. Reprinted in Parsons, *Mathematics in Philosophy*, Ithaca: Cornell University Press, 1983, 221–250.

Terence Parsons. Assertion, denial, and the liar paradox. *Journal of Philosophical Logic*, 13 (2):137–152, 1984. doi: 10.1007/BF00453019.

Douglas Patterson. Understanding the liar. In *Beall 2007*, pages 197–224. 2007. In Beall 2007: 197–224.

Douglas Patterson. Inconsistency theories of semantic paradox. *Philosophy and Phenomenological Research*, 79(2):387–422, 2009. doi: 10.1111/j.1933-1592.2009.00283.x.

Uwe Petersen. Logic without contraction as based on inclusion and unrestricted abstraction. *Studia Logica*, 64(3):365–403, 2000. doi: 10.1023/A:1005293713265.

Lavinia Picollo. The old-fashioned yablo paradox. *Análisis Filosófico*, 32(1):21–29, 2012.

Graham Priest. Logic of paradox revisited. *Journal of Philosophical Logic*, 13(2):153–179, 1984. doi: 10.1007/BF00453020.

Graham Priest. Yablo's paradox. *Analysis*, 57(4):236–242, 1997. doi: 10.1111/1467-8284.00081.

Graham Priest. *In Contradiction*. Oxford University Press, Oxford, 2 edition, 2006.

Graham Priest. *An Introduction to Non-Classical Logic*. Cambridge University Press, Cambridge, 2 edition, 2008.

Graham Priest, Jc Beall, and Bradley Armour-Garb, editors. *The Law of Non-Contradiction*. Oxford University Press, Oxford, 2004.

Shahid Rahman, Tero Tulenheimo, and Emmanuel Genot, editors. *Unity, Truth and the Liar: The Modern Relevance of Medieval Solutions to the Liar Paradox*. Springer Verlag, Berlin, 2008.

Agustín Rayo and Gabriel Uzquiano, editors. *Absolute Generality*. Oxford University Press, Oxford, 2006.

Stephen Read. The liar paradox from john buridan back to thomas bradwardine. *Vivarium*, 40(2):189–218, 2002. doi: 10.1163/156853402320901812.

Stephen Read. Symmetry and paradox. *History and Philosophy of Logic*, 27(4):307–318, 2006. doi: 10.1080/01445340600593942.

Greg Restall. *On Logics Without Contraction*. Ph.d. thesis, University of Queensland, 1994.

Greg Restall. Modal models for bradwardine's theory of truth. *Review of Symbolic Logic*, 1 (2):225–240, 2008. doi: 10.1017/S1755020308080180.

Jr. Richard G. Heck. Tarski, truth, and semantics. *Philosophical Review*, 106(4):533–554, 1997. doi: 10.2307/2998511.

Jr. Richard G. Heck. Self-reference and the languages of arithmetic. *Philosophia Mathematica*, 15(1):1–29, 2007. doi: 10.1093/philmat/nkl028.

Jr. Richard G. Heck. A liar paradox. *Thought*, 1(1):36–40, 2012. doi: 10.1002/tht3.5.

David Ripley. Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, 5(2):354–378, 2012. doi: 10.1017/S1755020312000056.

David Ripley. Paradoxes and failures of cut. *Australasian Journal of Philosophy*, 91(1): 139–164, 2013a. doi: 10.1080/00048402.2011.630010.

David Ripley. Revising up: Strengthening classical logic in the face of paradox. *Philosophers' Imprint*, 13(5):1–13, 2013b. Available online.

David Ripley. Comparing substructural theories of truth. *Ergo*, 2(13):299–328, 2015. doi: 10.3998/ergo.12405314.0002.013.

Kevin Scharp. *Replacing Truth*. Oxford University Press, Oxford, 2013.

Thomas Schindler. Axioms for grounded truth. *Review of Symbolic Logic*, 7(1):73–83, 2014. doi: 10.1017/S1755020313000282.

Peter Schroeder-Heister. Cut-elimination in logics with definitional reflection. In David Pearce and Heinrich Wansing, editors, *Nonclassical Logics and Information Processing*, pages 146–171. Springer, Berlin, 1992.

Peter Schroeder-Heister. On the notion of assumption in logical systems. In Renate Bluhm and Christian Nimtz, editors, *Selected Papers Contributed to the Sessions of GAP5*, pages 27–48. Mentis, Paderborn, 2004.

Lionel Shapiro. The rationale behind the revision theory. *Philosophical Studies*, 129(3):477–515, 2006. doi: 10.1007/s11098-004-2497-1.

Lionel Shapiro. Deflating logical consequence. *Philosophical Quarterly*, 61(243):320–342, 2011a. doi: 10.1111/j.1467-9213.2010.678.x.

Lionel Shapiro. Expressibility and the liar's revenge. *Australasian Journal of Philosophy*, 89 (2):297–314, 2011b. doi: 10.1080/00048401003695156.

Lionel Shapiro. Naive structure, contraction, and paradox. *Topoi*, 34(1):75–87, 2015. doi: 10.1007/s11245-014-9235-x.

Stewart Shapiro. Simple truth, contradiction, and consistency. In Priest et al. (2004), pages 336–354.

Keith Simmons. *Universality and the Liar*. Cambridge University Press, Cambridge, 1993.

Scott Soames. *Understanding Truth*. Oxford University Press, Oxford, 1999.

Roy A. Sorensen. Yablo's paradox and kindred infinite liars. *Mind*, 107(425):137–155, 1998. doi: 10.1093/mind/107.425.137.

Roy A. Sorensen. *A Brief History of Paradox*. Oxford University Press, Oxford, 2003.

Alfred Tarski. Der wahrheitsbegriff in den formalisierten sprachen. *Studia Philosophica*, 1: 261–405, 1935. References are to the English translation by J. H. Woodger, "The concept of truth in formalized languages", in *Logic, Semantics, Metamathematics*, 2nd ed., Indianapolis: Hackett, 1983, ed. J. Corcoran, pp. 152–278.

Alfred Tarski. The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3):341–375, 1944. doi: 10.2307/2102968.

Paula Teijeiro. Circularity is still scary. *Análisis Filosófico*, 32(1):31–35, 2012.

Neil Tennant. Proof and paradox. *Dialectica*, 36(2–3):265–296, 1982. doi: 10.1111/j.1746-8361.1982.tb00820.x.

Bas C. van Fraassen. Presupposition, implication, and self-reference. *Journal of Philosophy*, 65(5):136–152, 1968. doi: 10.2307/2024557.

Bas C. van Fraassen. Truth and paradoxical consequence. In Martin (1970), pages 13–23.

Albert Visser. Four valued semantics and the liar. *Journal of Philosophical Logic*, 13(2): 181–212, 1984. doi: 10.1007/BF00453021.

Alan Weir. Naive set theory, paraconsistency, and indeterminacy: Part i. *Logique et Analyse*, 41(161–163):219–266, 1998.

Alan Weir. Naive set theory, paraconsistency, and indeterminacy: Part ii. *Logique et Analyse*, 42(167–168):283–340, 1999.

Alan Weir. Naive truth and sophisticated logic. In Beall and Armour-Garb (2005), pages 218–249.

Alan Weir. A robust non-transitive logic. *Topoi*, 34(1):99–107, 2015. doi: 10.1007/s11245-013-9176-9.

Peter W. Woodruff. Paradox, truth, and logic part 1: Paradox and truth. *Journal of Philosophical Logic*, 13(2):213–232, 1984. doi: 10.1007/BF00453022.

Stephen Yablo. Hop, skip, and jump: The agonistic conception of truth. *Philosophical Perspectives*, 7:371–396, 1993a. doi: 10.2307/2214130.

Stephen Yablo. Paradox without self-reference. *Analysis*, 53(4):251–252, 1993b. doi: 10.2307/3328245.

Elia Zardini. Truth without contra(di)ction. *Review of Symbolic Logic*, 4(4):498–535, 2011. doi: 10.1017/S1755020311000177.

Elia Zardini. Naive modus ponens. *Journal of Philosophical Logic*, 42(4):575–593, 2013. doi: 10.1007/s10992-012-9239-1.

(IV) Paradoxo de Russell[1]

Título Original: Russell's Paradox
Autor: Andrew David Irvine e Harry Deutsch
Tradução: Paloma de Souza Xavier
Revisão: Ederson Safra Melo

O paradoxo de Russell é o mais famoso dos paradoxos lógicos ou paradoxos da teoria dos conjuntos. Também conhecido como paradoxo de Russell-Zermelo, ele surge na teoria ingênua dos conjuntos ao se considerar o conjunto de todos os conjuntos que não são membros de si mesmos. Esse conjunto parece ser membro de si mesmo se, e somente se, não for membro de si mesmo. Daí o paradoxo.

Alguns conjuntos, como o conjunto de todas as xícaras de chá, não são membros de si mesmos. Outros conjuntos, como o conjunto de todos os não xícaras de chá, são membros de si mesmos. Chamemos de 'R' o conjunto de todos os conjuntos que não são membros de si mesmos. Se R é membro de si mesmo, então, por definição, não deve ser membro de si mesmo. Da mesma forma, se R não é membro de si mesmo, então, por definição, deve ser membro de si mesmo.

Embora também tenha sido notada por Ernst Zermelo, a contradição não foi considerada importante até ser descoberta de forma independente por Bertrand Russell na primavera de 1901. Desde então, o paradoxo tem motivado um grande volume de trabalho em lógica,

[1] IRVINE, Andrew David; DEUTSCH, Harry, "Russell's Paradox", In: ZALTA, E. N. (ed.). The Stanford Encyclopedia of Philosophy (Spring 2021 Edition). Stanford, CA: The Metaphysics Research Lab, 2021. Disponível em: https://plato.stanford.edu/archives/spr2021/entries/russell-paradox/.
A seguir está a tradução da entrada sobre o Paradoxo de Russell de Andrew David Irvine e Harry Deutsch na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/spr2021/entries/russell-paradox/. Esta versão traduzida pode diferir da versão atual da entrada, que pode ter sido atualizada desde o momento desta tradução. A versão atual está localizada em https://plato.stanford.edu/entries/russell-paradox/. Agradecemos ao Prof. Dr. Edward N. Zalta pela permissão para traduzir e publicar esta entrada.

teoria dos conjuntos, e na filosofia e fundamentos da matemática.

1. O paradoxo

Indispensável a qualquer teoria de conjuntos é a formulação das condições sob as quais os conjuntos são formados. Além da simples listagem dos membros de um conjunto, assumiu-se inicialmente que qualquer condição bem definida (ou propriedade precisamente especificada) poderia ser utilizada para determinar um conjunto. Por exemplo, se T é a propriedade de ser uma xícara de chá, então o conjunto S de todas as xícaras de chá pode ser definido como $S = \{x : T(x)\}$, ou seja, o conjunto de todos os indivíduos x tal que x possui a propriedade T. Mesmo uma propriedade contraditória pode ser utilizada para determinar um conjunto. Por exemplo, a propriedade de ser simultaneamente T e não T determinaria o conjunto vazio, ou seja, o conjunto que não possui membros.

Mais precisamente, a teoria ingênua dos conjuntos assume o chamado Axioma da Compreensão ingênua ou irrestrita, que estabelece que, para qualquer fórmula $\phi(x)$ que contenha x como variável livre, existirá o conjunto $x : \phi(x)$ cujos membros são exatamente os objetos que satisfazem $\phi(x)$. Assim, se a fórmula $\phi(x)$ representa "x é primo", então $\{x : \phi(x)\}$ será o conjunto dos números primos. Se $\phi(x)$ representar "$\sim (x = x)$", então $\{x : \phi(x)\}$ será o conjunto vazio.

Mas, a partir da suposição deste axioma, segue-se a contradição de Russell. Por exemplo, se considerarmos que $\phi(x)$ representa $x \in x$ e definirmos $R = \{x :\sim \phi(x)\}$, então R é o conjunto cujos membros são exatamente os objetos que não são membros de si mesmos.

R é um membro de si mesmo? Se for, então tem de satisfazer a condição de não ser membro de si mesmo e, portanto, não é membro de si mesmo. Se não for, então não deve satisfazer a condição de não ser membro de si mesmo e, portanto, tem de ser membro de si mesmo. Uma vez que, pela lógica clássica, um dos dois casos deve ser válido — ou R é um membro de si mesmo, ou não é —, segue-se que a teoria implica uma contradição.

Como Russell nos diz, foi depois de aplicar o mesmo tipo de raciocínio encontrado no argumento diagonal de Cantor a uma "suposta classe de todos os objetos imagináveis" que ele foi conduzido à contradição:

> A classe abrangente que estamos considerando, que deve englobar tudo, tem de englobar a si mesmo como um dos seus membros. Em outras palavras,

se existe algo como "tudo", então, "tudo" é algo e é um membro da classe "tudo". Mas, normalmente, uma classe não é um membro de si mesmo. A humanidade, por exemplo, não é um homem. Agora, forme o conjunto de todas as classes que não são membros de si mesmas. Esta é uma classe: é membro de si mesma ou não? Se for, então pertence ao grupo das classes que não são membros de si mesmas, ou seja, não é membro de si mesma. Se não for, então não pertence a esse grupo, ou seja, é membro de si mesmo. Assim, das duas hipóteses — ser ou não ser membro de si mesmo — cada uma implica sua contraditória. Isso é uma contradição. (Russell, 1919, p. 136)

As respostas padrão ao paradoxo tentam limitar, de alguma forma, as condições sob as quais os conjuntos são formados. O objetivo é, normalmente, eliminar R (e conjuntos contraditórios semelhantes) e, ao mesmo tempo, manter todos os outros conjuntos necessários para a matemática. Isto é frequentemente feito substituindo o Axioma da Compreensão irrestrito pelo Axioma da Separação, mais restritivo, ou seja, o axioma que, dado qualquer conjunto (consistente) S e qualquer fórmula $\phi(x)$ com x livre, haverá um conjunto $\{x \in S : \phi(x)\}$ cujos membros são exatamente os membros de S que satisfazem $\phi(x)$. Se agora deixarmos $\phi(x)$ representar a fórmula $x \notin x$, verifica-se que o conjunto correspondente, $x \in S : x \notin x$ não será contraditório, uma vez que é constituído apenas pelos membros encontrados em S que não são membros de si mesmos. Portanto, o conjunto falha em incluir a si mesmo.

2. História do Paradoxo

Russell parece ter descoberto seu paradoxo no final da primavera de 1901, enquanto trabalhava em seus The Principles of Mathematics (1903). A data exata da descoberta não é clara. Russell afirma inicialmente que se deparou com o paradoxo "em junho de 1901" (1944, 13). Mais tarde, ele relata que a descoberta ocorreu "na primavera de 1901" (1959, 75). Ainda mais tarde, ele relata que se deparou com o paradoxo, não em junho, mas em maio daquele ano (1969, 221). Cesare Burali-Forti, um assistente de Giuseppe Peano, havia descoberto uma antinomia semelhante em 1897, quando reparou que, uma vez que o conjunto dos ordinais é bem ordenado, ele também deve ter um ordinal. No entanto, este ordinal tem de ser simultaneamente um elemento do conjunto de todos os ordinais e, ainda assim, maior do que todos esses elementos.

Ao contrário do paradoxo de Burali-Forti, o paradoxo de Russell não envolve ordinais nem cardinais, baseando-se, ao invés disso, apenas nas noções primitivas de conjunto e inclusão de conjuntos. Zermelo percebeu uma contradição semelhante entre 1897 e 1902, possivelmente antecipando Russell em alguns anos (Ebbinghaus e Peckhaus 2007, 43-48; Tappenden 2013, 336), embora Kanamori conclua que a descoberta poderia muito bem ter ocorrido já em 1902 (Kanamori 2009, 411). Como destaca Linsky, o argumento de Zermelo, embora semelhante ao de Russell, é melhor entendido como um dos vários argumentos de Zermelo, Schröder e Cantor que "de fato anteciparam" o argumento matemático desenvolvido por Russell, mas que acabaram sendo diferentes em pequenos, mas significativos, aspectos do argumento de Russell (Linsky 2013, 11). Em todo o caso, os argumentos foram considerados de menor importância até se perceber como eram prejudiciais para os fundamentos da aritmética de Gottlob Frege.

Russell escreveu a Frege com notícias sobre seu paradoxo em 16 de junho de 1902. (Para a relevante correspondência, ver Russell (1902) e Frege (1902) em van Heijenoort (1967). O paradoxo foi significativo para o trabalho lógico de Frege, pois, na prática, mostrou que os axiomas que ele estava usando para formalizar a sua lógica eram inconsistentes. Especificamente, o Axioma V de Frege exige que uma expressão como $\phi(x)$ seja considerada tanto uma função do argumento x como uma função do argumento ϕ. (Mais precisamente, a Lei de Frege afirma que o curso de valores de um conceito f é idêntico ao de valores de um conceito g se, e somente se, f e g concordam no valor de cada argumento, isto é, se, e somente se, para cada objeto x, $f(x) = g(x)$. De fato, foi esta ambiguidade que permitiu a Russell construir R de modo que pudesse ser e não ser membro de si mesmo.

A carta de Russell chegou justamente quando o segundo volume dos *Grundgesetze der Arithmetik* (As Leis Básicas da Aritmética), 1893, 1903) de Frege estava sendo impresso. Percebendo de imediato a dificuldade que o paradoxo colocava, Frege acrescentou ao *Grundgesetze* um apêndice escrito às pressas discutindo a descoberta de Russell. No apêndice, Frege observa que as consequências do paradoxo de Russell não são imediatamente claras. Por exemplo, "É sempre permitido falar da extensão de um conceito, de uma classe? E, se não for, como reconhecemos os casos excepcionais? Podemos sempre inferir que, se a extensão de um conceito coincide com a de outro, então todos os objetos que se enquadram no primeiro conceito também se enquadram no segundo? Estas são as questões", observa Frege, "levantadas pela comunicação do Sr. Russell" (1903, p. 127). Devido a essas preocupações, Frege acabou se sentindo forçado a abandonar muitos dos seus pontos de vista sobre lógica e matemática.

Mesmo assim, como Russell salienta, Frege enfrentou a notícia do paradoxo com uma firmeza notável:

> Quando penso em atos de integridade e virtude, percebo, que não há nada em meu conhecimento que se compare à dedicação de Frege à verdade. A obra de toda a sua vida estava prestes a ser concluída, grande parte do seu trabalho havia sido ignorado em benefício de homens infinitamente menos capazes, seu segundo volume estava prestes a ser publicado e, ao descobrir que a sua suposição fundamental estava errada, ele reagiu com um prazer intelectual que claramente sobrepujou qualquer sentimento de desilusão pessoal. Foi algo quase sobre-humano e uma poderosa demonstração do que as pessoas são capazes quando sua dedicação está voltada para o trabalho criativo e para o conhecimento, em vez de esforços mais grosseiros para dominar e ser reconhecidas. (van Heijenoort, 1967, p. 127)

É claro que Russell também estava preocupado com as consequências da contradição. Ao saber que Frege concordava com ele sobre o significado do resultado, começou imediatamente a escrever um apêndice para sua obra The Principles of Mathematics, que seria lançada em breve. Intitulado "Apêndice B: A Doutrina dos Tipos", o apêndice representa a primeira tentativa de Russell de apresentar um método fundamentado para evitar o que depois se tornaria conhecido como "o paradoxo de Russell".

3. Primeiras respostas ao paradoxo

O significado do paradoxo de Russell pode ser compreendido quando se percebe que, usando a lógica clássica, qualquer proposição pode ser derivada a partir de uma contradição. Por exemplo, assumindo P e $\sim P$, qualquer proposição arbitrária, Q, pode ser provada da seguinte forma: a partir de P obtemos $P \vee Q$ pela regra da Adição; depois, a partir de $P \vee Q$ e $\sim P$ obtemos Q pela regra do Silogismo Disjuntivo. Como a teoria dos conjuntos serve de base para todos os ramos da matemática, muitas pessoas começaram a se preocupar com o fato de que a inconsistência da teoria dos conjuntos poderia significar que nenhuma prova matemática seria completamente confiável. Somente ao eliminar o paradoxo de Russell é que a matemática como um todo poderia recuperar a sua consistência.

