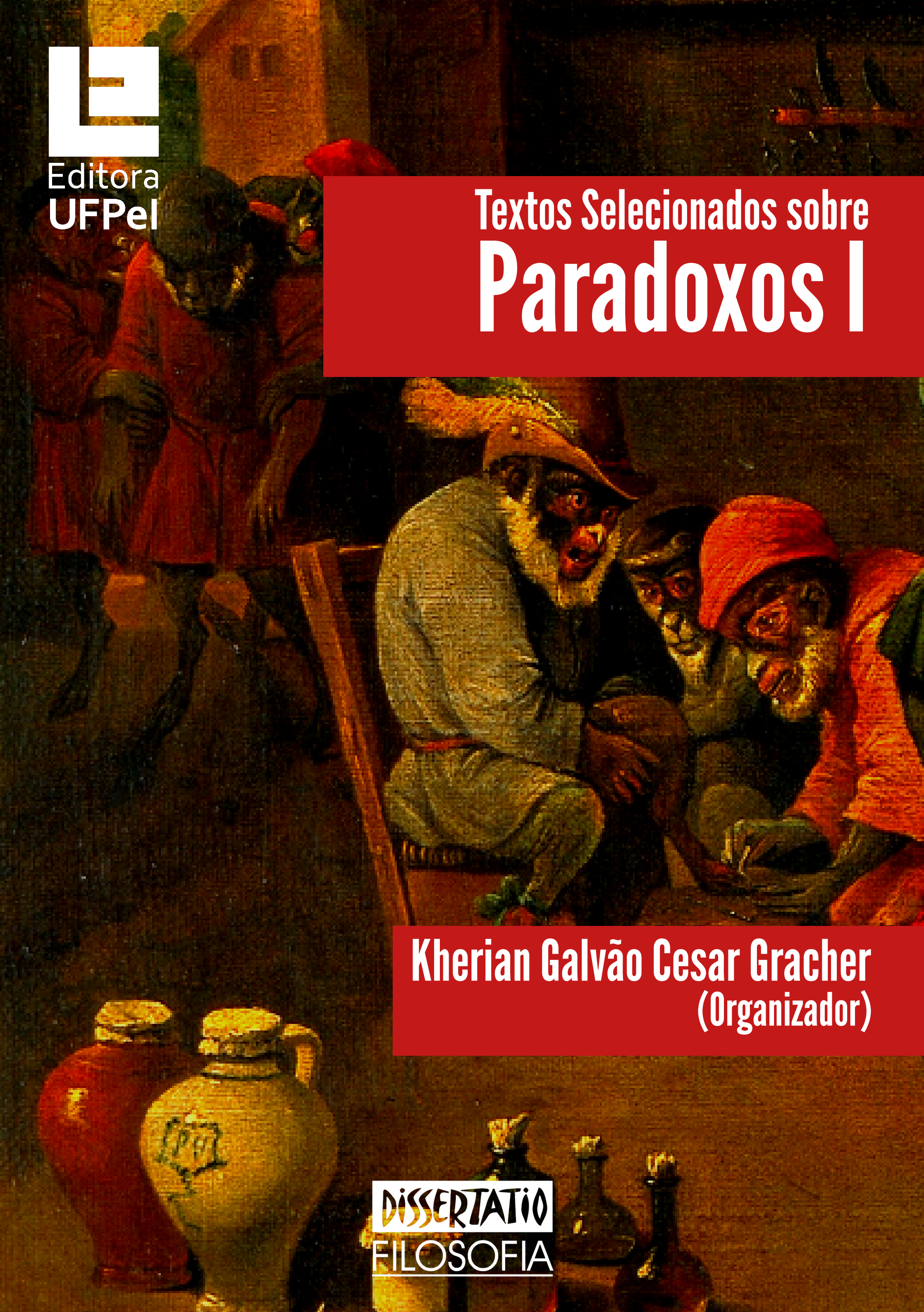

Editora UFPel

Textos Selecionados sobre Paradoxos I

Kherian Galvão Cesar Gracher
(Organizador)

DISSERTATIO FILOSOFIA

TEXTOS SELECIONADOS SOBRE PARADOXOS I

SÉRIE INVESTIGAÇÃO FILOSÓFICA

TEXTOS SELECIONADOS SOBRE PARADOXOS I

Kherian Galvão Cesar Gracher

Pelotas, 2025.