O paradoxo de Russell deriva, em última análise, da ideia de que qualquer condição ou propriedade pode ser usada para definir um conjunto. Por exemplo, a propriedade de ser divisível apenas por si mesmo e pelo número um distingue o conjunto dos números primos do conjunto dos números inteiros. A propriedade de possuir glândulas mamárias distingue o conjunto dos mamíferos dos répteis, aves e outros organismos vivos. A propriedade de ser simultaneamente quadrado e não quadrado (ou qualquer outra conjunção de propriedades contraditórias) define o conjunto vazio, e assim por diante.

Georg Cantor, o criador da teoria moderna dos conjuntos, foi um dos primeiros céticos em relação a um axioma de Compreensão (ou Abstração) irrestrita. Mesmo antes da descoberta de Russell, Cantor já havia rejeitado a Compreensão irrestrita a favor do que era, de fato, uma distinção entre conjuntos e classes, reconhecendo que algumas propriedades (como a propriedade de ser um ordinal) produziam coleções que eram simplesmente grandes demais para serem conjuntos, e que qualquer suposição contrária levaria à inconsistência. (Detalhes podem ser encontrados em Moore (1982), Hallett (1984) e Menzel (1984)).

A resposta de Russell ao paradoxo veio com a sua teoria dos tipos, neomeada de forma bastante apropriada. Acreditando que a autoaplicação estava no cerne do paradoxo, a ideia básica de Russell era que poderíamos evitar o compromisso com R (o conjunto de todos os conjuntos que não são membros de si mesmos) organizando todas as proposições (ou, mais precisamente, todas as funções proposicionais, funções que atribuem proposições como valores) em uma hierarquia. Assim, só seria possível referir-se a todos os objetos para os quais uma determinada condição (ou predicado) se aplica se eles estiverem no mesmo nível ou pertencerem ao mesmo "tipo".

Esta solução para o paradoxo de Russell é motivada, em grande parte, pela adoção do chamado princípio do círculo vicioso. O princípio afirma, de fato, que nenhuma função proposicional pode ser definida antes que se verifique o seu âmbito de aplicação. Em outras palavras, antes que uma função possa ser definida, é necessário primeiro determinar exatamente quais objetos pertencem ao seu domínio. Por exemplo, antes de definir o predicado "é um número primo", é preciso primeiro definir o conjunto de objetos que podem satisfazer este predicado, ou seja, o conjunto dos números naturais. Como explicam Whitehead e Russell,

> Uma análise dos paradoxos a serem evitados mostra que todos eles resultam de uma espécie de círculo vicioso. Esses círculos viciosos decorrem da suposição de que uma coleção de objetos pode conter membros que só podem ser

> definidos por meio da coleção como um todo. Assim, por exemplo, supõe-se que a coleção de proposições contenha uma proposição que afirma que "todas as proposições são ou verdadeiras ou falsas". Parece, no entanto, que tal afirmação só poderia ser legítima se "todas as proposições" se referisse a alguma coleção já definida, o que não é possível se novas proposições forem criadas por afirmações sobre "todas as proposições". Teremos, portanto, de afirmar que tais afirmações são destituídas de sentido. (...) O princípio que nos permite evitar totalidades ilegítimas pode ser formulado da seguinte maneira: "Tudo o que envolve a totalidade de uma coleção não deve ser um de seus membros"; ou, de forma equivalente: "Se, supondo que uma certa coleção tivesse um todo, ela tivesse membros apenas definíveis em termos desse todo, então essa coleção não possui um todo." Chamaremos isso de "princípio do círculo vicioso", pois ele nos permite evitar os círculos viciosos implicados na suposição de totalidades ilegítimas. (1910, 2ª ed. 37)

Se Whitehead e Russell estiverem certos, segue-se que o âmbito de aplicação de uma função jamais poderá incluir qualquer objeto que pressuponha a própria função. Como resultado, as funções proposicionais (juntamente com suas proposições correspondentes) acabarão sendo organizadas em uma hierarquia, tal como propõe Russell.

Embora Russell tenha apresentado pela primeira vez sua teoria dos tipos em The Principles of Mathematics (1903), reconheceu de imediato que seria necessário aprofundá-la, já que sua formulação inicial parecia resolver apenas alguns dos paradoxos. Entre as alternativas que considerou, estava a chamada teoria da substituição (Galaugher 2013). Isso levou, por sua vez, a uma formulação mais madura da teoria dos tipos, cinco anos depois, no artigo de 1908, Lógica Matemática Baseada na Teoria dos Tipos, e na obra monumental que coescreveu com Alfred North Whitehead, Principia Mathematica (1910, 1912, 1913). A teoria dos tipos de Russell aparece, assim, em duas versões: a "teoria simples", de 1903, e a "teoria ramificada", de 1908. Ambas foram criticadas por serem excessivamente ad hoc, não conseguindo eliminar o paradoxo de forma satisfatória.

Em resposta ao paradoxo de Russell, David Hilbert também ampliou seu programa de construção de uma base consistente e axiomática para a matemática, de modo a incluir uma fundamentação axiomática para a lógica e a teoria dos conjuntos (Peckhaus 2004). Subjacente a essa abordagem formalista estava a ideia de permitir apenas o uso de objetos finitos, bem definidos e construtíveis, juntamente com regras de inferência consideradas absolutamente certas.

Por fim, Luitzen Brouwer desenvolveu o intuicionismo, cuja ideia básica era que não se pode afirmar a existência de um objeto matemático a menos que se possa definir um procedimento para construí-lo.

Em conjunto, todas essas respostas ajudaram a concentrar a atenção nas conexões entre lógica, linguagem e matemática. Também auxiliaram os lógicos a desenvolver uma consciência explícita da natureza dos sistemas formais e dos tipos de resultados metalógicos e metamatemáticos que se revelaram centrais para a pesquisa nos fundamentos da lógica e da matemática ao longo dos últimos cem anos.

4. O Paradoxo de Russell na lógica contemporânea

O paradoxo de Russell é, às vezes, visto como um desenvolvimento negativo – como a ruína do Grundgesetze de Frege e como um dos pecados conceituais originais que levaram à nossa expulsão do paraíso de Cantor. W. V. Quine descreve o paradoxo como uma “antinomia” que “traz uma surpresa que só pode ser acomodada por nada menos que um repúdio à nossa herança conceitual” (1966, p. 11). Quine está se referindo ao princípio da Compreensão Ingênua mencionado anteriormente. Em símbolos, o princípio afirma que:

(CI) $\exists A \forall x (x \in A \equiv \phi)$

Onde A não ocorre livre na fórmula ϕ. Isso quer dizer: “Existe um conjunto A tal que, para qualquer objeto x, x é um elemento de A se, e somente se, a condição expressa por ϕ for satisfeita.” O paradoxo de Russell surge ao se tomar ϕ como a fórmula: $x \notin x$.

Apesar do comentário de Quine, é possível ver o paradoxo de Russell sob uma luz mais positiva. Por um lado, embora o assunto ainda seja controverso, pesquisas posteriores revelaram que o paradoxo não prejudica necessariamente a derivação da aritmética de Frege a partir da lógica. A versão de Frege do princípio de Compreensão (seu Axioma V) pode simplesmente ser abandonada. Por outro lado, Church apresenta uma formulação elegante da teoria simples dos tipos, que se mostrou frutífera até mesmo em áreas distantes dos fundamentos da matemática. Além disso, o desenvolvimento de teorias axiomáticas de conjuntos (em oposição às ingênuas), que apresentam várias maneiras engenhosas, e significativas tanto do ponto de vista filosófico quanto matemático, de lidar com o paradoxo de Russell, preparou o terreno para resultados impressionantes na metamatemática da teoria dos conjuntos. Esses resultados incluíram os teoremas de Gödel e Cohen sobre a independência do axioma da escolha e a hipótese do continuum de Cantor. Então, vamos ver, grosso modo, como alguns desses métodos – especificamente, os chamados métodos 'não tipados' – lidam com o paradoxo de Russell.

Zermelo substitui CI pelo seguinte esquema axiomático de Separação (ou axioma de Ausência):

(ZA) $\forall A \exists B \forall x(x \in B \equiv (x \in A \wedge \phi))$

Novamente, para evitar circularidade, B não pode ser livre em ϕ. Isso exige que, para ser membro de B, x deve ser um elemento de um conjunto existente A. Como se pode prever, isso requer uma série de axiomas adicionais de existência de conjuntos, nenhum dos quais seria necessário se NC tivesse sido mantido.

Como a ZA evita o paradoxo de Russell? A princípio, pode-se pensar que não. Afinal, se tomarmos A como V – o universo completo dos conjuntos – e ϕ como $x \notin x$, parece surgir uma contradição novamente. Mas, nesse caso, tudo o que a contradição demonstra é que V não é um conjunto. Tudo o que a contradição demonstra é que "V" é um nome vazio (ou seja, que não tem referência, que V não existe), já que a ontologia do sistema de Zermelo consiste exclusivamente em conjuntos.

Esse mesmo ponto pode ser apresentado de outra maneira, envolvendo uma forma relativizada do argumento de Russell. Seja B um conjunto qualquer. Pelo axioma de separação (ZA), o conjunto $Rb = x \in B : x \notin x$ existe, mas ele não pode ser um elemento de B. Pois, se for um elemento de B, então podemos perguntar se ele pertence ou não a Rb; e ele pertence se, e somente se, não pertencer. Assim, algo, a saber, Rb, está "faltando" em cada conjunto B. Logo, novamente, V não é um conjunto, já que nada pode estar faltando em V. Mas observe a seguinte sutileza: ao contrário do argumento anterior, que envolve a aplicação direta do axioma de separação a V, o argumento em questão sugere a ideia de que, embora V não seja um conjunto, "V" não é um nome vazio. A próxima estratégia para lidar com o paradoxo de Russell explora essa sugestão.

O método não tipado de John von Neumann (1925) para lidar com paradoxos — em particular, com o paradoxo de Russell — é simples e engenhoso. Von Neumann introduz uma distinção entre pertencimento e não pertencimento e, com base nisso, traça uma distinção entre conjuntos e classes. Um objeto é um membro *(simpliciter)* se for membro de alguma classe; e é um não-membro se não for membro de nenhuma classe. (Na verdade, von Neumann desenvolve uma teoria de funções — tomadas como primitivas — em vez de uma teoria de classes, na qual a distinção membro/não-membro corresponde à distinção entre um objeto que pode ser argumento de alguma função e um que não pode. Em sua forma moderna, devida a Bernays e Gödel, trata-se de uma teoria de classes de tipo único.)

Os conjuntos são então definidos como membros, e os não membros são rotulados como "classes próprias". Assim, por exemplo, a classe de Russell, R, não pode ser membro de nenhuma classe e, portanto, deve ser uma classe própria. Se R for considerado um elemento de alguma classe A, então decorre de um dos axiomas de von Neumann que R não é equivalente a V. Mas R é equivalente a V e, portanto, não pode ser um elemento de A. Dessa forma, o método de von Neumann está intimamente relacionado ao resultado mencionado anteriormente sobre o conjunto RBR_{BRB}, para um B arbitrário. O método de von Neumann, embora admirado por nomes como Gödel e Bernays, tem sido subestimado nos últimos anos.

De maneira semelhante, Quine (1937) e (1967) oferecem outro método não tipado (em letra, senão em espírito) para bloquear o paradoxo de Russell — um método repleto de anomalias interessantes. A ideia básica de Quine é introduzir um axioma de compreensão estratificada. Esse axioma evita a circularidade ao introduzir uma hierarquia (ou estratificação) que é semelhante à teoria dos tipos em certos aspectos, e distinta dela em outros.

Em contraste com as estratégias de Zermelo, von Neumann e Quine, que são, de certo modo, puramente conjuntista, também houve tentativas de evitar o paradoxo de Russell por meio da alteração da lógica subjacente. Houve muitas tentativas desse tipo, e não analisaremos todas, mas uma se destaca por ser, no momento, ao mesmo tempo radical e relativamente popular (embora não entre os teóricos de conjuntos propriamente ditos): trata-se da abordagem paraconsistente, que limita o efeito geral de uma contradição isolada sobre toda uma teoria. A lógica clássica exige que qualquer contradição trivialize uma teoria, tornando todas as sentenças da teoria demonstráveis. Isso ocorre porque, na lógica clássica, o seguinte é um teorema:

(Princípio de explosão) $A \to (\sim A \to B)$.

Agora, praticamente a única maneira de evitar o Princípio de Explosão (PE) é abandonar o modus ponens, dadas as definições usuais dos conectivos. Assim, alterar a lógica sentencial básica dessa maneira é, de fato, radical — mas possível. Infelizmente, mesmo renunciando ao PE, isso não é suficiente para preservar alguma forma do princípio de Compreensão Ingênua (CI). Também é necessário abrir mão do seguinte teorema adicional da lógica sentencial básica:

(Contração) $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$.

Pode-se então argumentar que o princípio de Compreensão Ingênua leva diretamente, não apenas a uma contradição isolada, mas à trivialidade. (no paradoxo de Curry observe-se também que não é suficiente apenas manter o nome *“modus ponens”*; é a própria regra que se modifica nas lógicas não tradicionais). Assim, parece que os problemas da CI não se limitam ao paradoxo de Russell, mas também incluem um paradoxo sem negação, devido a Curry.

Outra sugestão poderia ser concluir que o paradoxo depende de uma instância do princípio do terceiro excluído, ou seja, ou R é um membro de R, ou não é. Esse é um princípio rejeitado por algumas abordagens não clássicas da lógica, incluindo o intuicionismo. No entanto, é possível formular o paradoxo sem apelar para o terceiro excluído, com base na lei da não contradição. Fazemos isso da seguinte forma: Dada a definição de R, segue-se que $R \in R \equiv\sim (R \in R)$. Portanto, $R \in R \rightarrow\sim (R \in R)$. dada a definição de R, segue-se que $R \in R \equiv\sim (R \in R)$. Portanto, $R \in R \rightarrow\sim (R \in R)$. Mas, também sabemos que $R \in R \rightarrow (R \in R\wedge \sim (R \in R))$. Mas, pela lei da não contradição, sabemos que $\sim (R \in R\wedge \sim (R \in R))$. Assim, por *modus tollens*, concluímos que $\sim (R \in R)$. Ao mesmo tempo, também sabemos que, uma vez que $R \in R \equiv\sim (R \in R)$, segue-se que $\sim (R \in R) \rightarrow R \in R$, e, portanto, que $R \in R$. Assim, podemos deduzir tanto $R \in R$ quanto sua negação, usando apenas métodos intuicionisticamente aceitáveis.

Parece, portanto, que os proponentes de lógicas não clássicas não podem reinvindicar que preservaram a CI em qualquer sentido significativo, além da preservação da forma puramente sintática do princípio, e tampouco o intuicionismo ou a paraconsistência, somados ao abandono da contração, oferecerão uma vantagem sobre as soluções não tipificadas de Zermelo, von Neumann ou Quine. (Discussões adicionais podem ser encontradas em Meyer, Routley e Dunn (1979), Irvine (1992), Priest (2006, cap. 18), Weber (2010), Weber (2012), e nas entradas sobre o paradoxo de Curry (seção 2.2) e lógica paraconsistente (seção 2.3).)

Também vale a pena observar que o paradoxo de Russell não foi o único paradoxo que o preocupou e, portanto, não foi a única motivação para as restrições tipológicas encontradas no *Principia Mathematica*. Em sua obra anterior, *The Principles of Mathematics*, Russell dedica um capítulo à “Contradição” (o paradoxo de Russell), apresentando-a em várias formas e descartando várias respostas que ele considera inválidas desde o início. Em seguida, sinaliza que discutirá “em breve” a doutrina dos tipos. Isso só ocorre várias centenas de páginas depois, no final do livro, no Apêndice B! Ali, Russell apresenta uma teoria incipiente e simplificada dos tipos — não a teoria dos tipos que encontramos no *Principia Mathematica*.

Por que a teoria posterior era necessária? A razão é que, no Apêndice B, Russell também apresenta outro paradoxo que, segundo ele, não poderia ser resolvido por meio da teoria simples dos tipos. Esse novo paradoxo diz respeito a proposições, e não a classes, e, juntamente com os paradoxos semânticos, levou Russell a formular sua versão ramificada da teoria dos tipos.

A nova versão proposicional do paradoxo não teve grande destaque no desenvolvimento subsequente da lógica e da teoria dos conjuntos, mas causou grande perplexidade em Russell. Por um lado, parece contradizer o teorema de Cantor. Russell escreve: ’Não podemos admitir que existam mais classes [ou domínios] de proposições do que proposições’ (1903, 527). A razão é que parece haver correlações fáceis, um a um, entre classes de proposições e proposições. Por exemplo, a classe m de proposições pode ser correlacionada com a proposição de que toda proposição em m é verdadeira. Isso, junto a um princípio detalhado de individuação para proposições (que afirma, por exemplo, que, se as classes m e n de proposições diferem, qualquer proposição sobre m será diferente de qualquer proposição sobre n), leva a uma contradição.

Essa discussão sobre o paradoxo tem sido relativamente pouca, embora tenha desempenhado um papel fundamental no desenvolvimento da lógica de sentido e denotação de Church. Embora tenhamos várias teorias de conjuntos para escolher, não temos algo como uma teoria bem desenvolvida das proposições russellianas, embora tais proposições sejam centrais para os pontos de vista dos millianistas e dos teóricos da referência direta. Esperava-se que essa teoria fosse necessária para os fundamentos da semântica, se não para os fundamentos da matemática. Assim, embora um dos paradoxos de Russell tenha levado ao desenvolvimento frutífero dos fundamentos da matemática, seu ’outro’ paradoxo ainda não resultou em nada remotamente semelhante nos fundamentos da semântica. Sem dúvida, Church (1974a) e Anderson (1989) tentaram desenvolver uma lógica intensional russelliana com base na teoria ramificada dos tipos, mas pode-se argumentar que a teoria ramificada é muito restritiva para servir de base à semântica da linguagem natural. Também houve algumas tentativas recentes de obter os primórdios de uma lógica intensional russelliana com base em teorias de conjuntos não tipificados (Cantini 2004; Deutsch 2014). É bastante irônico que, embora as proposições russellianas de granulação fina sejam favorecidas na filosofia da linguagem, o desenvolvimento formal da lógica intensional seja dominado pela gramática de Montague, com sua teoria de proposições de pouco refinada.

Também vale a pena observar que vários princípios que parecem pertencer puramente à teoria dos conjuntos são, na verdade, instâncias aplicadas de teoremas da lógica pura (isto é, da teoria da quantificação de primeira ordem com identidade). Há uma lista parcial desses exemplos em Kalish, Montague e Mar (2000). O paradoxo de Russell é uma instância do T269 nessa lista:

(T269) $\sim \exists y \forall x(Fxy \equiv \sim Fxx)$.

Lendo a letra do predicado diádico "F" como "é membro de", isso expressa que não existe um y tal que, para qualquer x, x é membro de y se e somente se x não for membro de x. Isso significa que o paradoxo de Russell se reduz ao T269?

Sem dúvida, a demonstração do T269 condensa a essência do argumento de Russell, seu padrão de raciocínio. Mas esse mesmo padrão também fundamenta uma lista interminável de "paradoxos" aparentemente frívolos, como o célebre paradoxo do barbeiro que barbeia todos e apenas aqueles que não se barbeiam, ou, de modo semelhante, o paradoxo do Deus benevolente mas eficiente que ajuda todos e apenas aqueles que não se ajudam.

Como esses "pseudoparadoxos", como às vezes são chamados, diferem, se é que diferem, do paradoxo de Russell? O esquema argumentativo é o mesmo e a conclusão – que não há tal barbeiro, nem tal Deus eficiente, nem tal conjunto de conjuntos que não são membros de si mesmos – também é a mesma: tais entidades simplesmente não existem. (No entanto, como mostrou von Neumann, não é necessário ir tão longe. O método de von Neumann nos instrui não que tais entidades como R não existam, mas apenas que não podemos dizer muito sobre elas, já que R e semelhantes não podem pertencer à extensão de nenhum predicado que se qualifique como uma classe.)

A resposta padrão a essa pergunta é que a diferença reside no objeto do discurso. Quine pergunta: "Por que ele [o paradoxo de Russell] conta como uma antinomia e o paradoxo do barbeiro não?"; e responde: "A razão é que existe, em nossos hábitos de pensamento, uma presunção avassaladora de que exista tal classe, mas nenhuma presunção de que exista tal barbeiro" (1966, p. 14). Ainda assim, a linguagem psicológica dos "hábitos de pensamento" não é particularmente esclarecedora. Mais especificamente, o paradoxo de

Russell dá origem, de modo significativo, à questão sobre quais conjuntos existem; mas é um contrassenso perguntar, com base no T269, quais barbeiros ou deuses existem.

Esse veredicto, entretanto, não é totalmente justo com os defensores do paradoxo do Barbeiro ou do T269 em geral. Eles insistirão que a questão levantada pelo T269 não é sobre quais barbeiros ou deuses existem, mas sim sobre quais objetos não paradoxais existem. Essa questão é praticamente a mesma levantada pelo próprio paradoxo de Russell. Assim, a partir dessa perspectiva, a relação entre o paradoxo do Barbeiro e o paradoxo de Russell é muito mais próxima do que muitos (seguindo Quine) estão dispostos a admitir (Salmon 2013).

Observamos que há uma fórmula lógica de primeira ordem que guarda com o princípio sobre os RBs a mesma relação que o T269 guarda com o paradoxo de Russell. Ela é a seguinte:

(T273) $\forall z \forall y (\forall x [Fxy \equiv (Fxz \wedge \sim Fxx)] \rightarrow \sim Fyz)$.

(Tomamos a liberdade de estender a numeração usada em Kalish, Montague e Mar (2000) até o T273.) Mas nem todos os paradoxos da teoria dos conjuntos estão relacionados de forma semelhante a teoremas da lógica de primeira ordem. O paradoxo de Burali-Forti é um exemplo, já que a noção de uma boa ordenação não é elementar, ou seja, não é definível em primeira ordem.

O paradoxo de Russell nunca foi superado, mas recentemente houve uma explosão de interesse por ele por parte de acadêmicos envolvidos em pesquisas de lógica matemática e em estudos filosóficos e históricos da lógica moderna. Um exame do conteúdo do volume *One Hundred Years of Russell's Paradox* (Cem anos do paradoxo de Russell), de 2004, revela lógicos matemáticos e filosóficos proeminentes, além de historiadores da lógica, debruçados sobre o paradoxo, propondo novos caminhos de volta ao paraíso de Cantor ou outras formas de resolver a questão. Suas investigações incluem modos radicalmente novos de escapar do dilema apresentado pelo paradoxo, novos estudos das teorias dos tipos (simples, ramificada e suas extensões), novas interpretações do paradoxo de Russell e das teorias construtivas, do paradoxo russelliano das proposições e da tentativa do próprio Russell de formular uma teoria sem tipos (a teoria da substituição), entre outros temas.

Tudo isso nos lembra que observações improváveis podem dar origem a trabalhos extremamente frutíferos. Como expressou Dana Scott: "Deve-se entender desde o início que o paradoxo de Russell *não* deve ser considerado um desastre. Ele e os paradoxos relacionados mostram que a noção ingênua de coleções totalizantes é insustentável. Isso *realmente é* um resultado interessante, sem dúvida" (1974, p. 207).

Referências Bibliográficas

C. Anthony Anderson. Russellian intensional logic. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes from Kaplan*, pages 67–103. Oxford University Press, Oxford, 1989.

Jon Barwise. *Admissible Sets and Structures*. Springer-Verlag, Berlin, 1975.

Jon Barwise and John Etchemendy. *The Liar: An Essay on Truth and Circularity*. Oxford University Press, Oxford, 1987.

Jon Barwise and Lawrence Moss. *Vicious Circles*. CSLI Publications, Stanford, 1996.

George Bealer. *Quality and Concept*. Oxford University Press, New York, 1982.

Michael Beaney. Russell and frege. In Nicholas Griffin, editor, *The Cambridge Companion to Bertrand Russell*, pages 128–170. Cambridge University Press, Cambridge, 2003.

Andrea Cantini. On a russellian paradox about propositions and truth. In Godehard Link, editor, *One Hundred Years of Russell's Paradox*, pages 259–284. Walter de Gruyter, Berlin and New York, 2004.

Andrea Cantini. Paradoxes, self-reference and truth in the 20th century. In Dov M. Gabbay and John Woods, editors, *Handbook of the History of Logic: Volume 5 – Logic From Russell to Church*, pages 875–1013. Elsevier/North Holland, Amsterdam, 2009.

Alonzo Church. Russellian simple type theory. *Proceedings and Addresses of the American Philosophical Association*, 47:21–33, 1974a.

Alonzo Church. Set theory with a universal set. In *Proceedings of the Tarski Symposium*, pages 297–308. 1974b. Repr. in International Logic Review, 15: 11–23.

Alonzo Church. A comparison of russell's resolution of the semantical antinomies with that of tarski. *Journal of Symbolic Logic*, 41:747–760, 1978. Repr. in A.D. Irvine (ed.), *Bertrand Russell: Critical Assessments*, vol. 2, Routledge, 1999, pp. 96–112.

Alberto Coffa. The humble origins of russell's paradox. *Russell*, 33–34:31–37, 1979.

Irving Copi. *The Theory of Logical Types*. Routledge and Kegan Paul, London, 1971.

William Demopoulos and Peter Clark. The logicism of frege, dedekind and russell. In Stewart Shapiro, editor, *The Oxford Handbook of Philosophy of Mathematics and Logic*, pages

129–165. Oxford University Press, Oxford, 2005.

Harry Deutsch. Resolution of some paradoxes of propositions. *Analysis*, 74:26–34, 2014.

Heinz-Dieter Ebbinghaus and Volker Peckhaus. *Ernst Zermelo: An Approach to His Life and Work*. Springer-Verlag, Berlin, 2007.

T. E. Forster. *Set Theory with a Universal Set*. Clarendon Press, Oxford, 2nd edition, 1995.

Gottlob Frege. The russell paradox. In *The Basic Laws of Arithmetic*, pages 127–143. University of California Press, Berkeley, 1964. Originally 1903; abridged and repr. in A.D. Irvine, *Bertrand Russell: Critical Assessments*, vol. 2, Routledge, 1999, pp. 1–3.

Gottlob Frege. Letter to russell. In Jean van Heijenoort, editor, *From Frege to Gödel*, pages 126–128. Harvard University Press, Cambridge, Mass., 1967. Originally 1902.

Dov M. Gabbay and John Woods, editors. *Handbook of the History of Logic: Volume 5 – Logic From Russell to Church*. Elsevier/North Holland, Amsterdam, 2009.

J. B. Galaugher. Substitution's unsolved 'insolubilia'. *Russell*, 33:5–30, 2013.

A. Garciadiego. *Bertrand Russell and the Origins of the Set-theoretic "Paradoxes"*. Birkhäuser, Boston, 1992.

I. Grattan-Guinness. How bertrand russell discovered his paradox. *Historia Mathematica*, 5: 127–137, 1978.

I. Grattan-Guinness. *The Search for Mathematical Roots: 1870–1940*. Princeton University Press, Princeton and Oxford, 2000.

Nicholas Griffin, editor. *The Cambridge Companion to Bertrand Russell*. Cambridge University Press, Cambridge, 2003.

Nicholas Griffin. The prehistory of russell's paradox. In Godehard Link, editor, *One Hundred Years of Russell's Paradox*, pages 349–371. Walter de Gruyter, Berlin and New York, 2004.

Nicholas Griffin, Bernard Linsky, and Kenneth Blackwell, editors. *Principia Mathematica at 100*. Bertrand Russell Research Centre, Hamilton, ON, 2011. Also published as Special Issue, Volume 31, Number 1 of *Russell*.

Michael Hallett. *Cantorian Set Theory and Limitation of Size*. Clarendon Press, Oxford, 1984.

Paul R. Halmos. *Naive Set Theory*. D. van Nostrand, Princeton, 1960.

A. D. Irvine. Gaps, gluts and paradox. *Canadian Journal of Philosophy (Supplementary Volume)*, 18:273–299, 1992.

A. D. Irvine, editor. *Philosophy of Mathematics*. Elsevier/North Holland, Amsterdam, 2009.

Donald Kalish, Richard Montague, and Gary Mar. *Logic: Techniques of Formal Reasoning*. Oxford University Press, New York, 2nd edition, 2000.

Akihiro Kanamori. Zermelo and set theory. *The Bulletin of Symbolic Logic*, 10:487–553, 2004.

Akihiro Kanamori. Set theory from cantor to cohen. In A. D. Irvine, editor, *Philosophy of Mathematics*, pages 395–459. Elsevier/North Holland, Amsterdam, 2009.

Kevin Klement. The origins of the propositional functions version of russell's paradox. *Russell*, 24:101–132, 2005.

Kevin Klement. The paradoxes and russell's theory of incomplete symbols. *Philosophical Studies*, 169:183–207, 2014.

Gregory Landini. The ins and outs of frege's way out. *Philosophia Mathematica*, 14:1–25, 2006.

Gregory Landini. Zermelo 'and' russell's paradox: Is there a universal set? *Philosophia Mathematica*, 21:180–199, 2013.

Azriel Levy. *Basic Set Theory*. Springer-Verlag and Heidelberg, Berlin and New York, 1979.

Godehard Link, editor. *One Hundred Years of Russell's Paradox*. Walter de Gruyter, Berlin and New York, 2004.

Bernard Linsky. Was the axiom of reducibility a principle of logic? *Russell*, 10:125–140, 1990. Repr. in A.D. Irvine (ed.), *Bertrand Russell: Critical Assessments*, vol. 2, Routledge, 1999, pp. 150–264.

Bernard Linsky. The resolution of russell's paradox in principia mathematica. *Philosophical Perspectives*, 16:395–417, 2002.

Bernard Linsky. Ernst schroeder and zermelo's anticipation of russell's paradox. In Karine Fradet and François Lepage, editors, *La crise des fondements : quelle crise?*, pages 7–23. Les Cahiers d'Ithaque, Montréal, 2013.

Edwin Mares. The fact semantics for ramified type theory and the axiom of reducibility. *Notre Dame Journal of Formal Logic*, 48:237–251, 2007.

Christopher Menzel. Cantor and the burali-forti paradox. *Monist*, 67:92–107, 1984.

Robert K. Meyer, Richard Routley, and Michael Dunn. Curry's paradox. *Analysis*, 39:124–128, 1979.

Gregory H. Moore. *Zermelo's Axiom of Choice*. Springer, New York, 1982.

Gregory H. Moore. The roots of russell's paradox. *Russell*, 8:46–56, 1988.

Roman Murawski. On chwistek's philosophy of mathematics. In Nicholas Griffin, Bernard Linsky, and Kenneth Blackwell, editors, *Principia Mathematica at 100*, pages 121–130. Bertrand Russell Research Centre, Hamilton, ON, 2011. Also in Russell (Special Issue), 31(1).

Volker Peckhaus. Paradoxes in göttingen. In Godehard Link, editor, *One Hundred Years of Russell's Paradox*, pages 501–515. Walter de Gruyter, Berlin and New York, 2004.

Graham Priest. *In Contradiction*. Oxford University Press, New York, 2nd edition, 2006.

W. V. O. Quine. New foundations for mathematical logic. *American Mathematical Monthly*, 44:70–80, 1937. Repr. in Quine, *From a Logical Point of View*, Harper & Row, 1953.

W. V. O. Quine. *The Ways of Paradox and Other Essays*. Random House, New York, 1966.

W. V. O. Quine. *Set Theory and Its Logic*. Belknap Press, Harvard, 1967.

Bertrand Russell. Appendix b: The doctrine of types. In *The Principles of Mathematics*, pages 523–528. Cambridge University Press, Cambridge, 1903.

Bertrand Russell. Mathematical logic as based on the theory of types. *American Journal of Mathematics*, 30:222–262, 1908. Repr. in Russell, *Logic and Knowledge*, Allen & Unwin, 1956, pp. 59–102; also in van Heijenoort (ed.), *From Frege to Gödel*, Harvard, 1967, pp. 152–182.

Bertrand Russell. *Introduction to Mathematical Philosophy*. George Allen and Unwin Ltd, London, 1919. Also published in New York by The Macmillan Co.

Bertrand Russell. My mental development. In Paul Arthur Schilpp, editor, *The Philosophy of Bertrand Russell*, pages 3–20. Tudor, New York, 3rd edition, 1951.

Bertrand Russell. *My Philosophical Development*. George Allen and Unwin, London, 1959. Also published by Simon & Schuster, New York.

Bertrand Russell. Letter to frege. In Jean van Heijenoort, editor, *From Frege to Gödel*, pages 124–125. Harvard University Press, Cambridge, Mass., 1967a. Originally 1902.

Bertrand Russell. *The Autobiography of Bertrand Russell*, volume 1. George Allen and Unwin, London, 1967b. 3 vols, 1967–1969; Vols 1 and 2: Boston: Little Brown and Co.; Vol. 3: Simon and Schuster, New York.

Nathan Salmon. A note on kripke's paradox about time and thought. *Journal of Philosophy*, 110:213–220, 2013.

Dana Scott. Axiomatizing set theory. In T. J. Jech, editor, *Proceedings of Symposia in Pure Mathematics*, volume 13, part 2, pages 207–214. American Mathematical Society, 1974.

Stewart Shapiro, editor. *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford University Press, Oxford, 2005.

Keith Simmons. Sets, classes and extensions: A singularity approach to russell's paradox. *Philosophical Studies*, 100:109–149, 2000.

Keith Simmons. A berry and a russell without self-reference. *Philosophical Studies*, 126:

253–261, 2005.

Roy A. Sorensen. Philosophical implications of logical paradoxes. In Dale Jacquette, editor, *A Companion to Philosophical Logic*, pages 131–142. Oxford University Press, New York, 2002.

Roy A. Sorensen. Russell's set. In *A Brief History of the Paradox*, pages 316–332. Oxford University Press, New York, 2003.

Graham Stevens. From russell's paradox to the theory of judgement: Wittgenstein and russell on the unity of the proposition. *Theoria*, 70:28–61, 2004.

Graham Stevens. *The Russellian Origins of Analytical Philosophy*. Routledge, London and New York, 2005.

Jamie Tappenden. The mathematical and logical background to analytic philosophy. In Michael Beaney, editor, *The Oxford Handbook of the History of Analytic Philosophy*, pages 318–354. Oxford University Press, Oxford, 2013.

Alasdair Urquhart. Russell's zig-zag path to the ramified theory of types. *Russell*, 8:82–91, 1988.

Alasdair Urquhart. The theory of types. In Nicholas Griffin, editor, *The Cambridge Companion to Bertrand Russell*, pages 286–309. Cambridge University Press, Cambridge, 2003.

Jean van Heijenoort, editor. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*. Harvard University Press, Cambridge and London, 1967.

John von Neumann. An axiomatization of set theory. In Jean van Heijenoort, editor, *From Frege to Gödel*, pages 393–413. Harvard University Press, Cambridge and London, 1967. Originally 1925.

Russell Wahl. The axiom of reducibility. In Nicholas Griffin, Bernard Linsky, and Kenneth Blackwell, editors, *Principia Mathematica at 100*, pages 45–62. Bertrand Russell Research Centre, 2011. Also in Russell (Special Issue), 31(1).

Zach Weber. Transfinite numbers in paraconsistent set theory. *Review of Symbolic Logic*, 3: 71–92, 2010.

Zach Weber. Transfinite cardinals in paraconsistent set theory. *Review of Symbolic Logic*, 5: 269–293, 2012.

Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*. Cambridge University Press, Cambridge, 1910. 3 vols, 1910–1913; 2nd edn: 1925 (Vol. 1), 1927 (Vols 2–3); abridged as *Principia Mathematica to *56*, Cambridge University Press, 1962.

(V) Paradoxo de Curry[1]

Título Original: Curry's Paradox
Autores: Lionel Shapiro e Jc Beall
Tradução: Pedro Almeida Brandão
Revisão: Kherian Gracher

O "Paradoxo de Curry", como o termo é utilizado por filósofos hoje, refere-se a uma variedade de paradoxos de autorreferência ou circularidade que traçam sua linhagem moderna a Curry (1942b) e Löb (1955).[2] A característica comum desses chamados "paradoxos de Curry" é a maneira como eles exploram uma noção de implicação, implicação lógica ou consequência [3], seja na forma de um conectivo, seja na forma de um predicado. O paradoxo de Curry surge em vários domínios diferentes. Como o Paradoxo de Russell, ele pode assumir a forma de um paradoxo da teoria de conjuntos ou da teoria de propriedades. Mas também pode assumir a forma de um paradoxo semântico, intimamente relacionado ao Paradoxo do Mentiroso. O Paradoxo de Curry difere tanto do Paradoxo de Russell quanto do Paradoxo do Mentiroso pelo fato de não envolver essencialmente a noção de negação. Versões comuns

---

[1] SHAPIRO, Lionel; BEALL, Jc, "Curry's Paradox", In: ZALTA, E. N. (ed.). The Stanford Encyclopedia of Philosophy (Winter 2021 Edition). Stanford, CA: The Metaphysics Research Lab, 2021. Disponível em: https://plato.stanford.edu/archives/win2021/entries/curry-paradox/.
A seguir está a tradução da entrada sobre o Paradoxo de Curry de Lionel Shapiro e Jc Beall na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/win2021/entries/curry-paradox/. Esta versão traduzida pode diferir da versão atual da entrada, que pode ter sido atualizada desde o momento desta tradução. A versão atual está localizada em https://plato.stanford.edu/entries/curry-paradox/. Agradecemos ao Prof. Dr. Edward N. Zalta pela permissão para traduzir e publicar esta entrada.

[2] O termo "Paradoxo de Curry" parece ter origem em Fitch 1952; outras formulações iniciais influentes incluem Moh 1954, Geach 1955 e Prior 1955. Estas serão discutidas na seção 2. Precursores do Paradoxo de Curry também são encontrados no trabalho de lógicos escolásticos tardios e medievais; para referências e discussão, ver Ashworth 1974: 125, Read 2001 e Hanke 2013.

[3] *Implication, entailment, or consequence*, no original. (N. do T.)

baseadas na teoria da verdade envolvem uma sentença que diz de si mesma que, se esta é verdadeira, então uma afirmação arbitrariamente escolhida é verdadeira ou - para usar um exemplo mais sinistro - diz de si mesma que, se esta é verdadeira, então toda falsidade é verdadeira. O paradoxo se dá uma vez que a existência de tal sentença parece implicar a verdade da afirmação arbitrariamente escolhida, ou - no exemplo mais sinistro - de toda falsidade. Neste verbete, mostramos como os vários paradoxos de Curry podem ser construídos, examinamos o espaço de soluções disponíveis e explicamos algumas maneiras pelas quais o paradoxo de Curry é significativo e apresenta desafios característicos.

1. Introdução: Duas Formas do Paradoxo

1.1 Um Argumento Informal

Suponha que seu amigo lhe diga: "Se o que estou dizendo com esta sentença é verdade, então o tempo é infinito". Acontece que existe um argumento curto e aparentemente convincente para a seguinte conclusão:

(P) A mera existência da afirmação de seu amigo implica (ou tem como consequência) que o tempo é infinito.

Muitos consideram que (P) é inaceitável [4] (e, nesse sentido, paradoxal), mesmo que o tempo realmente seja infinito. Ou, se isso não for ruim o suficiente, considere outra versão, desta vez envolvendo uma afirmação sabidamente falsa. Suponha que seu amigo diga, em vez disso: "Se o que estou dizendo com esta sentença é verdade, então todos os números são primos". Agora, *mutatis mutandis*, o mesmo argumento curto e aparentemente convincente leva a (Q):

(Q) A mera existência da afirmação de seu amigo implica (ou tem como consequência) que todos os números são primos.

Eis aqui o argumento para (P). Seja k a sentença autorreferente que seu amigo pronunciou, um pouco simplificada de forma que fique: "Se k é verdadeira, então o tempo é infinito". À luz do que k diz, sabemos o seguinte:

(1) Sob a suposição de que k é verdadeira, segue-se que *se k é verdadeira então o tempo é infinito*.

Mas, claro, também temos:

[4] *Beyond belief* no original. A escolha por "inacreditável" em português poderia trazer ambiguidades. (N. do T.)

(2) Sob a suposição de que k é verdadeira, segue-se que k *é verdadeira*.