Editora UFPel
Chefia:
Ana da Rosa Bandeira | EDITORA-CHEFE
Seção de Pré-produção:
Isabel Cochrane | ADMINISTRATIVO
Suelen Aires Böettge | ADMINISTRATIVO
Seção de Produção:
Eliana Peter Braz | PREPARAÇÃO DE ORIGINAIS
Marisa Helena Gonsalves de Moura | CATALOGAÇÃO
Anelise Heidrich | REVISÃO
Suelen Aires Böettge | ADMINISTRATIVO
Fernanda Figueredo Alves | PROJETO GRÁFICO E DIAGRAMAÇÃO
Carolina Abukawa (Bolsista) | PROJETO GRÁFICO E DIAGRAMAÇÃO
Angélica Knuth (Bolsista) | PROJETO GRÁFICO E DIAGRAMAÇÃO
Seção de Pós-produção:
Madelon Schimmelpfennig Lopes | ADMINISTRATIVO
Eliana Peter Braz | ADMINISTRATIVO

NEPFIL online

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 I. [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: 978-65-998644-6-9

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

A Newton da Costa
ao mostrar que o impossível também pode ser verdadeiro,
ampliou o mundo – e o contradisse com elegância.

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

Paradoxos:
ou o que podemos aprender pensando absurdos

Paradoxos são fenômenos por vezes engraçados, por vezes difíceis de se compreender. Contudo, há algo que irremediavelmente prende nossa atenção a eles, tornando-os ferramentas fascinantes para testar e desafiar nossas intuições. Uma conclusão que soa absurda, contraditória ou impossível não é apenas motivo de espanto: é também um convite para pensar com mais rigor sobre as ideias que sustentam nosso raciocínio.

Não é à toa que, desde os primeiros debates da filosofia grega, os paradoxos ocupam lugar central na investigação filosófica. Zenão, por exemplo, elaborou argumentos para mostrar que o movimento não poderia existir – uma provocação que desconcertou filósofos de sua época e ainda hoje ecoa nas discussões sobre continuidade e infinito. Essa mistura de surpresa e perplexidade é a marca registrada dos paradoxos. Como observa Sainsbury:

> "um paradoxo é uma conclusão aparentemente inaceitável derivada por meios aparentemente aceitáveis a partir de premissas aparentemente aceitáveis."[1]

Mas qual é, afinal, a *natureza* de um paradoxo? Devemos compreendê-lo como um problema fundamentalmente epistêmico, lógico ou metafísico? Há boas razões para considerar cada uma dessas perspectivas. Muitos autores caracterizam os paradoxos como conjuntos de proposições, cada qual **epistemicamente** plausível isoladamente, mas inconsistentes quando tomadas em conjunto. Lycan formula essa ideia de modo direto:

> "um paradoxo é um conjunto inconsistente de proposições, cada uma das quais é muito plausível."[2]

Schiffer reforça essa caracterização:

---

[1]SAINSBURY, R. M. *Paradoxes*. 3. ed. Cambridge: Cambridge University Press, 2009, p. 26. ISBN 9780521720798.

[2]LYCAN, William G. "What, exactly, is a paradox?". *Analysis*, v. 70, n. 4, p. 615–622, 2010, p. 621. DOI: 10.1093/analys/anq069.

> “um paradoxo filosófico é um conjunto de proposições aparentemente mutuamente incompatíveis, cada uma das quais goza de algum grau significativo de plausibilidade quando considerada isoladamente.”[3]

Na mesma linha, Sorensen define os paradoxos como conjuntos de sentenças inconsistentes em conjunto, mas críveis quando tomadas separadamente:

> “um paradoxo é um conjunto de sentenças que são inconsistentes em conjunto, mas credíveis quando consideradas separadamente.”[4]

Rescher apresenta formulação semelhante:

> “um paradoxo surge quando um conjunto de proposições individualmente plausíveis é coletivamente inconsistente.”[5]

Em uma leitura **epistêmica**, portanto, o paradoxo surge do atrito no modo como organizamos nossas crenças e inferências. Ele surge quando princípios que parecem inofensivos – como aceitar o fechamento do conhecimento, confiar na consistência de nossas crenças ou aplicar regras de previsão – entram em colisão entre si, gerando inconsistência. O valor desses paradoxos está em mostrar que nem sempre podemos manter todas as nossas intuições epistêmicas ao mesmo tempo: em algum ponto, é preciso revisar, abandonar ou qualificar certos compromissos que antes pareciam inquestionáveis.