Sob a suposição de que k é verdadeira, portanto, derivamos uma condicional junto com seu antecedente. Usando *modus ponens* dentro do escopo da suposição, agora derivamos o consequente da condicional sob essa mesma suposição:

(3) Sob a suposição de que k é verdadeira, é o caso que *o tempo é infinito*.

A regra da prova condicional agora nos permite afirmar uma condicional com nossa suposição como antecedente:

(4) Se k é verdadeiro, então o tempo é infinito.

Mas, já que (4) é apenas k ela mesma, temos então:

(5) k é verdadeira.

Finalmente, juntando (4) e (5) pelo *modus ponens*, obtemos:

(6) O tempo é infinito.

Parece que estabelecemos que o tempo é infinito sem usar suposições além da existência da sentença autorreferente k, junto com os princípios aparentemente óbvios sobre verdade que nos levaram a (1) e também de (4) a (5). O mesmo vale para (Q), uma vez que poderíamos ter usado a mesma forma de argumento para chegar à conclusão falsa de que todos os números são primos.

1.2 Uma Restrição às Teorias

Um desafio imposto pelo Paradoxo de Curry é identificar o que dá errado no argumento informal anterior para (P), (Q) ou algo semelhante. Mas, a começar pela apresentação inicial de Curry em Curry 1942b (consulte o documento suplementar *Curry sobre o Paradoxo de Curry*), a discussão sobre o Paradoxo de Curry geralmente teve um foco diferente. Ela centrou-se em vários sistemas formais — mais frequentemente teorias de conjuntos ou teorias da verdade. Nesse contexto, o que gera o paradoxo é uma prova de que o sistema possui uma característica particular. Tipicamente, a característica em questão é a *trivialidade*. Uma teoria é dita trivial, ou *absolutamente inconsistente*, quando valida toda e qualquer afirmação [5] expressável na linguagem da teoria.[6]

Um argumento que estabelece que uma teoria formal particular é trivial representará um problema se uma das seguintes condições for o caso: (i) desejamos usar a teoria formal em

[5] *Affirms every claim*, no original. (N. do T.)

[6] O artigo de Curry é intitulado "A Inconsistência de Certas Lógicas Formais"[7] (1942b). Por "inconsistência", ele se refere à inconsistência *absoluta*, ou seja, trivialidade.

nossas investigações, como usamos teoria de conjuntos ao fazer matemática, ou (ii) desejamos usar a teoria formal para modelar características da linguagem ou do pensamento, em particular as afirmações às quais alguns falantes ou pensantes[8] estão comprometidos. De ambas as formas, a trivialidade da teoria alvo mostraria que ela é inadequada para o seu propósito pretendido. Então, esse é um segundo desafio imposto pelo paradoxo de Curry.

Para explicar o sentido em que o paradoxo de Curry impõe restrições a teorias, precisamos dizer o que é uma *sentença de Curry*. Informalmente, uma sentença de Curry é uma sentença que é equivalente, à luz de alguma teoria, a uma condicional *com ela mesma como antecedente*. Por exemplo, pode-se pensar no argumento da seção 1.1 como recorrendo a uma teoria informal da verdade. Assim, a sentença "k é verdadeira" serve como uma sentença de Curry para essa teoria. Isso porque, dado o que nossa teoria informal nos diz sobre o que está envolvido na verdade de k, "k é verdadeira" deve ser equivalente a "Se k é verdadeira, então o tempo é infinito" (já que essa condicional é k ela mesma).

Nas seções seguintes, a notação $\vdash_{\mathcal{T}} \alpha$ é usada para dizer que a teoria $\mathcal{T}$ contém a sentença α, e $\Gamma \vdash_{\mathcal{T}} \alpha$ é usado para dizer que α segue das premissas coletadas em Γ de acordo com $\mathcal{T}$ (ou seja, de acordo com $\vdash_{\mathcal{T}}$, a relação de consequência de $\mathcal{T}$).[9] Exceto na seção 4.2.1, no entanto, onde estaremos preocupados apenas com afirmações sobre o que segue de acordo com a teoria a partir de uma única premissa, isto é, afirmações expressas por sentenças da forma $\gamma \vdash_{\mathcal{T}} \alpha$. (O contexto deixará claro onde tal sentença está sendo usada e onde está apenas sendo mencionada.)

Duas sentenças (na linguagem da teoria $\mathcal{T}$) serão chamadas de *intersubstituíveis* de acordo com $\mathcal{T}$ desde que a verdade de qualquer afirmação da forma $\Gamma \vdash_{\mathcal{T}} \alpha$ não seja afetada pela substituição de uma pela outra em α ou em qualquer uma das sentenças de Γ. Finalmente, assumimos que a linguagem contém um conectivo $\rightarrow$ que serve, em algum sentido adequado, como uma condicional. Para os propósitos da definição seguinte, não colocamos nenhum requisito específico sobre o comportamento desta condicional. Agora podemos definir a noção de uma *sentença de Curry para um par sentença-teoria*.

---

[8] *Thinkers*, no original. Embora o termo "pensadores" seja mais comum, "pensantes" expressa melhor a ideia de aquele que é capaz de pensar. (N. do T.)

[9] Tipicamente, a relação de consequência de uma teoria será uma *relação de fechamento*[10], no sentido de que $\Gamma \vdash_{\mathcal{T}} \alpha$ se e somente se α está entre as sentenças em $Cn(\Gamma)$, onde Cn é uma operação de fechamento. Ver entrada sobre lógica proposicional algébrica. No entanto, Ripley (2015b) mostra que as relações de consequência de muitas das teorias que resultam em respostas "livres de contração" ao paradoxo de Curry (ver seção 4.2.1) não são relações de fechamento.

Definição 1 (Sentença de Curry): Seja π uma sentença da linguagem de $\mathcal{T}$. Uma sentença de Curry para π e $\mathcal{T}$ é qualquer sentença κ tal que κ e $\kappa \to \pi$ são intersubstituíveis de acordo com $\mathcal{T}$.[11]

As várias versões do paradoxo de Curry surgem da existência de argumentos em favor da seguinte afirmação muito geral. (Esses argumentos, que se baseiam em suposições sobre a condicional $\to$, serão discutidos em detalhes na seção 3).

Afirmação Problemática[12]: Para toda teoria $\mathcal{T}$ e qualquer sentença π na linguagem de $\mathcal{T}$, se existe uma sentença de Curry para π e $\mathcal{T}$, então $\vdash_{\mathcal{T}} \pi$.

Um argumento que parece estabelecer a Afirmação Problemática será considerado paradoxal desde que também haja razões convincentes para acreditar que essa afirmação é falsa. Um contraexemplo para a Afirmação Problemática seria qualquer teoria $\mathcal{T}$ e sentença π tais que existe uma sentença de Curry para π e $\mathcal{T}$, mas não é o caso que $\vdash_{\mathcal{T}} \pi$.

Como observado acima, o paradoxo de Curry é frequentemente entendido como um desafio à existência de teorias não triviais. Dada a Afirmação Problemática, uma teoria será trivial sempre que uma sentença de Curry puder ser formulada para *qualquer* sentença na linguagem da teoria. De fato, a trivialidade segue de uma condição mais fraca, que a definição a seguir torna explícita.

Definição 2 (Teoria Curry-Completa): Uma teoria $\mathcal{T}$ é Curry-completa se, para cada sentença π na linguagem de $\mathcal{T}$, existe algum π' tal que (i) existe uma sentença de Curry para π' e $\mathcal{T}$ e (ii) se $\vdash_{\mathcal{T}} \pi'$ então $\vdash_{\mathcal{T}} \pi$.

Enquanto uma instância de π' satisfazendo a condição (ii) seria π ela mesma, outra instância seria uma sentença "explosiva" $\perp$ que está contida em uma teoria apenas se *toda*

[11] O termo "sentença de Curry" é por vezes usado em um sentido mais restrito, a saber, para uma sentença que diz de si mesma (apenas) que *se esta é verdadeira*, então p (ou, alternativamente, um absurdo) é verdadeiro (por exemplo, Beall 2009: 33; Zardini 2011: 503). Algumas discussões (por exemplo, Restall 1994: vii; Humberstone 2006) delimitam a noção relevante usando um *bicondicional* em vez de intersubstituibilidade, ou seja, como qualquer κ tal que $\vdash_{\mathcal{T}} \kappa \leftrightarrow (\kappa \to \pi)$. A desvantagem aqui é que o comportamento de $\to$ está atrelado ao de $\leftrightarrow$, embora eles desempenhem papéis muito diferentes na geração de paradoxos. Note que a mera existência de uma sentença de Curry para um par teoria-sentença não constitui um paradoxo, nem é uma objeção à teoria em questão. Por exemplo, $\pi \to \pi$ será uma sentença de Curry para $\pi \to \pi$ e a teoria que consiste nos teoremas da lógica clássica, já que $\pi \to \pi$ será intersubstituível com $(\pi \to \pi) \to (\pi \to \pi)$. Os autores agradecem a Lorenzo Rossi por levantar esta questão.

[12] *Troubling Claim*, no original. (N. do T.)

sentença estiver contida na teoria.[13]

A Afirmação Problemática agora tem uma consequência imediata: uma teoria Curry-completa deve conter toda sentença em sua linguagem.

Corolário Problemático[14]: Toda teoria Curry-completa é trivial.

Novamente, qualquer argumento que pareça estabelecer o Corolário Problemático será considerado paradoxal desde que haja razões convincentes para acreditar que existem teorias não triviais (de fato, verdadeiras) que são Curry-completas.

1.3 Visão Geral

No restante deste verbete, o Paradoxo de Curry será entendido como impondo uma restrição paradoxal às teorias, a saber, aquela enunciada pelo Corolário Problemático acima. Apresentar uma versão do Paradoxo de Curry, entendido dessa forma, envolve fazer duas coisas:

- argumentar que $\mathcal{T}$ é Curry-completa, para alguma teoria alvo $\mathcal{T}$ aparentemente não trivial, e
- fornecer um argumento para a Afirmação Problemática.[15]

As seções 2 e 3 discutem essas duas tarefas nessa ordem. Por enquanto, a ideia básica pode ser transmitida usando o exemplo da sentença autorreferencial k: "Se k é verdadeira, então o tempo é infinito". Primeiramente, dado nosso entendimento de verdade, reconhecemos que a sentença "k é verdadeira" é intersubstituível com "Se k é verdadeira, então o tempo é infinito". Em segundo lugar, o argumento informal da seção 1.1 deriva uma conclusão paradoxal a partir desta equivalência. Leitores interessados principalmente nos princípios lógicos envolvidos nesse argumento e em argumentos relacionados, bem como nas opções para resistir a tais argumentos, podem desejar passar para a seção 3.

[13]Se assumimos que a relação $\vdash_{\mathcal{T}}$ é transitiva, a condição (ii) pode ser substituída pela condição de que $\pi' \vdash_{\mathcal{T}} \pi$. No entanto, uma das respostas ao paradoxo de Curry consideradas abaixo rejeita a transitividade.

[14]*Troubling Corollary,* no original. (N. do T.)

[15]A seção 6 levará em consideração uma família de paradoxos um pouco mais ampla.

2. Construindo Sentenças de Curry

Na forma pela qual é comumente apresentado hoje, o paradoxo de Curry afeta teorias da verdade "ingênuas" (aquelas que apresentam um predicado de verdade "transparente") e teorias de conjuntos "ingênuas" (aquelas que apresentam abstração de conjuntos irrestrita[16]). Esta seção explicará como cada tipo de teoria pode dar origem a sentenças de Curry. Começamos, no entanto, com uma versão que diz respeito a teorias de propriedades, uma versão que se assemelha mais de perto à formulação de Curry. (O documento suplementar *Curry sobre o Paradoxo de Curry* caracteriza brevemente os alvos das versões do paradoxo do próprio Curry.)

Uma teoria de propriedades apresenta *abstração de propriedades irrestrita*, desde que para qualquer condição que possa ser expressa na linguagem da teoria, exista uma propriedade que (de acordo com a teoria) é exemplificada precisamente pelas coisas que satisfazem essa condição. Considere uma teoria $\mathcal{T}_{\mathcal{P}}$ formulada em uma linguagem com um dispositivo de abstração de propriedades $[x : \phi x]$ e uma relação de exemplificação ϵ. Por exemplo, se $\phi(t)$ afirma que o objeto que o termo t representa é triangular, então $t \epsilon [x : \phi x]$ afirma que esse objeto exemplifica a propriedade da triangularidade. Então, dado a abstração de propriedades irrestrita, deveríamos ter o seguinte princípio.

(Propriedade): Para cada sentença aberta ϕ com uma variável livre e cada termo t, as sentenças $t \epsilon [x : \phi x]$ e ϕt são intersubstituíveis de acordo com $\mathcal{T}_{\mathcal{P}}$.

Efetivamente, Curry (1942b) esboça dois "métodos de construção" de sentenças de Curry usando sua contraparte da (Propriedade). Ele diz que o primeiro é "baseado no paradoxo de Russell", enquanto o segundo é "baseado no paradoxo de Epimenides". Embora ambos os métodos sejam baseados na teoria de propriedades, o primeiro método produz um precursor das versões conjuntistas do paradoxo de Curry, enquanto o segundo gera um precursor das versões em teoria da verdade.

2.1 O Primeiro Método e Sentenças de Curry Conjuntistas

A versão do paradoxo de Russell a qual o primeiro método de Curry se assemelha é aquela que diz respeito à exemplificação de propriedades. Seu tema é a propriedade de ser

[16] *Irrestricted set abstraction*, no original. (N. do T.)

tal que *falha-se em exemplificar a si mesmo*[17]. Obtemos uma sentença de Curry em teoria de propriedades considerando, ao invés, a propriedade de ser tal que *exemplifica-se a si mesmo apenas se o tempo for infinito*.[18] Digamos que introduzimos o nome h para essa propriedade, estipulando $h =_{\text{def}} [x : x \epsilon x \rightarrow \pi]$, onde a sentença π afirma que o tempo é infinito. Aplicando o princípio (Propriedade) à sentença $h \epsilon h$, obtemos:

($h \epsilon h$) e ($h \epsilon h \rightarrow \pi$) são intersubstituíveis de acordo com $\mathcal{T}_{\mathcal{P}}$.

Em outras palavras, $h \epsilon h$ é uma sentença de Curry para π e $\mathcal{T}_{\mathcal{P}}$.

O primeiro método de Curry subsequentemente deu origem a sentenças de Curry conjuntistas. Uma teoria de conjuntos apresenta *abstração de conjuntos irrestrita* desde que para qualquer condição expressável na linguagem da teoria, exista um conjunto que (de acordo com a teoria) contém todas e somente as coisas que satisfazem essa condição. Seja $\mathcal{T}_{\mathcal{S}}$ nossa teoria de conjuntos, formulada em uma linguagem que expressa abstração de conjuntos utilizando $\{x : \phi x\}$ e a relação de pertencimento utilizando $\in$. Então, a contraparte da (Propriedade) é:

(Conjunto): Para cada sentença aberta ϕ com uma variável livre e cada termo t, as sentenças $t \in \{x : \phi x\}$ e ϕt são intersubstituíveis de acordo com $\mathcal{T}_{\mathcal{S}}$.

Para obter uma sentença de Curry conjuntista, considere o conjunto consistindo de qualquer coisa que seja membro de si mesma apenas se o tempo for infinito. Chamemos de c este conjunto, estipulando $c =_{\text{def}} [x : x \in x \rightarrow \pi]$. Aplicando o princípio (Conjunto) à sentença $c \in c$, obtemos:

($c \in c$) e ($c \in c \rightarrow \pi$) são intersubstituíveis de cordo com $\mathcal{T}_{\mathcal{S}}$

Em outras palavras, $c \in c$ é uma sentença de Curry para π e $\mathcal{T}_{\mathcal{S}}$

A versão conjuntista do Paradoxo de Curry foi introduzida por Fitch em 1952[19] e também é apresentada em Moh 1954 e Prior 1955.

2.2 O Segundo Método e Sentenças de Curry em Teoria da Verdade

Apesar de seu comentário sobre o "paradoxo de Epimênides", uma forma do paradoxo do Mentiroso, o segundo método de Curry é uma variante de um paradoxo semântico rela-

[17] *Being such that one fails to exemplify oneself*, no original. (N. do T.)

[18] A letra h é escolhida para corresponder ao $\mathfrak{H}$ que Curry 1942b usa para denotar essa propriedade; o mesmo vale para u abaixo.

[19] Fitch considera o paradoxo em suas versões tanto em teoria de propriedades quanto em teorias de conjuntos (Fitch 1952: 89).

cionado: o paradoxo de Grelling.[20] Em sua forma original, o paradoxo de Grelling considera uma propriedade possuída por muitas palavras, a saber, a propriedade que uma palavra tem quando ela *falha em exemplificar a propriedade que representa* (Grelling $\&$ Nelson 1908). Por exemplo, a palavra "ofensividade" possui essa propriedade: ela falha em exemplificar a propriedade que representa, já que não é ofensiva (ver entrada sobre paradoxos e lógica contemporânea). Curry considera, em vez disso, a propriedade que uma palavra tem uma vez que que ela *exemplifica a propriedade que representa apenas se o tempo for infinito*. Agora suponha que nossa teoria introduza um nome u para essa propriedade. Curry então mostra como construir uma sentença que (informalmente) diz que o nome u exemplifica a propriedade que representa. Ele mostra que essa sentença servirá como uma sentença de Curry para uma teoria de propriedades e a denotação de nomes.[21]

Embora este método de obter uma sentença de Curry seja baseado em uma *característica semântica das expressões*, ele ainda se baseia em abstração de propriedades. Não obstante, este pode ser visto como precursor de uma versão inteiramente semântica. (Em vez de considerar a *propriedade* introduzida acima, poder-se-ia considerar o *predicado* "aplica-se a si mesmo apenas se o tempo é infinito".) Dessa forma, como Geach (1955) e Löb (1955) foram os primeiros a mostrar, sentenças de Curry podem ser obtidas usando apenas princípios semânticos, sem recorrer à abstração de propriedades. A abordagem deles corresponde ao argumento informal, na seção 1.1, envolvendo a sentença autorreferencial k que diz: "Se k é verdadeira, então o tempo é infinito".

Para este propósito, seja $\mathcal{T}_{\mathcal{T}}$ uma teoria da verdade onde T é o predicado de verdade. Assuma o princípio da "transparência":

(Verdade): Para cada sentença α, as sentenças $T(\alpha)$ e α são intersubstituíveis de acordo com $\mathcal{T}_{\mathcal{T}}$.

Para obter uma sentença de Curry usando este princípio, suponha que exista uma sentença ξ que seja $T\langle\xi\rangle \rightarrow \pi$.[22] Então segue imediatamente de (Verdade) que:

$T\langle\xi\rangle$ e $T\langle\xi\rangle \rightarrow \pi$ são intersubstituíveis de acordo com $\mathcal{T}_{\mathcal{T}}$.

[20] Isso é mencionado na resenha contemporânea de Church (Church 1942).

[21] Estritamente falando, em sua apresentação, os "nomes" são numerais que denotam números, não propriedades; as propriedades são então códigos numéricos dados.

[22] Geach, na verdade, utiliza uma diagonalização baseada em uma teoria da sintaxe para obter a autorreferência necessária. Ele adiciona que "poderíamos em vez disso usar os conhecidos dispositivos Gödelianos" para simular autorreferência usando diagonalização baseada em uma teoria da aritmética. Ver o verbete sobre os teoremas de incompletude de Gödel.

Em outras palavras, $T\langle\xi\rangle$ é uma sentença de Curry para π e $\mathcal{T}_{\mathcal{T}}$.

Geach observa que o paradoxo semântico que resulta de uma sentença como $T\langle\xi\rangle$ se assemelha "ao Paradoxo de Curry em teoria de conjuntos". Löb, que não menciona o trabalho de Curry, atribui o paradoxo a uma observação de um revisor sobre a prova do que agora é conhecido como teorema de Löb, concernente à provabilidade[23] (ver o verbete sobre os teoremas de incompletude de Gödel). O revisor, que agora sabemos ter sido Leon Henkin (Halbach & Visser 2014: 257), sugeriu que o método usado por Löb em sua prova "leva a uma nova derivação de paradoxos em linguagem natural", a saber, o argumento informal da seção 1.1 acima.[24]

3. Derivando o Paradoxo

Suponha que usamos um dos métodos acima para mostrar, para alguma teoria da verdade, conjuntos ou propriedades, que a teoria é Curry-completa (por virtude de, digamos, conter uma sentença de Curry para cada sentença da linguagem, ou para uma sentença explosiva). Para concluir que a teoria em questão é trivial, basta agora fornecer um argumento para a Afirmação Problemática. Esta afirmação diz que para cada teoria $\mathcal{T}$, se há uma sentença de Curry para π e $\mathcal{T}$, então $\vdash_{\mathcal{T}} \pi$. Tal argumento fará uso das suposições sobre o comportamento lógico do condicional $\rightarrow$ mencionado na Definição 1. Assumindo que a Afirmação Problemática deve ser resistida, isso impõe, consequentemente, restrições ao comportamento deste condicional.

3.1 O Lema do Paradoxo de Curry

Para começar, aqui está um resultado limitativo muito geral, uma variante próxima do Lema que aparece em Curry 1942b.[25]

Lema do Paradoxo de Curry: Suponha que a teoria $\mathcal{T}$ e a sentença π sejam tais que (i) existe uma sentença de Curry para π e $\mathcal{T}$, (ii) todas as instâncias da regra de identidade

[23] *Provability*, no original. (N. do T.)

[24] Sobre a relação entre o Paradoxo de Curry e o teorema de Löb, ver van Benthem (1978), que comenta sobre o "estranho antiparalelo entre a prova de Gödel, na qual um conhecido paradoxo semântico inspirou um resultado formal em termos de provabilidade, em vez de verdade, e a prova de Löb, onde um paradoxo semântico ... foi extraído de um resultado formal sobre provabilidade" (1978: 59).

[25] Não deve ser confundido com um resultado em teoria da prova conhecido como Lema de Curry, seguindo Anderson & Belnap (1975: 136).

(Id) $\alpha \vdash_{\mathcal{T}} \alpha$ são válidas, e (iii) o condicional $\rightarrow$ satisfaz ambos os seguintes princípios:

(MP) Se $\vdash_{\mathcal{T}} \alpha \rightarrow \beta$ e $\vdash_{\mathcal{T}} \alpha$ então $\vdash_{\mathcal{T}} \beta$

(Cont) Se $\alpha \vdash_{\mathcal{T}} \alpha \rightarrow \beta$ então $\vdash_{\mathcal{T}} \alpha \rightarrow \beta$

Então $\vdash_{\mathcal{T}} \pi$.

Aqui MP é uma versão do *modus ponens*, e Cont é um princípio de *contração*: duas ocorrências da sentença α são "contraídas" em uma. (Em breve encontraremos princípios relacionados que são mais comumente referidos como contração.[14]) O Lema do Paradoxo de Curry implica que qualquer teoria Curry-completa deve violar um ou mais princípios dentre Id, MP ou Cont sob pena de trivialidade.

Para provar o Lema, mostra-se que Id, MP e Cont, juntamente com a "Curry-intersubstitutividade" de k e $k \rightarrow \pi$, são suficientes para estabelecer $\vdash_{\mathcal{T}} \pi$. A derivação a seguir se assemelha ao argumento informal da seção 1.1. Esse argumento também incluiu um subargumento para o princípio Cont, que será examinado abaixo.

1. $k \vdash_{\mathcal{T}} k$ Id
2. $k \vdash_{\mathcal{T}} k \rightarrow \pi$ 1 Curry-intersubstitutividade
3. $\vdash_{\mathcal{T}} k \rightarrow \pi$ 2 Cont
4. $\vdash_{\mathcal{T}} k$ 3 Curry-intersubstitutividade
5. $\vdash_{\mathcal{T}} \pi$ 3, 4 MP

A seção 4 discutirá maneiras pelas quais cada um dos dois princípios relativos a $\rightarrow$ assumidos no Lema do Paradoxo de Curry poderia ser justificado ou rejeitado.

3.1.1 Premissas Alternativas

Existem contrapartes do Lema do Paradoxo de Curry que invocam conjuntos alternativos de princípios lógicos (ver, por exemplo, Rogerson & Restall 2004 e Bimbó 2006). Provavelmente a versão mais comum substitui as *regras* Id e Cont pelas *leis* correspondentes:

(IdL) $\vdash_{\mathcal{T}} \alpha \rightarrow \alpha$

(ContL) $\vdash_{\mathcal{T}} (\alpha \rightarrow (\alpha \rightarrow \beta)) \rightarrow (\alpha \rightarrow \beta)$

A derivação agora procede da seguinte forma:

1. $\vdash_{\mathcal{T}} k \to k$ IdL

2. $\vdash_{\mathcal{T}} k \to (k \to \pi)$ 1 Curry-intersubstitutividade

3. $\vdash_{\mathcal{T}} (k \to (k \to \pi)) \to (k \to \pi)$ 2 ContL

4. $\vdash_{\mathcal{T}} k \to \pi$ 2, 3 MP

5. $\vdash_{\mathcal{T}} k$ 4 Curry-intersubstitutividade

6. $\vdash_{\mathcal{T}} \pi$ 4, 5 MP

Uma segunda contraparte comum do Lema do Paradoxo de Curry é devida a Meyer, Routley e Dunn (1979).[26] Ela utiliza dois princípios relativos à conjunção: a forma de lei do *modus ponens* e a idempotência da conjunção.

(MPL) $\vdash_{\mathcal{T}} ((\alpha \to \beta) \wedge \alpha) \to \beta$

(Idem$_\wedge$) As sentenças α e $\alpha \wedge \alpha$ são intersubstituíveis de acordo com T

Desta vez, a derivação procede da seguinte forma:

1. $\vdash_{\mathcal{T}} ((k \to \pi) \wedge k) \to \pi$ MPL

2. $\vdash_{\mathcal{T}} (k \wedge k) \to \pi$ 1 Curry-intersubstitutividade

3. $\vdash_{\mathcal{T}} k \to \pi$ 2 Idem$_\wedge$

4. $\vdash_{\mathcal{T}} k$ 4 Curry-intersubstitutividade

5. $\vdash_{\mathcal{T}} \pi$ 3, 4 MP

Formulando o Lema do Paradoxo de Curry usando Cont, em vez de ContL ou MPL, facilitará chamar atenção (na próxima seção) para diferenças significativas dentro da classe de respostas que rejeitam ambos os últimos princípios.[27]

[26] Também está intimamente relacionado ao método usado na prova original de Löb do teorema de Löb (1955).

[27] Uma terceira contraparte do Lema do Paradoxo de Curry substitui Cont por uma regra que, à luz da Lei de Peirce $\vdash_{\mathcal{T}} ((\alpha \to \beta) \to \alpha)) \to \alpha$, poderia ser chamada de Regra de Peirce:

(RP) Se $\alpha \to \beta \vdash_{\mathcal{T}} \alpha$, então $\vdash_{\mathcal{T}} \alpha$.

A prova do Lema é, então, estruturalmente análoga àquela da seção 3.1. Aqui também, poderia-se em vez disso usar IdL e a Lei de Peirce. Cf. Bunder 1986, Rogerson $\&$ Restall 2004. Ao contrário dos princípios usados até agora, a Lei de Peirce e a Regra de Peirce falham na lógica intuicionista.

4. Respostas ao Paradoxo de Curry

As respostas ao paradoxo de Curry podem ser divididas em duas classes, baseadas em se aceitam ou não o Corolário Problemático, que diz que todas as teorias Curry-completas são triviais.

● **Respostas de Curry-incompletude**[28] aceitam o Corolário Problemático. No entanto, elas negam que as teorias-alvo (de propriedades, conjuntos ou verdade) sejam Curry-completas. Respostas de Curry-incompletude podem adotar a lógica clássica (e geralmente o fazem).

● **Respostas de Curry-completude**[29] rejeitam o Corolário Problemático; elas insistem que podem existir teorias Curry-completas não triviais. Qualquer teoria desse tipo deve violar um ou mais dos princípios lógicos assumidos no Lema do Paradoxo de Curry. Uma vez que a lógica clássica valida esses princípios, essas respostas invocam uma lógica não clássica.[30]

Também existe a opção de defender uma resposta de Curry-incompletude para paradoxos de Curry surgindo em um domínio, digamos, teoria de conjuntos, enquanto se defende uma resposta de Curry-completude para paradoxos de Curry em outro domínio, por exemplo, teoria de propriedades (e.g., Field 2008; Beall 2009).

4.1 Respostas de Curry-Incompletude

A teoria hierárquica de Tarski, a teoria revisionista da verdade (Gupta & Belnap 1993) e as abordagens contextualistas (Burge 1979, Simmons 1993, e Glanzberg 2001, 2004) são exemplos de teorias da verdade tradicionais que fornecem respostas ao paradoxo de Curry baseadas na Curry-incompletude. Todas essas teorias restringem o princípio da transparência "ingênuo" (Verdade). Para uma visão geral, ver verbete sobre o Paradoxo do Mentiroso. No contexto da teoria de conjuntos, as respostas de Curry-incompletude incluem as teorias dos tipos russellianas[31] e diversas teorias que restringem o princípio "ingênuo" de abstração de conjuntos (Conjunto). Ver verbete sobre o Paradoxo de Russell e teorias de conjuntos

[28] *Curry-incompleteness responses*, no original. (N. do T.)
[29] *Curry-completeness responses*, no original. (N. do T.)
[30] No entanto, uma vez que uma resposta rejeita o suficiente princípios lógicos padrão envolvendo $\rightarrow$, pode ser indeterminado se esta conta como uma resposta de Curry-completude ou de Curry-incompletude. Isso porque o conectivo $\rightarrow$ que aparece na definição de uma sentença de Curry (Definição 1) foi estipulado como um *condicional*.
[31] *Russellian type theories*, no original. (N. do T.)

axiomáticas alternativas.

Em geral, as considerações pertinentes para avaliar a maioria das respostas de Curry-incompletude não parecem ser específicas ao Paradoxo de Curry, mas se aplicam igualmente ao paradoxo do Mentiroso (no domínio da teoria da verdade) e ao paradoxo de Russell (nos domínios da teoria de conjuntos e teoria de propriedades).[32] Por essa razão, o restante desta entrada se concentrará nas respostas de Curry-completude, embora a seção 6.3 retorne brevemente à distinção no contexto dos chamados Paradoxos de Curry-validade.[33]

4.2 Respostas de Curry-completude

As respostas de Curry-completude ao paradoxo de Curry sustentam que existem teorias que são Curry-completas, porém não triviais; tais teorias devem violar um ou mais dos princípios lógicos assumidos no Lema do Paradoxo de Curry. Uma vez que a regra Id em geral não foi posta em questão (mas ver French 2016 e Nicolai $\&$ Rossi, *a ser publicado*), isto significou negar que que o condicional $\rightarrow$ de uma teoria Curry-completa não trivial satisfaça ambos MP e Cont. Consequentemente, as respostas foram divididas em duas categorias:

(I) A estratégia mais comum tem sido aceitar que o condicional de tal teoria obedece a MP, mas negar que obedece a Cont. Como Cont é um princípio de contração, tais respostas podem ser chamadas de *livres de contração*[34]. Esta estratégia foi primeiramente proposta por Moh (1954), que é citado com aprovação por Geach (1955) e Prior (1955).

(II) Uma segunda e muito mais recente estratégia consiste em aceitar que o condicional de uma teoria obedece a Cont, mas negar que obedece a MP (às vezes chamado de regra de "desprendimento"). Tais respostas podem ser chamadas de *livres de desprendimento*[35].

[32] As justificativas dadas para negar que uma teoria-alvo $\mathcal{T}$ seja Curry-completa são tipicamente paralelas às justificativas para negar que existe uma sentença λ tal que esta e sua negação $\neg\lambda$ são intersubstituíveis de acordo com $\mathcal{T}$. E os respectivos movimentos terão ramificações filosóficas paralelas. Uma exceção pode ser a resposta inicial ao Paradoxo de Curry por Fitch (1952, 1969). Fitch apresenta várias possíveis restrições às regras de inferência de um sistema de dedução natural que são motivadas especificamente pelo Paradoxo de Curry. Como seu efeito é garantir violações de (Propriedade) e (Conjunto), estas são respostas de Curry-incompletude. Em particular, os sistemas de Fitch permitem uma derivação de $\vdash_{\mathcal{T}_{\mathcal{P}}} h \in h \rightarrow \pi$. No entanto, baseado em características globais dessa derivação, ambas as "restrição especial" e "restrição à não-recorrência" de Fitch bloqueariam a sua extensão a uma derivação de $\vdash_{\mathcal{T}_{\mathcal{P}}} h \in h$. Para uma discussão útil, ver Anderson 1975 e Rogerson 2007.

[33] *Validity Curry paradoxes*, no original. (N. do T.)

[34] *Contraction-free*, no original. (N. do T.)

[35] *Detachment-free*, no original. (N. do T.)

Esta estratégia é defendida, de diferentes maneiras, por Ripley (2013) e Beall (2015).

Cada categoria de respostas de Curry-completude pode, por sua vez, ser subdividida em função da maneira pela qual bloqueia as pretendidas derivações de Cont e MP.

4.2.1 Respostas livres de contração

O princípio Cont, rejeitado pelas respostas livres de contração segue de dois princípios padrões. Tratam-se de *a)* provas condicionais de premissa única e *b)* uma versão um pouco mais geral do *modus ponens*, envolvendo no máximo uma premissa γ:

(MP') Se $\gamma \vdash_{\mathcal{T}} \alpha \to \beta$ e $\gamma \vdash_{\mathcal{T}} \alpha$, então $\gamma \vdash_{\mathcal{T}} \beta$

(PC) Se $\alpha \vdash_{\mathcal{T}} \beta$ então $\vdash_{\mathcal{T}} \alpha \to \beta$

1. $\alpha \vdash_{\mathcal{T}} \alpha \to \beta$
2. $\alpha \vdash_{\mathcal{T}} \alpha$ Id
3. $\alpha \vdash_{\mathcal{T}} \beta$ 1, 2 MP'
4. $\vdash_{\mathcal{T}} \alpha \to \beta$ 3 PC

As respostas livres de contração, portanto, devem rejeitar um ou outro desses dois princípios para o condicional de uma teoria Curry-completa não trivial. Consequentemente, duas subcategorias de teóricos na categoria (I) podem ser identificadas:

(Ia) Uma resposta *livre de contração forte*[36] nega que $\to$ obedece a MP' (por exemplo, Mares $\&$ Paoli 2014; Slaney 1990; Weir 2015; Zardini 2011).

(Ib) Uma resposta *livre de contração fraca*[37] aceita que $\to$ obedece a MP', mas nega que obedece a PC (por exemplo, Field 2008; Beall 2009; Nolan 2016).

A razão pela qual as respostas livres de contração na categoria (Ib) são consideradas "fracas" é que, como os passos 1–3 mostram, elas aceitam o princípio de contração segundo o qual se $\alpha \vdash_{\mathcal{T}} \alpha \to \beta$, então $\alpha \vdash_{\mathcal{T}} \beta$.

Proponentes das respostas livres de contração fortes sustentam que MP' não expressa adequadamente a forma relevante do *modus ponens*. Eles geralmente apresentam sua própria versão dessa regra em um framework "subestrutural", especificamente um que nos permite distinguir entre o que se segue de uma premissa *considerada uma vez* e o que se segue

[36] *Strongly contraction-free response*, no original. (N. do T.)
[37] *Weakly contraction-free response*, no original. (N. do T.)

da mesma premissa *considerada duas vezes*. (Ver verbete sobre lógicas subestruturais). Consequentemente, MP' precisa ser substituído por:

(MP") Se $\gamma \vdash_{\mathcal{T}} \alpha \rightarrow \beta$ e $\gamma \vdash_{\mathcal{T}} \alpha$, então $\gamma, \gamma \vdash_{\mathcal{T}} \beta$

E a regra da "contração estrutural" precisa ser rejeitada:

(sCont) Se $\Gamma, \gamma, \gamma \vdash_{\mathcal{T}} \beta$, então $\Gamma, \gamma \vdash_{\mathcal{T}} \beta$

É porque rejeitam a contração estrutural que as respostas livres de contração fortes podem afirmar que preservam *modus ponens* apesar de rejeitar MP'. (Ver Shapiro 2011, Zardini 2013, e Ripley 2015a).

Respostas livres de contração fortes também precisam bloquear uma derivação de MP' usando um par de princípios envolvendo conjunção:

(MP'$_\wedge$) Se $\gamma \vdash_{\mathcal{T}} \alpha \rightarrow \beta$ e $\delta \vdash_{\mathcal{T}} \alpha$, então $\gamma \wedge \delta \vdash_{\mathcal{T}} \beta$

(Idem$_\wedge$) As sentenças α e $\alpha \wedge \alpha$ são intersubstituíveis de acordo com T

1. $\gamma \vdash_{\mathcal{T}} \alpha \rightarrow \beta$

2. $\gamma \vdash_{\mathcal{T}} \alpha$

3. $\gamma \wedge \gamma \vdash_{\mathcal{T}} \beta$ 1, 2 MP'$_\wedge$

4. $\gamma \vdash_{\mathcal{T}} \beta$ 3 Idem$_\wedge$

Evitar essa derivação de MP' exige negar que exista uma conjunção $\wedge$ que obedeça tanto a MP'$_\wedge$ quanto a Idem$_\wedge$. Segundo muitos proponentes de respostas livres de contração fortes (por exemplo, Mares $\&$ Paoli 2014; Zardini 2011), um tipo de conjunção — o tipo "multiplicativo" ou "de fusão" — obedece a MP'$_\wedge$, mas não a Idem$_\wedge$, enquanto outro tipo — o tipo "aditivo" — obedece a Idem$_\wedge$, mas não a MP'$_\wedge$ (ver verbete sobre lógica linear, bem como Ripley 2015a). Se o framework subestrutural discutido acima for utilizado, a falha de [38] MP'$_\wedge$ se dá pelo fato de que, para a conjunção aditiva, $\gamma, \delta \vdash_{\mathcal{T}} \beta$ não é equivalente a $\gamma \wedge \delta \vdash_{\mathcal{T}} \beta$.

Quanto às respostas livres de contração fracas, a falha de PC foi por vezes motivada usando a "semântica de mundos" do tipo que implica uma distinção entre mundos logicamente *possíveis* e *impossíveis* (por exemplo, Beall 2009; Nolan 2016). Para refutar PC, precisamos da verdade de $\alpha \vdash_{\mathcal{T}} \beta$ e da falsidade de $\vdash_{\mathcal{T}} \alpha \rightarrow \beta$. Nas abordagens de

[38] *The failure of*, no original. (N. do T.)

"mundos" alvo, $\vdash_{\mathcal{T}}$ é definido como preservação de verdade sobre um subconjunto próprio de mundos (em um modelo), a saber, os "mundos possíveis" do modelo. Portanto, $\alpha \vdash_{\mathcal{T}} \beta$ ser verdadeiro significa não haver nenhum mundo possível (em nenhum modelo) no qual α é verdadeiro e β não o é. Por sua vez, para refutar $\vdash_{\mathcal{T}} \alpha \rightarrow \beta$ precisamos de um mundo possível no qual $\alpha \rightarrow \beta$ não seja verdadeiro. Como isso acontece? Porque os conectivos são definidos de uma forma que leva em conta *todos* os (tipos de) mundos no modelo (possíveis e, se houver, impossíveis), há uma opção para $\alpha \rightarrow \beta$ não ser verdadeiro[39] em um mundo possível em virtude de α ser verdadeiro e β não o ser em um mundo *impossível*. E é justamente isso que acontece nas abordagens alvo. (A maneira exata de definir as condições de veradade-em-um-mundo e falsidade-em-um-mundo para o conectivo depende da exata abordagem de "mundos" em questão).