Por outro lado, há paradoxos cujo impacto é sobretudo **lógico-matemático**. Os de Zenão e Russell, por exemplo, revelam não apenas limites de intuições humanas, mas verdadeiras quebras nas estruturas formais da lógica e da matemática. Margaret Cuonzo expressa esse ponto da seguinte forma:

> “Argumentos são pedaços de raciocínio nos quais uma afirmação (a conclusão) é sustentada por outras afirmações (as premissas). Quando o raciocínio é correto, premissas verdadeiras sempre levarão a conclusões verdadeiras.

---

[3]SCHIFFER, Stephen. *The Things We Mean*. Oxford: Oxford University Press, 2003, p. 68. ISBN 0199241273.
[4]SORENSEN, Roy. *Vagueness and Contradiction*. Oxford: Oxford University Press, 2003, p. 364. ISBN 0199255854.
[5]RESCHER, Nicholas. *Paradoxes: Their Roots, Range, and Resolution*. Chicago: Open Court, 2001, p. 6. ISBN 0812694376.

Mas, no caso dos paradoxos, parece que algo deu errado, pois premissas verdadeiras e raciocínio correto levam a uma conclusão obviamente falsa ou contraditória."[6]

Ou seja, aqui o problema não está (apenas) na plausibilidade epistêmica das premissas, mas na própria *arquitetura formal* que as conecta: certos esquemas inferenciais e princípios de fundo – como o uso irrestrito de compreensão em teoria de conjuntos, a autorreferência aliada ao esquema-T para verdade, ou a idealização do contínuo em argumentos sobre movimento –, quando combinados, produzem incompatibilidades. Nesses casos, a "quebra" exige intervir nos mecanismos lógicos ou matemáticos: restringir axiomas (*e.g.*, compreensão), distinguir rigorosamente linguagem-objeto e metalinguagem, revisar regras inferenciais (ou mesmo considerar lógicas não clássicas), ou refinar as noções técnicas de infinito, continuidade e modelo.

Há ainda quem veja nos paradoxos tensões de ordem **metafísica**, pois eles expõem os limites de conceitos fundamentais como movimento, tempo, identidade, vaguidade ou verdade. Barwise e Etchemendy resumem bem essa perspectiva:

> "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."[7]

Ou seja, quando encaramos os paradoxos em chave metafísica, o ponto central não está em falhas de inferência ou em sistemas formais insuficientes, mas na própria constituição dos conceitos fundamentais com que estruturamos a realidade. O paradoxo atua como um espelho que evidencia tensões em noções como *movimento*, *tempo*, *identidade* ou *verdade*, mostrando que, levadas aos seus limites, tais noções podem gerar resultados incompatíveis ou contraditórios. Nesses casos, não basta corrigir regras lógicas ou revisar axiomas: é preciso repensar os próprios conceitos que organizam nossa visão do mundo. É nesse

[6]CUONZO, Margaret. *Paradox*. Cambridge, MA: MIT Press, 2014, p. 7. (Essential Knowledge Series). DOI: 10.7551/mitpress/9904.001.0001.

[7]BARWISE, Jon; ETCHEMENDY, John. *The Liar: An Essay on Truth and Circularity*. New York; Oxford: Oxford University Press, 1987, pp. 4–5. ISBN 0195059441.

sentido que os paradoxos metafísicos funcionam como diagnósticos de falhas conceituais profundas, revelando onde a ontologia implícita de nossas teorias se rompe.

Essa pluralidade de enfoques repercute também nas tentativas de *classificar* paradoxos. Doris Olin distingue, por exemplo, paradoxos do **Tipo I**, em que um único argumento leva a uma conclusão absurda (como Aquiles e a tartaruga), e paradoxos do **Tipo II**, que envolvem dois argumentos simétricos com conclusões incompatíveis (como certas formulações do Mentiroso):

> "um paradoxo do tipo I, como Aquiles e a tartaruga, tem um argumento e uma conclusão; um paradoxo do tipo II, como o Navio de Teseu, envolve dois argumentos e duas conclusões."[8]

Outros autores reservam o termo *antinomia* para designar paradoxos que desembocam diretamente em uma contradição lógica. Pleitz propõe uma definição explícita:

> "(Def. Antinomia) Uma antinomia é um paradoxo com uma conclusão contraditória."[9]

Cook, por sua vez, chama a atenção para o fato de que a distinção entre paradoxo e antinomia não é absoluta:

> "Dada qualquer antinomia, podemos construir um paradoxo simplesmente combinando os argumentos em um único argumento. [...] Da mesma forma, podemos transformar qualquer paradoxo em uma antinomia. Como resultado, restringiremos nossa atenção aos paradoxos."[10]

Assim, ao longo deste volume, adotaremos uma concepção ampla: paradoxos são *raciocínios que tensionam os limites de nossas intuições e de nossas teorias, seja pela via da contradição, seja pela via da conclusão absurda*. Armour-Garb justifica o interesse filosófico dessa postura:

---

[8] OLIN, Doris. *Paradox*. Montreal: McGill-Queen's University Press, 2003, p. 7. ISBN 0773525847.
[9] PLEITZ, Martin. *Logic, Language, and the Liar Paradox*. Paderborn: Mentis, 2018, p. 18. ISBN 3957438497.
[10] COOK, Roy T. *Paradoxes*. Cambridge: Polity Press, 2013, pp. 15–16. ISBN 9780745665511.

> “estudamos os paradoxos porque eles revelam certas suposições que fizemos, que precisam ser questionadas, e nos forçam a repensar alguns dos compromissos mais centrais que possuímos.”[11]

É justamente essa pluralidade que organiza o presente volume. Cada capítulo explora um tipo distinto de paradoxo, revelando como problemas epistemológicos, linguísticos, históricos e lógico-matemáticos podem ser iluminados por tais raciocínios enigmáticos.

No capítulo (I) (p. 21–50), encontramos o **Paradoxo da Cognoscibilidade**, mais conhecido como Paradoxo de Fitch. A ideia inicial é simples: se toda verdade pode ser conhecida, então não deveria haver verdades permanentemente inacessíveis. O argumento, contudo, mostra que essa suposição leva a uma conclusão surpreendente: se toda verdade é cognoscível, então *todas as verdades já são conhecidas*. O paradoxo toca diretamente em correntes filosóficas que vinculam verdade e possibilidade de conhecimento, como o verificacionismo, o antirrealismo semântico e o construtivismo matemático. Ao longo do capítulo, o leitor encontrará não apenas a formulação original do paradoxo, mas também as principais estratégias de resposta: restringir a tese de cognoscibilidade, rever os princípios lógicos usados na derivação ou reconsiderar a própria noção de verdade em jogo. Trata-se de um exemplo marcante de como um raciocínio aparentemente trivial pode abalar teorias amplamente aceitas sobre o conhecimento.

No capítulo (II) (p. 51–88), tratamos do **Paradoxo Sorites**, nascido da vagueza em predicados como “monte” ou “careca”. A engrenagem é simples e traiçoeira: partindo do *princípio de tolerância* (“se n grãos formam um monte, então $n-1$ também”), encadeamos passos tão pequenos que parecemos poder descer de $10,000$ grãos até 1 sem perder o estatuto de “monte”. O capítulo apresenta as duas formulações canônicas (indutiva e condicional) e explicita a tensão entre tolerância e a ausência de fronteiras nítidas. Em seguida, mapeia as principais respostas: (i) *linguagens ideais*/regimentação, que tornam os termos precisos ao custo de empobrecer o uso ordinário; (ii) *epistemicismo*, que postula um corte objetivo porém incognoscível; (iii) *abordagens semânticas* com lacunas de verdade (supervalorações) ou graus/medidas; e (iv) estratégias de *aceitação parcial* do sorites sob restrições. Por fim, o texto conecta o Sorites ao *Mentiroso* (via esquemas unificadores) e extrai lições filosóficas sobre *significado como uso*, *verdade e o esquema T* e *imperscrutabilidade da referência*, preparando terreno para debates que atravessam semântica, lógica e metafísica.

[11] ARMOUR-GARB, Bradley. “Introduction”. In: ARMOUR-GARB, Bradley (ed.). *Reflections on the Liar*. Oxford: Oxford University Press, 2017, p. 4. ISBN 9780199896042.