4.2.2 Respostas livres de desprendimento

Respostas livres de desprendimento devem bloquear uma derivação direta de MP baseada em um princípio de transitividade juntamente com a recíproca da prova condicional de premissa única:

(Trans) Se $\alpha \vdash_{\mathcal{T}} \beta$ e $\vdash_{\mathcal{T}} \alpha$, então $\vdash_{\mathcal{T}} \beta$

(RPC)[40] Se $\vdash_{\mathcal{T}} \alpha \rightarrow \beta$, então $\alpha \vdash_{\mathcal{T}} \beta$

1. $\vdash_{\mathcal{T}} \alpha \rightarrow \beta$
2. $\vdash_{\mathcal{T}} \alpha$
3. $\alpha \vdash_{\mathcal{T}} \beta$ 1 RPC
4. $\vdash_{\mathcal{T}} \beta$ 2, 3 Trans

Existem duas subcategorias de teóricos na categoria (II):

(IIa) Uma resposta *livre de desprendimento forte*[41] nega que $\rightarrow$ obedece a RPC (Goodship 1996; Beall 2015).

[39] *Being untrue*, no original. (N. do T.)
[40] Recíproca da prova condicional. (N. do T.)
[41] *Strongly detachment-free response*, no original. (N. do T.)

(IIb) Uma resposta *livre de desprendimento fraca*[42] aceita que $\rightarrow$ obedece a RPC, mas rejeita Trans (Ripley 2013).

A razão pela qual as respostas livres de desprendimento na categoria (IIb) são consideradas “fracas” é que RPC, que essas respostas aceitam, pode ser visto como um tipo de princípio de desprendimento[43] para o condicional.

Uma estratégia para responder à alegação de que as respostas livres de desprendimento são contra-intuitivas tem sido apelar para uma conexão entre consequência e nossa aceitação e rejeição de sentenças. De acordo com essa conexão, sempre que é o caso que $\alpha \vdash_{\mathcal{T}} \beta$, isso significa (ou pelo menos implica) que é incoerente, à luz da teoria $\mathcal{T}$, aceitar α enquanto se rejeita β (ver Restall 2005). Agora suponha que, à luz de uma teoria $\mathcal{T}$, é incoerente rejeitar α e também é incoerente aceitar α enquanto se rejeita β. Então, Ripley (2013) argumenta, não há nada de incoerente, segundo os critérios da teoria, em rejeitar β, desde que também não se aceite α. Há, portanto, espaço para abandonar Trans e adotar uma resposta livre de desprendimento fraca ao paradoxo de Curry. A defesa de Beall da abordagem livre de desprendimento forte baseia-se em considerações relacionadas. Ele argumenta, em suma, que um princípio mais fraco que RPC pode desempenhar o papel relevante na restrição de combinações de aceitação e rejeição de sentenças, incluindo α, β e $\alpha \rightarrow \beta$.

4.2.3 Aplicação ao Argumento Informal

As abordagens para o Paradoxo de Curry acima distinguidas apontam falhas em diferentes inferências e subconclusões do argumento paradoxal informal na seção 1.1. Uma resposta livre de contração forte corresponde ao bloqueio do passo (3) desse argumento, uma vez que rejeita MP’. Uma resposta livre de contração fraca, em vez disso, bloqueia o passo (4), já que rejeita PC. Nenhum tipo de resposta livre de desprendimento aceitará o raciocínio no passo (3). Como aceitam Cont, as respostas livres de desprendimento nos permitem derivar a *conclusão* de (4), da qual as respostas livres de desprendimento fracas permitem ainda derivar a conclusão de (3) por meio de RPC. Contudo, ambas as resposta livre de desprendimento encontram falhas no movimento final, por meio de MP, até (6).

[42] *Weakly detachment-free response*, no original. (N. do T.)
[43] *Detachment principle*, no original. (N. do T.)

5. A Significância do Paradoxo de Curry

Nesta seção, explicamos algumas lições distintivas que podem ser aprendidas ao considerar o Paradoxo de Curry. Para discussões sobre os tipos de significância que as versões do Paradoxo de Curry compartilham com paradoxos relacionados, consulte os verbetes sobre o Paradoxo de Russell e o Paradoxo do Mentiroso.

5.1 Frustrando Esperanças por Soluções de Paradoxos de Negação

Começando com Church (1942), Moh (1954), Geach (1955), Löb (1955) e Prior (1955), discussões sobre o Paradoxo de Curry tem enfatizado o fato de que ele difere do Paradoxo de Russell e do Paradoxo do Mentiroso, no que ele não "envolve essencialmente negação" (Anderson 1975: 128).[44] Uma razão pela qual o status "livre de negação" do paradoxo de Curry é importante é que isso torna o paradoxo resistente a algumas resoluções que poderiam ser adequadas para tais "paradoxos de negação".

Geach argumenta que o paradoxo de Curry representa um problema para qualquer proponente de uma teoria da verdade ingênua ou uma teoria de conjuntos ingênua que, diante de paradoxos de negação,

> ...poderia esperar evitar [esses paradoxos] usando um sistema lógico no qual 'p se e somente se não-p' fosse um teorema para algumas interpretações de 'p' sem que pudéssemos inferir daí qualquer sentença arbitrária... (Geach 1955: 71)

O problema, diz ele, é que o Paradoxo de Curry "não pode ser resolvido apenas adotando um sistema que contém um tipo estranho de negação". Em vez disso, "se queremos manter a visão ingênua de verdade ou a visão ingênua de classes, então devemos modificar as regras de inferência elementares relacionadas a 'se'"(1955: 72). O ponto de vista de Geach

[44] Contrapartes do Lema do Paradoxo de Curry podem ser formuladas utilizando princípios envolvendo negação. Por exemplo, $\neg\alpha \vdash_{\mathcal{T}} \alpha \to \beta$ junto à clássica regra *reductio*

Se $\neg\alpha \vdash_{\mathcal{T}} \alpha$, então $\vdash_{\mathcal{T}} \alpha$

resultam na Regra de Peirce da nota [27]. (É claro que, em uma linguagem cujo único "condicional" $\alpha \to \beta$ é o condicional material $\neg\alpha \vee \beta$, a versão correspondente do Paradoxo de Curry envolverá negação. Nesse contexto, o Paradoxo de Curry apresenta um desafio menos distintivo.

sobre a significância do paradoxo de Curry é ecoada por Meyer, Routley e Dunn (1979: 127). Eles concluem que o Paradoxo de Curry frustra aqueles que "tinham esperança de que enfraquecer os princípios clássicos de negação" resolveria o Paradoxo de Russell.[45]

Em resumo, o ponto é que existem lógicas não-clássicas com princípios de negação fracos que resolvem o Paradoxo de Russell e o Paradoxo do Mentiroso, mas que permanecem vulneráveis ao Paradoxo de Curry. Estas são lógicas com as seguintes características:

(a) Elas podem servir como base para uma teoria não-trivial segundo a qual alguma sentença é intersubstituível com sua própria negação.

(b) Elas não podem servir como base para uma teoria não-trivial que seja Curry-completa.

Apesar de não estar claro quais lógicas Geach tinha em mente, existem de fato lógicas não-clássicas que atendem a essas duas condições. Teorias baseadas nessas lógicas permanecem, consequentemente, vulneráveis ao paradoxo de Curry.

5.1.1 Soluções Paraconsistentes Frustradas

Meyer, Routley e Dunn (1979) chamam atenção para uma classe de lógicas que satisfazem as condições (a) e (b). Elas estão entre as lógicas paraconsistentes, que são lógicas segundo as quais uma sentença juntamente com sua negação não implicará qualquer sentença arbitrária. Lógicas paraconsistentes podem ser usadas para obter teorias que resolvem os paradoxos de Russell e do Mentiroso, ao abraçar a inconsistência da negação sem sucumbir à trivialidade.

De acordo com uma tal teoria $\mathcal{T}$, sentenças λ e $\neg\lambda$ podem ser intersubstituíveis, desde que tenhamos tanto $\vdash_{\mathcal{T}} \lambda$ quanto $\vdash_{\mathcal{T}} \neg\lambda$. Tais teorias são "glutty", no sentido de que afirmam alguma sentença junto com sua negação (ver verbete sobre dialeteísmo). No entanto, várias lógicas paraconsistentes proeminentes não podem servir como base para teorias Curry-completas, sob pena de trivialidade. Diz-se por vezes que tais lógicas falham em ser "Curry paraconsistentes" (Slaney 1989).[47]

[45] Assim como Geach, eles acrescentam que o Paradoxo de Curry mostra que "as propriedades de implicação... bastam para acabar com a gente"[46]. (Por "implicação", eles querem dizer um condicional; ver nota [50].)

[47] Meyer et al. (1979) citam as lógicas de relevância E e R; o mesmo é verdade para a lógica de relevância livre de contração RWX (Slaney 1989). Respostas *glutty* ao paradoxo que *admitem* teorias Curry-completas incluem Priest 2006, Beall 2009 e Beall 2015.

5.1.2 Soluções Paracompletas Frustradas

Muitas das lógicas não-clássicas que foram propostas para fundamentar respostas ao Paradoxo de Russell e ao Paradoxo do Mentiroso são *lógicas paracompletas*, lógicas que rejeitam a lei do terceiro excluído. Essas lógicas possibilitam teorias "gappy". Em particular, onde λ e $\neg\lambda$ são intersubstituíveis de acordo com uma tal teoria $\mathcal{T}$, não será o caso que $\vdash_{\mathcal{T}} \lambda \vee \neg\lambda$. Algumas dessas lógicas paracompletas também atendem às condições (a) e (b).

Um exemplo é a lógica $Ł_3$, baseada nas tabelas de verdade trivalorada de Łukasiewicz (ver, por exemplo, Priest 2008). Uma vez que satisfaz a condição (a), $Ł_3$ oferece uma possível resposta ao Paradoxo de Russell e ao Paradoxo do Mentiroso — em particular, uma resposta "gappy". No entanto, considere o condicional iterado $\alpha \rightarrow (\alpha \rightarrow \beta)$, que abreviamos como $\alpha \Rightarrow \beta$. Suponha que uma sentença de Curry para π e uma teoria $\mathcal{T}$ baseada em $Ł_3$ seja redefinida como sendo[48] qualquer sentença κ intersubstituível com $\kappa \Rightarrow \pi$. Então $\mathcal{T}$ cumprirá todas as condições do Lema do Paradoxo de Curry, como foi inicialmente notado por Moh (1954). Portanto, desde que haja um κ intersubstituível com $\kappa \Rightarrow \pi$ de acordo com $\mathcal{T}$, então $\vdash_{\mathcal{T}} \pi$. Consequentemente, $Ł_3$ não servirá como fundamento de uma resposta ao Paradoxo de Curry.[49]

Resumindo: o Paradoxo de Curry continua sendo um obstáculo para certas abordagens que, em outros casos, serviram como caminhos viáveis para a resolução de paradoxos semânticos por meio de teorias "glutty" ou "gappy". Como resultado, a necessidade de evitar o Paradoxo de Curry desempenhou papel significativo no desenvolvimento de lógicas não-clássicas (por exemplo, Priest 2006; Field 2008).

5.2 Em Direção à uma Estrutura Geral de Paradoxo

O status "livre de negação" do Paradoxo de Curry é importante por uma segunda razão. Prior faz a seguinte importante observação:

[48] *Redefined to be any sentence* κ, no original. (N. do T.)

[49] Ver também Restall 1993. Geach cita o artigo de Moh, mas há poucas razões para pensar que ele tinha $_3$ em mente como uma lógica com a qual se "poderia ... esperar" evitar o Paradoxo do Mentiroso, mas que permanece vulnerável ao Paradoxo de Curry. Isso porque as sentenças de Curry em relação ao condicional primitivo $\rightarrow$, que falha em satisfazer Cont, não representam problema para uma teoria baseada em $_3$. Respostas *gappy* para o paradoxo que admitem teorias Curry-completas não-triviais incluem Kripke 1975, White 1979 e Field 2008.

> Podemos ... dizer não apenas que o Paradoxo de Curry não envolve negação, mas que até mesmo o Paradoxo de Russell pressupõe apenas as propriedades da negação que esta compartilha com a implicação. (Prior 1955: 180)[50]

O que ele tem em mente é que o Paradoxo de Russell e o Paradoxo de Curry podem ser entendidos como resultando da *mesma estrutura geral*, que pode ser instanciada tanto usando negação quanto usando uma condicional.[51]

A estrutura geral pode ser explicitada definindo um tipo de conectivo unário que engendra[52] o Paradoxo de Curry e mostrando como esse tipo é exemplificado tanto pela negação quanto por um conectivo unário definido em termos de uma condicional.

Definição 3: (Conectivo de Curry) Seja π uma sentença na linguagem da teoria $\mathcal{T}$. O conectivo unário $\odot$ é um conectivo de Curry para π e $\mathcal{T}$ se satisfaz dois princípios:

(P1) Se $\vdash_{\mathcal{T}} \alpha$ e $\vdash_{\mathcal{T}} \odot\alpha$, então $\vdash_{\mathcal{T}} \pi$.

(P2) Se $\alpha \vdash_{\mathcal{T}} \odot\alpha$, então $\vdash_{\mathcal{T}} \odot\alpha$

Lema do Paradoxo de Curry Generalizado: Suponha que $\mathcal{T}$ seja tal que Id é válido e que, para algum par de sentenças π e μ, (i) μ e $\odot\mu$ são intersubstituíveis de acordo com $\mathcal{T}$ e (ii) $\odot$ é um conectivo de Curry para π e $\mathcal{T}$. Nesse caso, $\vdash_{\mathcal{T}} \pi$.[53]

Prova:

1. $\mu \vdash_{\mathcal{T}} \mu$ Id

2. $\mu \vdash_{\mathcal{T}} \odot\mu$ 1 Curry-intersubstitutividade

---

[50]Ver também Curry & Feys 1958: 259. Por “implicação”, Prior quer dizer uma condicional. Alguns anos antes, Quine havia notado que o uso de “condicional” no lugar de “implicação” está ”encorajadoramente em ascensão” (1953: 451).

[51]O fato de que o Paradoxo de Curry chama atenção para essa estrutura geral pode ser o que Curry et al. (1972) tinham em mente quando se referem ao Paradoxo de Curry como “o paradoxo de Russell generalizado”. No entanto, o ponto deles é provavelmente diferente, ou seja, que o Paradoxo de Russell se torna um *caso especial* do paradoxo de Curry (em sua forma conjuntista ou em teoria de propriedades), desde que $\neg\alpha$ seja definido como $\alpha \rightarrow \bot$, como é possível nas lógicas intuicionista e clássica.

[52]*Gives rise to*, no original. (N. do T.)

[53]O mesmo resultado é obtido se o requerimento de que $\odot$ seja um conectivo de Curry for substituído pelo requerimento de que seja um conectivo de Curry’, definido substituindo P2 por:
(P2’) Se $\odot\alpha \vdash_{\mathcal{T}} \alpha$, então $\vdash_{\mathcal{T}} \alpha$. (Ver nota 16.)
Para o que são, de fato, várias outras maneiras de generalizar o Lema do Paradoxo de Curry, ver Bimbó 2006 e Humberstone 2006.

3. $\vdash_{\mathcal{T}} \odot\mu$ 2 P2

4. $\vdash_{\mathcal{T}} \mu$ 3 Curry-intersubstitutividade

5. $\vdash_{\mathcal{T}} \pi$ 3, 4 P1

O Lema do Paradoxo de Curry Generalizado pode agora ser instanciado de duas maneiras diferentes, de forma a produzir ou o Paradoxo de Curry, ou um paradoxo de negação.

● Para obter o Paradoxo de Curry, seja o conectivo unário $\odot$ tal que $\odot\alpha$ seja $\alpha \to \pi$ e seja μ uma sentença intersubstituível com $\mu \to \pi$ de acordo com $\mathcal{T}$. Então P1 equivale à instância de MP usada em nossa derivação do Lema do Paradoxo de Curry, enquanto P2 não é nada além de nossa regra Cont.

(MP) Se $\vdash_{\mathcal{T}} \alpha \to \beta$ e $\vdash_{\mathcal{T}} \alpha$ então $\vdash_{\mathcal{T}} \beta$

(Cont) Se $\alpha \vdash_{\mathcal{T}} \alpha \to \beta$ então $\vdash_{\mathcal{T}} \alpha \to \beta$

● Para obter um paradoxo de negação, deixe que $\odot\alpha$ seja $\neg\alpha$ e seja μ uma sentença intersubstituível com $\neg\mu$ de acordo com $\mathcal{T}$.[54] Então P1 equivale a uma instância de *ex contradictione quodlibet* (ou "explosão"), enquanto P2 é um princípio *reductio*.

(ECQ) Se $\vdash_{\mathcal{T}} \alpha$ e $\vdash_{\mathcal{T}} \neg\alpha$ então $\vdash_{\mathcal{T}} \beta$

(Red) Se $\alpha \vdash_{\mathcal{T}} \neg\alpha$ então $\vdash_{\mathcal{T}} \neg\alpha$

O ponto de Prior é que as características da negação que são relevantes para o Paradoxo de Russell ou para o Paradoxo do Mentiroso são *esgotadas pelo seu status como um conectivo de Curry*. Isso deixa claro por que esses paradoxos não dependem de propriedades[55] da negação, como a lei do terceiro excluído ou a eliminação da dupla negação, que não se mantêm[56] em teorias não-clássicas onde a negação permanece um conectivo de Curry (por exemplo, em teorias intuicionistas, onde ECQ e Red ambos se mantêm).[57]

[54]No caso do Paradoxo de Russell na teoria de conjuntos, μ pode ser $\{x : x \notin x\} \in \{x : x \notin x\}$. No caso do Paradoxo do Mentiroso, μ pode ser $T\langle\lambda\rangle$, onde λ é $\neg T\langle\lambda\rangle$.

[55]*Features of negation*, no original. (N. do T.)

[56]*Fail to hold*, no original. (N. do T.)

[57]Ver também Curry & Feys 1958: 259. Eles acrescentam: "Se insiste-se em considerar $[\odot\alpha$, definido como $\alpha \to \pi]$ como uma espécie de negação, então essa negação é uma negação minimal". Como notado abaixo, isso é, estritamente falando, incorreto. A explicação é que a prova de Curry da afirmação correspondente, em Curry 1952, assume que o condicional é caracterizado usando uma extensão do cálculo de sequentes em Curry 1950: 32–33, que inclui a regra de troca estrutural[58]. Isso é o que garante que $\alpha \vdash (\alpha \to \pi) \to \pi$, ou seja, $\alpha \vdash_{\mathcal{T}} \odot\odot\alpha$.

Além disso, um conectivo de Curry não precisa ser muito parecido com uma negação. Ele pode falhar até mesmo em ser uma *negação minimal* (ver verbete sobre negação), já que não precisa obedecer à lei da dupla introdução:

(DI) $\alpha \vdash_{\mathcal{T}} \odot \odot \alpha$.

Por exemplo, suponha que $\odot\alpha$ seja $\alpha \rightarrow \pi$. Então, para que $\odot$ obedeça à DI, teria que ser o caso que $\alpha \vdash_{\mathcal{T}} (\alpha \rightarrow \pi) \rightarrow \pi$. Esse princípio é violado por várias teorias não-clássicas para as quais $\odot$, quando definido desta maneira, *qualifica-se* como um conectivo de Curry.[59]

Resumindo: o Paradoxo de Curry aponta para uma estrutura geral instanciada por uma ampla gama de paradoxos. Esta estrutura ela mesma não envolve negação, mas também é exibida por paradoxos que (ao contrário do Paradoxo de Curry) envolvem essencialmente a negação, como o Paradoxo de Russell e o Paradoxo do Mentiroso.

A questão de saber quais paradoxos exibem uma estrutura comum torna-se importante à luz do "princípio de solução uniforme", influentemente defendido por Priest (1994). Segundo este princípio, paradoxos que pertencem ao "mesmo tipo" devem receber o "mesmo tipo de solução". Suponha que delimitamos um tipo de paradoxo da seguinte forma:

Definição 4: (Paradoxo de Curry Generalizado) Temos um Paradoxo de Curry em todos os casos onde as hipóteses anunciadas no Lema do Paradoxo de Curry Generalizado aparentam ser o caso.[60]

Assumindo que alguém aceita o princípio de solução uniforme, a questão se torna saber o que conta como propor uma solução uniforme para todos os paradoxos de Curry generalizados. Em particular, é suficiente mostrar, para cada instância do tipo assim delimitado, que o que parece ser um conectivo de Curry na verdade não o é? Pareceria que isso deveria ser suficiente. Não é claro por que "uniformidade" deveria adicionalmente exigir que todos os aparentes conectivos de Curry não se qualifiquem como tais *em virtude de violarem a mesma condição*. Por exemplo, suponha que a negação e nosso conectivo unário definido usando $\rightarrow$ ambos aparentam satisfazer o princípio generalizado P2; no primeiro caso, porque $\neg$ parece obedecer a Red e, no último caso, porque $\rightarrow$ parece obedecer a Cont. A menos que essas duas aparências compartilhem uma fonte comum (por exemplo, uma recorrência implícita à contração estrutural, como sustentado por Zardini 2011), não há nada

[59] Estas incluem as lógicas de relevância T e E de Anderson $\&$ Belnap 1975, e muitas das lógicas exploradas em Brady 2006.

[60] *Appear to hold*, no original.

objetavelmente não uniforme em aceitar uma em seu valor nominal enquanto se descarta a outra como enganosa. (Para discussão sobre a questão filosófica aqui, aplicada a uma classe diferente de paradoxos, ver diálogo em Smith 2000 e Priest 2000).

Se este é o caso, o desiderato de que os paradoxos de Curry generalizados sejam resolvidos de forma uniforme não precisa discriminar entre as diferentes soluções logicamente revisionárias que têm sido perseguidas. Estas incluem as seguintes três opções:

● Pode-se sustentar que é apenas o princípio P1 que falha quando $\odot\alpha$ é instanciado como $\neg\alpha$ (para obter um paradoxo de negação), enquanto é P2 sozinho que falha quando $\odot\alpha$ é instanciado como $\alpha \rightarrow \pi$ (para obter um Paradoxo de Curry). Nessa abordagem, ECQ e Cont falham, enquanto Red e MP se mantêm. (Priest 1994, 2006).

● Pode-se sustentar que apenas P2 falha para ambas as instanciações de $\odot$. Nessa abordagem, Red e Cont falham, enquanto ECQ e MP se mantêm (Field 2008; Zardini 2011).

● Pode-se sustentar que apenas P1 falha para ambas as instanciações de $\odot$. Nessa abordagem, ECQ e MP falham, enquanto Red e Cont se mantêm (Beall 2015; Ripley 2013).

Assim, por exemplo, a própria abordagem de Priest seria considerada como resolvendo uniformemente o Paradoxo de Curry e o Paradoxo do Mentiroso *como exemplos do paradoxo de Curry generalizado*. Isso seria o caso apesar do fato de Priest avaliar as sentenças do Mentiroso como tanto verdadeiras quanto falsas, enquanto ele rejeita a afirmação de que as sentenças de Curry são verdadeiras.

Em todo caso, o Paradoxo de Curry apresenta desafios em conexão com a questão de que tipo de uniformidade deveria ser exigida das soluções para vários paradoxos (ver também Zardini 2015). Priest chama atenção para um tipo de paradoxo mais restrito do que os paradoxos de Curry generalizados, um tipo cujas instâncias incluem os paradoxos de negação mas *excluem* o Paradoxo de Curry. Este tipo é identificado pelo "*Inclosure Schema*" de Priest (2002); ver verbete sobre autorreferência. Uma disputa em andamento trata da questão de se poderia haver uma versão do Paradoxo de Curry que se qualifique como um " *inclosure paradox*", mas que resista à solução dialeteísta uniforme de Priest para tais paradoxos (ver diálogo em Beall 2014b, Weber et al. 2014, e Beall 2014a, bem como Pleitz 2015).

6. Curry-validade

Na última década (até a data desta versão do verbete), testemunhou-se um aumento significativo da atenção dada aos paradoxos de Curry, e talvez especialmente aos chamados *Paradoxos de Curry-validade* ou *v-Curry* (Whittle 2004; Shapiro 2011; Beall & Murzi 2013).[61] V-Curry envolve sentenças de Curry que especificamente invocam a relação de consequência ou "validade" de uma teoria, através da utilização seja de uma condicional, seja de um predicado que pretende expressar a relação $\vdash_{\mathcal{T}}$ na linguagem da própria teoria $\mathcal{T}$.

6.1 Forma de Conectivo

Para uma forma do paradoxo de v-Curry, seja o condicional mencionado na definição de uma sentença de Curry (Definição 1) um *conectivo de consequência* $\Rightarrow$. Uma sentença com $\Rightarrow$ como seu operador principal deve ser interpretada da seguinte maneira: "Que p implica (de acordo com $\mathcal{T}$) que q". Agora, imediatamente obtemos versões do Paradoxo de Curry em teoria de propriedades, teoria de conjuntos ou teoria da verdade, desde que $\Rightarrow$ atenda às condições MP e Cont do Lema do Paradoxo de Curry.

O que torna esta instância do Lema do Paradoxo de Curry particularmente problemática é que ela apresenta um obstáculo para uma resposta comum ao Paradoxo de Curry, a saber, a resposta *livre de contração fraca* discutida na seção 4.2.1. Essa resposta dependia da rejeição da regra CP da prova condicional de premissa única, uma direção do "teorema de dedução" de premissa única. Mas esta é uma regra que parece difícil de resistir para um conectivo de consequência (Shapiro 2011; Weber 2014; Zardini 2013). Se β é uma consequência de α de acordo com a relação de consequência da teoria $\mathcal{T}$, onde esta teoria tem $\Rightarrow$ como seu próprio conectivo de consequência, então $\mathcal{T}$ deve certamente conter a afirmação de consequência $\alpha \Rightarrow \beta$. Da mesma forma, esta variedade de Paradoxo de Curry apresenta um obstáculo para respostas *livres de desprendimento*, que requerem a rejeição da regra MP. Se uma teoria com seu próprio conectivo de consequência contém tanto α quanto a condicional de consequência $\alpha \Rightarrow \beta$, então ela deve certamente conter β

[61] Trabalhos mais recentes incluem Bacon 2015, Barrio et al. *a ser publicado*, Cook 2014, Field 2017, Mares & Paoli 2014, Meadows 2014, Murzi 2014, Murzi & Rossi *a ser publicado*, Murzi & Shapiro 2015, Nicolai & Rossi 2017, Rosenblatt 2017, Tajer & Pailos 2017, Priest 2015, Shapiro 2013 & 2015, Wansing & Priest 2015, Weber 2014, Zardini 2013 & 2014.

também. Ou, pelo menos, é o que parece. Certamente, o proponente de uma resposta livre de desprendimento *fraca* argumentará que MP para $\Rightarrow$ ilicitamente incorpora [62] transitividade (ver seção 4.2.2). Ainda assim, o que parece inescapável é a recíproca de PC: a regra RPC, que é a outra direção do teorema de dedução de premissa única. Se uma teoria contém o condicional de consequência $\alpha \Rightarrow \beta$, então certamente β segue de α de acordo com a teoria. Isso ainda excluiria uma resposta livre de desprendimento *forte*.

6.2 Forma de Predicado

Uma segunda forma do paradoxo de v-Curry surge para uma teoria $\mathcal{T}_V$, cujo objeto inclui a relação de consequência de premissa única $\vdash_{\mathcal{T}_V}$, que se verifica, de acordo com essa mesma teoria, entre sentenças em sua linguagem.[63]

Seja essa relação expressa pelo predicado $Val(x, y)$, e suponha ainda que existe uma sentença χ que seja ou $Val(\langle\chi\rangle, \langle\pi\rangle)$, ou ao menos intersubstituível com esta de acordo com $\mathcal{T}_V$. Uma forma do paradoxo de v-Curry emprega dois princípios que governam Val, que chamamos de "desprendimento de validade"[64] e "prova de validade", seguindo Beall $\&$ Murzi (2013).

(DV) Se $\gamma \vdash_{\mathcal{T}_V} Val(\langle\alpha\rangle, \langle\beta\rangle)$ e $\gamma \vdash_{\mathcal{T}_V} \alpha$ então $\gamma \vdash_{\mathcal{T}_V} \beta$

(PV) Se $\alpha \vdash_{\mathcal{T}_V} \beta$ então $\vdash_{\mathcal{T}_V}$ Val$(\langle\alpha\rangle, \langle\beta\rangle)$

Usando estes princípios, obtemos o seguinte breve argumento para $\vdash_{\mathcal{T}_V} \pi$.

1. $\chi \vdash_{\mathcal{T}_V} \chi$ Id
2. $\chi \vdash_{\mathcal{T}_V} Val(\langle\chi\rangle, \langle\pi\rangle)$ Curry-intersubstitutividade
3. $\chi \vdash_{\mathcal{T}_V} \pi$ 1, 2 DV

---

[62] *Build in transitivity*, no original. (N. do T.)

[63] A forma de predicado de v-Curry é discutida em detalhes por Beall $\&$ Murzi (2013); é mencionada por Whittle (2004) e Shapiro (2011). As formas de conectivo e de predicado são equivalentes desde que $Val(\langle\alpha\rangle, \langle\beta\rangle)$ e $\alpha \Rightarrow \beta$ sejam intersubstituíveis de acordo com $\mathcal{T}_V$. Consequentemente, embora a forma de predicado seja um paradoxo *semântico* (no sentido de que concerne a uma característica das expressões que depende de sua interpretação), ela pode ser obtida construindo uma sentença de Curry κ em teoria de propriedades — ou conjuntos — intersubstituível, de acordo com $\mathcal{T}_V$, com $\kappa \Rightarrow \pi$ e garantindo que a última sentença seja, por sua vez, intersubstituível com $Val(\langle\kappa\rangle, \langle\pi\rangle)$.

[64] *Validity detachment*, no original. (N. do T.)

4. $\vdash_{\mathcal{T}_V} Val(\langle\chi\rangle, \langle\pi\rangle)$ 3 PV

5. $\vdash_{\mathcal{T}_V} \chi$ 4 Curry-intersubstitutividade

6. $\vdash_{\mathcal{T}_V} \pi$ 4, 5 DV

Quando aplicada a esta *forma de predicado* de v-Curry, uma resposta livre de contração fraca resistiria à "contração" dada do passo 2 ao passo 4, rejeitando a regra PV, e uma resposta livre de desprendimento rejeitaria DV, mesmo na forma "sem premissas"[65] usada no passo 6. Novamente, no entanto, ambos PV e DV "sem premissas" parecem inescapáveis à luz da interpretação pretendida do predicado Val (Beall $\&$ Murzi 2013; Murzi 2014; Murzi $\&$ Shapiro 2015; Priest 2015; Zardini 2014).[66] Finalmente, mesmo que DV seja rejeitado como envolvendo ilicitamente a transitividade, o que parece inescapável é a recíproca de PV. Se assim for, isso excluiria pelo menos uma resposta livre de desprendimento *forte*.

Uma versão possivelmente mais poderosa do argumento v-Curry[67] é apresentada por Shapiro (2013) e Field (2017: 7). Esse argumento pode tomar tanto a forma de conectivo quanto a de predicado, mas não depende de PC ou PV. Aqui apresentamos a forma predicativa usando Val. Como anteriormente, derivamos primeiro que $\chi \vdash_{\mathcal{T}_V} \pi$ através de DV. À luz do significado de Val, a conclusão de que $\chi \vdash_{\mathcal{T}_V} \pi$ mostra que $Val(\langle\chi\rangle, \langle\pi\rangle)$ é verdadeiro, ou seja, que χ é verdadeiro. Mas se χ é verdadeiro e $\chi \vdash_{\mathcal{T}_V} \pi$, então pareceria que π deve ser também verdadeiro. Dado que respostas livres de desprendimento *fracas* (não-transitivas) a v-Curry permitem a derivação de $\chi \vdash_{\mathcal{T}_V} \pi$, esse argumento também apresenta uma objeção a tais respostas.

6.3 Significância

Se, de fato, os paradoxos v-Curry não são suscetíveis a respostas livres de contração fracas ou respostas livres de desprendimento fortes, então (assumindo que a regra Id é mantida) o espaço para respostas Curry-completas fica restrito a respostas *livres de contração fortes* e respostas *livres de desprendimento fracas*. As primeiras respostas, como explicado na seção 4.2.1, são tipicamente apresentadas reformulando o *modus ponens* (ou desprendimento para o predicado de validade) em um sistema de dedução subestrutural e rejeitando

[65] *Zero-premise form*, no original. (N. do T.)
[66] O status de regras como PV e DV é um tópico controverso: ver Field 2017.
[67] *v-Curry reasoning*, no original. (N. do T.)

a regra de contração estrutural sCont. As últimas, como explicado na seção 4.2.2, rejeitam o princípio estrutural da transitividade. Por essa razão, paradoxos v-Curry são por vezes considerados como motivando relações de consequências *subestruturais* (por exemplo, Barrio et al. *a ser publicado*; Beall & Murzi 2013; Ripley 2015a; Shapiro 2011, 2015).[68]

O debate vívido e abrangente sobre os paradoxos de v-Curry teve como resultado um progresso genuíno em nosso entendimento dos paradoxos de Curry. Ao final, o que ficou claro é que, embora os paradoxos v-Curry possam convidar diferentes resoluções em relação aos paradoxos não-v-Curry, eles permanecem no mesmo molde dos paradoxos de Curry generalizados. Em particular, no modelo geral da seção 5.2, pode-se tomar $\odot$ como expressando consequência (seja como um predicado, seja como conectivo) à luz da relação $\vdash_{\mathcal{T}}$ ela mesma. Isso é o cerne de v-Curry. Na medida em que existem (muitas) diferentes relações de consequência (formais) definíveis em nossa linguagem (por exemplo, consequência lógica em virtude do vocabulário lógico, consequência epistêmica em virtude do vocabulário lógico-e-epistêmico, e assim por diante), existem, portanto, muitos diferentes paradoxos de v-Curry que podem surgir. Ainda assim, o espaço de soluções para estes paradoxos é o espaço de soluções para os paradoxos de Curry generalizados discutidos neste verbete.

No entanto, restam pelo menos duas razões pelas quais os paradoxos de v-Curry merecem atenção separada. Primeiramente, como mencionado anteriormente, duas categorias de soluções Curry-completas — as opções livres de contração fracas e livres de desprendimento fortes — se mostraram especialmente problemáticas no caso dos paradoxos de v-Curry. Em segundo lugar, suponha que alguém trate de um Paradoxo de Curry ordinário (em teoria de propriedades, teoria de conjuntos ou semântico) de maneira Curry-completa. Ainda podem haver motivos para tratar o correspondente paradoxo de v-Curry (conectivo ou predicado) de maneira incompleta, talvez em virtude de olhar para a relação de consequência de uma teoria como estando essencialmente além do alcance de qualquer conectivo ou predicado na linguagem da teoria (ver, por exemplo, Myhill 1975; Whittle 2004). Assim, uma solução "não-uniforme" aos paradoxos de Curry ordinários e suas contrapartes v-Curry pode

[68] De fato, Nicolai & Rossi (*a ser publicado*) argumentam que os paradoxos de v-Curry motivam respostas segundo as quais um terceiro princípio estrutural, a saber, o princípio da reflexividade $\phi \vdash_{\mathcal{T}_V} \phi$ codificado por Id, pode falhar em casos onde ϕ contém um predicado expressando a relação $\vdash_{\mathcal{T}_V}$.

— mais uma vez — ser uma não-uniformidade motivada.[69][70]

7. Apêndice: Curry sobre o Paradoxo de Curry

7.1 Os Sistemas Alvo de Curry

Quando Curry (1942b) introduziu o paradoxo para demonstrar a inconsistência de "certos sistemas de lógica formal", os sistemas que ele tinha em mente eram teorias de aplicação funcional, especificamente o cálculo lambda não tipado de Church e a própria lógica combinatória de Curry (Church 1932; Curry 1930; Seldin 2006). Além de variáveis, a sintaxe de ambos os sistemas consiste apenas em termos que denotam entidades (por exemplo, números, propriedades e proposições). Uma operação binária sobre termos forma um termo que representa o resultado da aplicação de uma entidade à outra. Por exemplo, aplicar a propriedade *ser par* ao número *dois* resulta em uma proposição verdadeira.

A hipótese fundamental de Curry é que os sistemas alvo são "combinatoriamente completos". Isso significa que "qualquer função que possamos definir intuitivamente por meio de uma variável pode ser formalmente representada como uma entidade do sistema" através do uso de um termo denotativo (Curry $\&$ Feys 1958: 5). Por exemplo, existe uma função que, aplicada a alguma entidade x, produz como resultado a aplicação de x a si mesmo.[72] No cálculo lambda, essa função é formalmente representada pela expressão $\lambda x.xx$. O princípio da completude combinatória é a contraparte de Curry do princípio de abstração de propriedades (Propriedade) mencionado na seção 2 do verbete principal.

7.2 A Resposta de Curry

Curry considerava seu paradoxo como uma restrição a uma teoria de aplicação funcional adequada. Ele insistiu que a lição *errada* seria abandonar a abstração de propriedades irrestrita que permite à teoria conter expressões como a que corresponde à h na seção 2.1. Informalmente falando, este termo denota a função que, quando aplicada a um dado

[69] *May be a motivated non-uniformity*, no original. (N. do T.)

[70] A não uniformidade considerada no início da seção 4 se tratava de fornecer diferentes respostas para os paradoxos de Curry ordinários surgidos em diferentes domínios (como teoria de propriedades e teoria de conjuntos). A não uniformidade considerada na seção 5.2 se tratava de diferentes conectivos falhando em se qualificar[71] como conectivos de Curry por diferentes razões.

[72] *Yields the result of*, no original. (N. do T.)

argumento x, produz como resultado a aplicação da função implicação à autoaplicação de x junto a uma proposição arbitrária p.

> A presença desses termos paradoxais é uma vantagem, pois permite que os paradoxos sejam representados no sistema onde é possível analisá-los. (Curry 1942a: 56n14a)

Curry também criticou as tentativas de resolver o paradoxo por meio da adoção de uma lógica não-clássica para o condicional. Ele afirmou que as lógicas disponíveis para desempenhar esse papel eram ou *ad hoc* ou falhavam em ser "adequadas para a matemática" (Curry $\&$ Feys 1958: 261).

Para entender a lição que Curry extraiu de seu paradoxo, deve-se ter em mente que sua formulação não envolve *sentenças* de Curry, mas sim *termos* de Curry. No lugar da sentença $h \epsilon h$ da seção 2.1, ele emprega um termo que denota o resultado da aplicação de uma certa função a si mesma. Segundo ele, o termo em questão é uma expressão denotativa significativa[73], mas falha em "denotar uma proposição" (Curry 1942a: 62). No entanto, os princípios lógicos usados para derivar o paradoxo são ditos ser princípios que se aplicam apenas a proposições.

A abordagem de Curry não está disponível como resposta às versões do paradoxo consideradas no verbete principal. Isso ocorre porque, como agora é padrão, expressões como $h \epsilon h$, $c \in c$ e $T\langle\xi\rangle$ são tratadas como sentenças, e não como nomes que denotam entidades. Apesar disso, uma contraparte de sua abordagem sustenta que essas sentenças, embora significativas, falham em *expressar* proposições. Diversos comentários de Kripke (1975; especialmente pp. 699–700) sugerem que essa é uma maneira de entender sua resposta ao paradoxo em teoria da verdade. Para uma defesa mais recente dessa abordagem, que é um exemplo do que a seção 4 chama de *abordagem de Curry-incompletude*, ver Goldstein 2000.

7.3 A Dívida Professada de Curry para com Carnap

Curry observa que "a ideia central" de sua derivação do paradoxo foi "sugerida por algum trabalho de R. Carnap" (Curry 1942b). Ele cita Carnap 1934, a parte relevante da qual

[73] *Meaningful denoting expression*, no original. (N. do T.)

aparece na tradução como §60a-d da edição expandida em inglês da *Sintaxe Lógica da Linguagem*[74] de Carnap (1937). Essa observação é intrigante, já que a discussão de Carnap não contém nada que se assemelhe à prova de Curry de seu resultado central, a saber, o Lema do qual o Lema do Paradoxo de Curry da seção 3.1 é uma variante. A discussão de Carnap se concentra sobre o paradoxo de propriedades de Russell e no paradoxo semântico de Grelling, as mesmas duas fontes em que Curry se apoia para obter suas sentenças paradoxais (ver seção 2).