No capítulo (III) (p. 89–127), o tema são os **paradoxos epistêmicos**, que desafiam nossa compreensão sobre o que pode ser sabido ou previsto. O primeiro é o *Paradoxo da Prova Surpresa*, em que a tentativa de antecipar racionalmente uma prova anunciada como inesperada parece anular a própria surpresa prometida. Em seguida, encontramos o *Paradoxo do Conhecedor*, no qual uma sentença autorreferencial leva a contradições quando assumimos princípios epistêmicos plausíveis. O capítulo mostra como esses enigmas colocam em xeque noções como fechamento epistêmico, bivalência e o princípio KK, além de revelarem a dificuldade em articular coerentemente conhecimento, crença e previsão. Ao mesmo tempo, ilustra o papel filosófico dos paradoxos epistêmicos: longe de serem apenas curiosidades lógicas, eles funcionam como testes de estresse para nossas teorias do saber.

O capítulo (IV) (p. 128–158) revisita os argumentos clássicos de **Zenão de Eleia** contra a pluralidade e o movimento. Em linguagem simples, Zenão mostra como uma sequência infinita de etapas "inocentes" parece bloquear mudanças que tomamos como óbvias: antes de alcançar um ponto, é preciso chegar à metade do caminho; antes disso, à metade da metade, e assim por diante – o que sugere que *Aquiles* jamais alcançaria a tartaruga, que *a dicotomia* impediria qualquer partida, que *a flecha* estaria em repouso a cada instante e que *o estádio* geraria resultados incompatíveis sobre velocidades relativas. O texto organiza o material em duas famílias. Primeiro, os **paradoxos da pluralidade** – *Densidade*, *Extensão Finita* e *Divisibilidade Completa* – que tensionam a ideia de que haja muitos corpos com tamanhos determinados ocupando um contínuo espacial. Em seguida, os **paradoxos do movimento** – *Dicotomia*, *Aquiles e a Tartaruga*, *A Flecha* e *O Estádio* –, além de peças correlatas como o *Paradoxo do Lugar* e *O Grão de Trigo*. Ao longo da exposição, o capítulo contrasta leituras antigas e respostas contemporâneas: o cálculo de *limites* e a teoria da *medida* explicam como somas infinitas podem ter valor finito; a noção física-matemática de *velocidade instantânea* responde à Flecha; e a análise de *supertarefas* esclarece o que significa (e o que não significa) "completar" infinitas etapas em tempo finito. O resultado é um panorama rigoroso que conecta as aporias eleáticas à metafísica do espaço-tempo e às ferramentas matemáticas hoje empregadas para lidar com o infinito e o contínuo.

Por fim, o capítulo (V) (p. 159–204) é dedicado ao **Paradoxo de Skolem**, que nasce do encontro entre o teorema de Löwenheim–Skolem e o teorema de Cantor. De um lado, Cantor mostrou que certos conjuntos, como os números reais, são não-enumeráveis; de outro, Löwenheim–Skolem garante que qualquer teoria de primeira ordem com modelos infinitos – inclusive a teoria dos conjuntos – possui também modelos enumeráveis. O aparente choque é imediato: como pode um modelo enumerável satisfazer a sentença de que existem conjuntos não-enumeráveis? O capítulo discute essa tensão em detalhe, mostrando como ela revela tanto os limites expressivos da lógica de primeira ordem quanto questões filosóficas sobre relatividade semântica e realismo matemático. Ao mesmo tempo, apresenta as principais estratégias de resposta, desde a análise técnica dos modelos até reflexões mais amplas sobre o estatuto do infinito. O paradoxo de Skolem, assim, exemplifica como mesmo a matemática formalizada está sujeita a dilemas conceituais profundos.

O percurso deste volume acompanha a própria diversidade dos paradoxos. Iniciamos com o desafio epistêmico do *Paradoxo da Cognoscibilidade*, que põe em xeque a relação entre verdade e possibilidade de conhecimento; em seguida, passamos ao *Paradoxo de Sorites*, que revela a fragilidade de conceitos vagos em nossa linguagem e o impacto da vagueza sobre a lógica. O caminho prossegue com outros paradoxos epistêmicos – como o do Conhecedor e o da Prova Surpresa –, que testam os limites da previsibilidade e da autorreferência no domínio do saber. Retornamos então às origens da tradição paradoxal com os argumentos de Zenão, cujas aporias sobre o movimento continuam a alimentar debates sobre infinito e continuidade. Por fim, encontramos o Paradoxo de Skolem, que expõe tensões internas na lógica contemporânea e questiona a própria noção de modelo matemático.

Reunidos, esses capítulos mostram que os paradoxos não são meras curiosidades, mas instrumentos poderosos para sondar as fronteiras de nossas intuições, de nossas teorias e de nossas práticas de raciocínio. São eles que, ao desafiar o que tomamos por certo, abrem caminho para a reformulação crítica de conceitos fundamentais.

Kherian Gracher
Organizador

(I) Paradoxo da Cognoscibilidade de Fitch[1]

Título Original: Fitch's Paradox of Knowability
Autores: Berit Brogaard e Joe Salerno
Tradução: Alan R. Antezana
Revisão: Daniel Alves da Silva Lopes Diniz

O paradoxo da cognoscibilidade de Fitch (também conhecido como o paradoxo da cognoscibilidade ou paradoxo Church–Fitch) concerne qualquer teoria comprometida com a tese de que todas as verdades podem ser conhecidas. Exemplos históricos de tais teorias incluem possivelmente o antirrealismo semântico de Michael Dummett (a tese de que toda verdade é verificável); o construtivismo matemático (a tese de que a verdade de uma fórmula matemática depende dos construtos mentais que matemáticos usam para demonstrar esta fórmula); o realismo interno de Hilary Putnam (a tese de que a verdade é aquilo em que acreditaríamos em circunstâncias epistêmicas ideais); a teoria pragmática da verdade de Charles Sanders Peirce (de que a verdade é aquilo com que concordaríamos no limite da investigação); o positivismo lógico (a visão de que significado é dado por condições de verificação); o idealismo transcendental de Kant (isto é, que todo o conhecimento é conhecimento de aparências); e o idealismo de George Berkeley (isto é, ser é ser perceptível).

O conceito em vigor da noção de "cognoscibilidade" permanece difícil de definir, mas deve estar entre identificar, de forma não-informativa, a verdade com o que Deus saberia,

[1]BROGAARD, Berit; SALERNO, Joe, "Fitch's Paradox of Knowability", In: ZALTA, E. N. (ed.). The Stanford Encyclopedia of Philosophy (Fall 2019 Edition). Stanford, CA: The Metaphysics Research Lab, 2019. Disponível em: https://plato.stanford.edu/archives/fall2019/entries/fitch-paradox/.
A seguir está a tradução da entrada sobre o Paradoxo da Cognoscibilidade de Fitch de Berit Brogaard e Joe Salerno na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/fall2019/entries/fitch-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/fitch-paradox/. Agradecemos ao Prof. Dr. Edward N. Zalta pela permissão para traduzir e publicar esta entrada.

e ingenuamente identificar a verdade com o que humanos atualmente sabem. Identificar a verdade com o que Deus sabe não melhora a inteligibilidade do conceito; e identificá-la com o que humanos atualmente sabem falha em apreciar a objetividade e o potencial de descobrimento da verdade. O caminho do meio, que podemos chamar de *antirrealismo moderado*, pode ser situado logicamente em algum lugar próximo ao princípio da cognoscibilidade:

$$\forall p(p \to \Diamond Kp), \qquad \text{(Princípio K)}$$

que diz formalmente que, para todas as proposições p, se p, então é possível saber que p.

O grande problema para o meio-termo é o *paradoxo de Fitch*. Ele é a prova que mostra (em uma lógica modal normal acrescida de um operador de cognoscibilidade) que, se “todas as verdades podem ser conhecidas”, então “todas as verdades são conhecidas”.

$$\forall p(p \to \Diamond Kp) \vdash \forall p(p \to Kp) \qquad \text{(Paradoxo K)}$$

Desse modo, a prova realiza a parte interessante do trabalho de fazer o antirrealismo moderado colapsar em idealismo ingênuo.

Qual é o paradoxo? Timothy Williamson (2000b) afirma que o paradoxo da cognoscibilidade não é um paradoxo, mas um “vexame”—um vexame para várias formas de antirrealismo que negligenciaram por tanto tempo um contraexemplo simples. Ele nota que é “uma afronta” a várias teorias filosóficas, mas não ao senso comum. Outros discordam. O paradoxo não é que a prova de Fitch é uma ameaça imediata ao meio-termo, mas sim que a prova de Fitch, empregando recursos modais e epistêmicos mínimos, faz o meio-termo colapsar na visão ingênua. O paradoxo, como articulado em Kvanvig (2006) e Brogaard & Salerno (2008), é que o antirrealismo moderado não parece ser exprimível como uma tese distinta, logicamente mais fraca que o idealismo ingênuo. Isso é interessante e problemático independentemente da atitude de alguém, favorável ou contrária ao antirrealismo moderado.