No entanto, Curry também direciona o leitor para Hilbert & Bernays 1939 para um "resumo" do trabalho de Carnap que inspirou seu resultado central. Essa referência sugere que ele pode ter tido em mente o precursor do lema diagonal generalizado citado por Carnap em §60c de *Sintaxe Lógica da Linguagem* (e apresentado em §35). Carnap escreve em §60c:

> É possível construir, para toda e cada ... propriedade formulável em S, uma sentença de S, $\mathfrak{S}_1$, tal que $\mathfrak{S}_1$ atribui essa propriedade ... a si mesma.

Se esta é a afirmação que Curry tinha em mente, a "ideia central" seria a aplicabilidade geral do resultado do ponto fixo de Gödel para qualquer propriedade definível, e não apenas para propriedade de *não ser provável*. Isso pode ter sugerido a Curry a possibilidade de um paradoxo que envolve implicação em vez de negação. (Ver os verbetes sobre paradoxos e lógica contemporânea e os teoremas de incompletude de Gödel.)

Referências Bibliográficas

Alan Ross Anderson. Fitch on consistency. In Alan Ross Anderson, Ruth Barcan Marcus, and R. M. Martin, editors, *The Logical Enterprise*, pages 123–141. Yale University Press, New Haven, CT, 1975.

Alan Ross Anderson and Jr. Belnap, Nuel D. *Entailment: the Logic of Relevance and Necessity*, volume 1. Princeton University Press, Princeton, NJ, 1975.

Alan Ross Anderson, Ruth Barcan Marcus, and R. M. Martin, editors. *The Logical Enterprise*. Yale University Press, New Haven, CT, 1975.

E. J. Ashworth. *Language and Logic in the Post-Medieval Period*. Reidel, Dordrecht, 1974.

Andrew Bacon. Paradoxes of logical equivalence and identity. *Topoi*, 34(1):89–98, 2015. doi: 10.1007/s11245-013-9193-8.

[74] Carnap, R., *Logical Syntax of Language*; 1937.

Eduardo Barrio, Lucas Rosenblatt, and Diego Tajer. Capturing naive validity in the cut-free approach. *Synthese*, forthcoming. doi: 10.1007/s11229-016-1199-5. first online 1 September 2016.

Jc Beall. *Spandrels of Truth*. Oxford University Press, Oxford, 2009. doi: 10.1093/acprof:oso/9780199268733.001.0001.

Jc Beall. End of inclosure. *Mind*, 123(491):829–849, 2014a. doi: 10.1093/mind/fzu075.

Jc Beall. Finding tolerance without gluts. *Mind*, 123(491):791–811, 2014b. doi: 10.1093/mind/fzu081.

Jc Beall. Free of detachment: Logic, rationality, and gluts. *Noûs*, 49(2):410–423, 2015. doi: 10.1111/nous.12029.

Jc Beall and Julien Murzi. Two flavors of curry's paradox. *Journal of Philosophy*, 110(3):143–165, 2013. doi: 10.5840/jphil2013110336.

Katalin Bimbó. Curry-type paradoxes. *Logique & Analyse*, 49(195):227–240, 2006.

Ross Brady. *Universal Logic*. CSLI Publications, Stanford, CA, 2006.

M. W. Bunder. Tautologies that, with an unrestricted comprehension axiom, lead to inconsistency or triviality. *Journal of Non-Classical Logic*, 3(2):5–12, 1986.

Tyler Burge. Semantical paradox. *Journal of Philosophy*, 76(4):169–198, 1979. doi: 10.2307/2025724.

Rudolf Carnap. Die antinomien und die unvollständigkeit der mathematik. *Monatshefte für Mathematik*, 41:263–84, 1934.

Rudolf Carnap. *The Logical Syntax of Language*. K. Paul Trench, London, 1937.

Alonzo Church. A set of postulates for the foundation of logic. *Annals of Mathematics*, 33(2):346–366, 1932. doi: 10.2307/1968337.

Alonzo Church. Review: The inconsistency of certain formal logics by haskell b. curry. *Journal of Symbolic Logic*, 7(4):170–71, 1942. doi: 10.2307/2268117.

Roy T. Cook. There is no paradox of logical validity! *Logica Universalis*, 8(3–4):447–467, 2014. doi: 10.1007/s11787-014-0094-4.

Haskell B. Curry. Grundlagen der kombinatorischen logik (teile i & ii). *American Journal of Mathematics*, 52:509–536, 789–834, 1930.

Haskell B. Curry. The combinatory foundations of mathematical logic. *Journal of Symbolic Logic*, 7(2):49–64, 1942a. doi: 10.2307/2266302.

Haskell B. Curry. The inconsistency of certain formal logics. *Journal of Symbolic Logic*, 7(3):115–117, 1942b. doi: 10.2307/2269292.

Haskell B. Curry. *A Theory of Formal Deducibility*. Number 6 in Notre Dame Mathematical Lectures. University of Notre Dame Press, Notre Dame, IN, 1950. [Curry 1950 available online].

Haskell B. Curry. On the definition of negation by a fixed proposition in inferential calculus. *Journal of Symbolic Logic*, 17(2):98–104, 1952. doi: 10.2307/2266240.

Haskell B. Curry and Robert Feys. *Combinatory Logic*, volume 1. North-Holland, Amsterdam, 1958.

Haskell B. Curry, J. Roger Hindley, and Jonathan P. Seldin. *Combinatory Logic*, volume 2 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1972.

Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008. doi: 10.1093/acprof:oso/9780199230747.001.0001.

Hartry Field. Disarming a paradox of validity. *Notre Dame Journal of Formal Logic*, 58(1): 1–19, 2017. doi: 10.1215/00294527-3699865.

Frederic B. Fitch. *Symbolic Logic: An Introduction*. Ronald Press Company, New York, 1952.

Frederic B. Fitch. A method for avoiding the curry paradox. In Nicholas Rescher, editor, *Essays in Honor of Carl. G. Hempel*, pages 255–265. Reidel, Dordrecht, 1969.

Rohan French. Structural reflexivity and the paradoxes of self-reference. *Ergo*, 3(5):113–131, 2016. doi: 10.3998/ergo.12405314.0003.005.

P. T. Geach. On insolubilia. *Analysis*, 15(3):71–72, 1955. doi: 10.1093/analys/15.3.71.

Michael Glanzberg. The liar in context. *Philosophical Studies*, 103(3):217–251, 2001. doi: 10.1023/A:1010314719817.

Michael Glanzberg. A contextual-hierarchical approach to truth and the liar paradox. *Journal of Philosophical Logic*, 33(1):27–88, 2004. doi: 10.1023/B:LOGI.0000019227.09236.f5.

Laurence Goldstein. A unified solution to some paradoxes. *Proceedings of the Aristotelian Society*, 100(1):53–74, 2000. doi: 10.1111/j.0066-7372.2003.00003.x.

Laura Goodship. On dialethism. *Australasian Journal of Philosophy*, 74(1):153–161, 1996. doi: 10.1080/00048409612347131.

Kurt Grelling and Leonard Nelson. Bemerkungen zu den paradoxien von russell und burali-forti. *Abhandlungen der Fries'schen Schule*, 2:301–334, 1908.

Anil Gupta and Nuel Belnap. *The Revision Theory of Truth*. MIT Press, Cambridge, MA, 1993.

Volker Halbach and Albert Visser. The henkin sentence. In Maria Manzano, Ildikó Sain, and Enrique Alonso, editors, *The Life and Work of Leon Henkin*, Studies in Universal Logic, pages 249–264. Springer International, Cham, 2014. doi: 10.1007/978-3-319-09719-0_17.

Miroslav Hanke. Implied-meaning analysis of the currian conditional. *History and Philosophy of Logic*, 34(4):367–380, 2013. doi: 10.1080/01445340.2013.812832.

David Hilbert and Paul Bernays. *Grundlagen der Mathematik*, volume II. Springer, Berlin, 1939.

Lloyd Humberstone. Variations on a theme of curry. *Notre Dame Journal of Formal Logic*, 47 (1):101–131, 2006. doi: 10.1305/ndjfl/1143468315.

Saul A. Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72(19):690–716, 1975. doi: 10.2307/2024634.

M. H. Löb. Solution of a problem of leon henkin. *Journal of Symbolic Logic*, 20(2):115–118, 1955. doi: 10.2307/2266895.

Edwin Mares and Francesco Paoli. Logical consequence and the paradoxes. *Journal of Philosophical Logic*, 43(2–3):439–469, 2014. doi: 10.1007/s10992-013-9268-4.

Toby Meadows. Fixed points for consequence relations. *Logique & Analyse*, 57(227):333–357, 2014.

Robert K. Meyer, Richard Routley, and J. Michael Dunn. Curry's paradox. *Analysis*, 39(3): 124–128, 1979. doi: 10.1093/analys/39.3.124.

Julien Murzi. The inexpressibility of validity. *Analysis*, 74(1):65–81, 2014. doi: 10.1093/analys/ant096.

Julien Murzi and Lorenzo Rossi. Naïve validity. *Synthese*, forthcoming. doi: 10.1007/s11229-017-1541-6. first online 27 September 2017.

Julien Murzi and Lionel Shapiro. Validity and truth-preservation. In Theodora Achourioti, Henri Galinon, José Martínez-Fernández, and Kentaro Fujimoto, editors, *Unifying the Philosophy of Truth*. Springer, Dordrecht, 2015. doi: 10.1007/978-94-017-9673-6_22.

John Myhill. Levels of implication. In Alan Ross Anderson, Ruth Barcan Marcus, and R. M. Martin, editors, *The Logical Enterprise*, pages 179–185. Yale University Press, New Haven, CT, 1975.

Carlo Nicolai and Lorenzo Rossi. Principles for object-linguistic consequence: from logical to irreflexive. *Journal of Philosophical Logic*, forthcoming. doi: 10.1007/s10992-017-9438-x. first online 20 June 2017.

Daniel Nolan. Conditionals and curry. *Philosophical Studies*, 173(10):2629–2649, 2016. doi: 10.1007/s11098-016-0666-7.

Martin Pleitz. Curry's paradox and the inclosure scheme. In Pavel Arazim and Michal Dančák, editors, *Logica Yearbook 2014*. College Publications, London, 2015.

Graham Priest. The structure of the paradoxes of self-reference. *Mind*, 103(409):25–34, 1994. doi: 10.1093/mind/103.409.25.

Graham Priest. On the principle of uniform solution: A reply to smith. *Mind*, 109(433):123–126, 2000. doi: 10.1093/mind/109.433.123.

Graham Priest. *Beyond the Limits of Thought*. Oxford University Press, Oxford, 2002. doi: 10.1093/acprof:oso/9780199254057.001.0001.

Graham Priest. *In Contradiction*. Oxford University Press, Oxford, 2006. doi: 10.1093/acprof:oso/9780199263301.001.0001. Expanded edition (first published 1987).

Graham Priest. *An Introduction to Non-Classical Logic: From If to Is*. Cambridge University Press, Cambridge, second edition, 2008. doi: 10.1017/CBO9780511801174.

Graham Priest. Fusion and confusion. *Topoi*, 34(1):55–61, 2015. doi: 10.1007/s11245-013-9175-x.

A. N. Prior. Curry's paradox and 3-valued logic. *Australasian Journal of Philosophy*, 33(3):177–82, 1955. doi: 10.1080/00048405585200201.

W. V. O. Quine. Mr. strawson on logical theory. *Mind*, 62(248):433–451, 1953. doi: 10.1093/mind/LXII.248.433.

Stephen Read. Self-reference and validity revisited. In Mikko Yrjönsuuri, editor, *Medieval Formal Logic*, pages 183–196. Kluwer Academic Publishers, Dordrecht, 2001. doi: 10.1007/978-94-015-9713-5_7.

Greg Restall. How to be really contraction free. *Studia Logica*, 52(3):381–91, 1993. doi: 10.1007/BF01057653.

Greg Restall. *On Logics Without Contraction*. PhD thesis, The University of Queensland, 1994. Restall 1994 available online.

Greg Restall. Multiple conclusions. In Petr Hájek, Luis Valdés-Villanueva, and Dag Westerståhl, editors, *Logic, Methodology and the Philosophy of Science: Proceedings of the Twelfth International Congress*, pages 189–205. College Publications, London, 2005. Restall 2005 available online.

David Ripley. Paradoxes and failures of cut. *Australasian Journal of Philosophy*, 91:139–164, 2013. doi: 10.1080/00048402.2011.630010.

David Ripley. Comparing substructural theories of truth. *Ergo*, 2(13):299–328, 2015a. doi: 10.3998/ergo.12405314.0002.013.

David Ripley. Contraction and closure. *Thought*, 4(2):131–138, 2015b. doi: 10.1002/tht3.166.

Susan Rogerson. Natural deduction and curry's paradox. *Journal of Philosophical Logic*, 36 (2):155–179, 2007. doi: 10.1007/s10992-006-9032-0.

Susan Rogerson and Greg Restall. Routes to triviality. *Journal of Philosophical Logic*, 33(4): 421–436, 2004. doi: 10.1023/B:LOGI.0000036853.44128.8f.

Lucas Rosenblatt. Naive validity, internalization, and substructural approaches to paradox. *Ergo*, 4(4):93–120, 2017. doi: 10.3998/ergo.12405314.0004.004.

Jonathan P. Seldin. The logic of curry and church. In Dov M. Gabbay and John Woods, editors, *Handbook of the History of Logic, Volume 5: Logic from Russell to Church*, pages 819–873. Elsevier, Amsterdam, 2006.

Lionel Shapiro. Deflating logical consequence. *The Philosophical Quarterly*, 61(243):320–42, 2011. doi: 10.1111/j.1467-9213.2010.678.x.

Lionel Shapiro. Validity curry strengthened. *Thought*, 2:100–107, 2013. doi: 10.1002/tht3.80.

Lionel Shapiro. Naive structure, contraction and paradox. *Topoi*, 34(1):75–87, 2015. doi: 10.1007/s11245-014-9235-x.

Moh Shaw-Kwei. Logical paradoxes for many-valued systems. *Journal of Symbolic Logic*, 19(1):37–40, 1954. doi: 10.2307/2267648.

Keith Simmons. *Universality and the Liar: An Essay on Truth and the Diagonal Argument*. Cambridge University Press, Cambridge, 1993.

John Slaney. Rwx in not curry paraconsistent. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic: Essays on the Inconsistent*, pages 472–480. Philosophia, Munich, 1989.

John Slaney. A general logic. *Australasian Journal of Philosophy*, 68(1):74–88, 1990. doi: 10.1080/00048409012340183.

Nicholas J. J. Smith. The principle of uniform solution (of the paradoxes of self-reference). *Mind*, 109(433):117–122, 2000. doi: 10.1093/mind/109.433.117.

Diego Tajer and Federico Pailos. Validity in a dialetheist framework. *Logique & Analyse*, 60 (238):191–202, 2017.

Johan van Benthem. Four paradoxes. *Journal of Philosophical Logic*, 7(1):49–72, 1978. doi: 10.1007/BF00245920.

Heinrich Wansing and Graham Priest. External curries. *Journal of Philosophical Logic*, 44 (4):453–471, 2015. doi: 10.1007/s10992-014-9336-4.

Zach Weber. Naïve validity. *The Philosophical Quarterly*, 64(254):99–114, 2014. doi: 10.

1093/pq/pqt016.

Zach Weber, David Ripley, Graham Priest, Dominic Hyde, and Mark Colyvan. Tolerating gluts. *Mind*, 123(491):813–828, 2014. doi: 10.1093/mind/fzu057.

Alan Weir. A robust non-transitive logic. *Topoi*, 34(1):99–107, 2015. doi: 10.1007/s11245-013-9176-9.

Richard B. White. The consistency of the axiom of comprehension in the infinite-valued predicate logic of łukasiewicz. *Journal of Philosophical Logic*, 8(1):509–534, 1979. doi: 10.1007/BF00258447.

Bruno Whittle. Dialetheism, logical consequence and hierarchy. *Analysis*, 64:318–26, 2004. doi: 10.1093/analys/64.4.318.

Elia Zardini. Truth without contra(di)ction. *Review of Symbolic Logic*, 4(4):498–535, 2011. doi: 10.1017/S1755020311000177.

Elia Zardini. Naive modus ponens. *Journal of Philosophical Logic*, 42(4):575–593, 2013. doi: 10.1007/s10992-012-9239-1.

Elia Zardini. Naive truth and naive logical properties. *Review of Symbolic Logic*, 7(2):351–384, 2014. doi: 10.1017/S1755020314000045.

Elia Zardini. Getting one for two, or the contractors' bad deal. towards a unified solution to the semantic paradoxes. In Theodora Achourioti, Henri Galinon, José Martínez-Fernández, and Kentaro Fujimoto, editors, *Unifying the Philosophy of Truth*. Springer, Dordrecht, 2015. doi: 10.1007/978-94-017-9673-6_23.

Sobre o Organizador e Tradutores

Organizador

Kherian Galvão Cesar Gracher é Bacharel em Filosofia (2008-12) pela Universidade Federal de Ouro Preto (UFOP); Mestre em Filosofia (2014-16), sob orientação do Prof. Dr. Décio Krause; e Doutor em Filosofia (2016-20), também sob orientação do Prof. Dr. Décio Krause e sob coorientação do Prof. Dr. Newton C. A. da Costa, pelo Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Catarina (PPGFIL-UFSC). Atualmente, é Professor Colaborador e desenvolve pesquisa de Pós-Doutorado no Programa de Pós-Graduação em Filosofia da Universidade Federal do Rio de Janeiro (PPGF-UFRJ), sob supervisão do Prof. Dr. Jean-Yves Béziau e financiado pela Fundação *Carlos Chagas Filho* de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).[1]

[1] Programa de Pós-Doutorado *Nota 10* (PDR10) – Processos: E-26/200.129/2022 e E-26/200.130/2022; Matrícula: 2021.04772.0

Tradutores

Caio Cezar Silva é doutorando pela Universidade Federal do Rio de Janeiro (PPGLM-UFRJ). Mestre em Filosofia pela mesma instituição e programa. Tem experiência na área de Filosofia, com ênfase em Causalidade, Disposições e Persistência. Foi bolsista CAPES, sob orientação do prof. Dr. Guido Imaguire. Atualmente é bolsista FAPERJ Nota 10, sob orientação do mesmo professor.

Benedito Monteiro cursa doutorado em Filosofia no Programa de Pós-graduação em Lógica e Metafísica (PPGLM) da Universidade Federal do Rio de Janeiro (UFRJ), com parte do programa realizado no exterior, na University of South Florida, sob a supervisão do Prof. Stephen Turner. Possui mestrado em Filosofia pelo Programa de Pós-graduação em Filosofia (PPGFIL) da Universidade Federal do Pará (UFPA). Realizou especialização em Saberes, Linguagens e Práticas Educacionais na Amazônia pelo Instituto Federal de Educação, Ciência e Tecnologia do Pará (IFPA). Concluiu a graduação em Filosofia pela Universidade do Estado do Pará (UEPA) e graduação em Letras-Português pela Universidade Metodista de São Paulo (UMESP).

Annelyze Reis possui licenciatura e bacharelado em História pela Pontifícia Universidade Católica do Rio de Janeiro (2018), com domínio adicional em Cultura Greco Latina, e mestrado em Filosofia também pela Pontifícia Universidade Católica do Rio de Janeiro (2021), defendido na linha de História da Filosofia Antiga. Desenvolve pesquisa sobre as relações entre os modos de apreensão do conhecimento e a metafísica aristotélica.

Paloma de Souza Xavier cursa doutorado em Filosofia na Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), com bolsa CAPES, sob orientação do prof. Ludovic Soutif (PUC-Rio) e coorientação da profa. Célia Teixeira (UFRJ). Possui mestrado em Filosofia pela Universidade Federal de Pernambuco (UFPE), também com bolsa CAPES, sob orientação do prof. Marcos Silva (UFPE). Concluiu a graduação em Filosofia pela mesma universidade. Sua pesquisa atual investiga os desdobramentos da noção de *deep disagreements* (desacordos profundos) no campo da epistemologia social, com ênfase na teoria da injustiça epistêmica.

Pedro Almeida Brandão é doutorando em Filosofia pela Universidade de Viena. Possui mestrado e graduação em Filosofia pela Universidade Federal do Rio de Janeiro. Tem interesse em Filosofia da Ciência, Metafísica Moderna e Matemática.

Editora
UFPel

DISSERTATIO
FILOSOFIA