1. Breve Histórico

A literatura sobre o paradoxo da cognoscibilidade emerge em resposta a uma prova publicada pela primeira vez por Frederic Fitch em seu artigo de 1963 intitulado *“A Logical Analysis of Some Value Concepts”* (“uma análise lógica de alguns conceitos de valor”). O Teorema 5, como assim foi chamado, ameaça colapsar diversas diferenças modais e epistêmicas. Seja ignorância a falha em conhecer alguma verdade. Então o Teorema 5 colapsa um comprometimento com a ignorância contingente em comprometimento com a ignorância necessária. Isso porque mostra que a existência de verdades de fato desconhecidas acarreta a existência de verdades necessariamente desconhecidas. Formalmente,

$$\exists p(p \wedge \neg Kp) \vdash \exists p(p \wedge \neg \Diamond Kp). \qquad \text{(Teorema 5)}$$

A conversa do Teorema 5 é trivial (uma vez que verdade implica possibilidade), então Fitch dedica a maior parte do tempo a apagar quaisquer diferenças lógicas entre a existência de ignorância contingente e a existência de incognoscibilidade necessária.

É todavia a contrapositiva do Teorema 5 que é geralmente referida como o paradoxo:

$$\forall p(p \to \Diamond Kp) \vdash \forall p(p \to Kp). \qquad \text{(Paradoxo K)}$$

Ela nos diz que, se qualquer verdade pode ser conhecida, então segue-se que toda verdade é, de fato, conhecida.

A primeira versão da prova foi transmitida a Fitch por um parecerista anônimo em 1945. Em 2005, nós descobrimos que Alonzo Church foi esse parecerista (Salerno 2009b). Seus relatórios foram publicados em sua totalidade em Church (2009). Fitch aparentemente não considerou o resultado como paradoxal. Ele publicou a prova em 1963 para evitar um tipo de "falácia condicional" que ameaçava o seu desejo bem informado de uma análise de valor. A análise diz, a grosso modo, que x tem valor para s somente se há uma verdade p tal que, se s soubesse que p, então s desejaria x. A existência de verdades incognoscíveis explica, em última instância, por que ele restringe as variáveis proposicionais a proposições cognoscíveis. É porque uma verdade incognoscível permite um antecedente impossível no contrafactual de Fitch, e então trivializa a análise. Uma vez que a teoria do valor de Fitch não é o contexto em que o paradoxo é amplamente discutido, não o discutiremos mais aqui.

Redescoberto em Hart e McGinn (1976) e Hart (1979), o resultado foi tomado como uma refutação do verificacionismo, a postura de que todas as asserções com significado (e, portanto, todas as verdades) são verificáveis. Afinal, se alguém aceita o princípio da cognoscibilidade $\forall p(p \to \Diamond Kp)$, também está comprometido com a asserção absurda de que todas as verdades são conhecidas. Mackie (1980) e Routley (1981), entre outros daquele período, apontam dificuldades com essa posição geral, mas, em última instância, aceitam que o resultado de Fitch é uma refutação da posição de que todas as verdades são cognoscíveis, e que várias formas de verificacionismo são ameaçadas por motivos relacionados. Desde o início da década de oitenta, todavia, tem havido um esforço considerável para analisar a prova como uma prova paradoxal. Por que, afinal, uma teoria epistêmica da verdade colapsaria o conhecimento possível no conhecimento atual? Intuitivamente, que a verdade deva ser entendida em termos das capacidades epistêmicas de agentes não oniscientes é uma posição ao menos coerente—uma posição distinta de, e mais plausível que, a tese de que todas as verdades são conhecidas. Além do mais, tem-se estranhado que formas sofisticadas da teoria epistêmica da verdade sejam vítimas de uma dedução tão breve. Logo, a prova de Church–Fitch veio a se tornar conhecida como o *paradoxo* da cognoscibilidade.

2. O Paradoxo da Cognoscibilidade

O raciocínio de Fitch envolve quantificação sobre sentenças. Nossas variáveis proposicionais p e q terão sentenças declarativas como substitutas. Seja K o operador epistêmico "é sabido por alguém em algum instante que". Seja $\Diamond$ o operador modal "é possível que".

Suponha que o princípio da cognoscibilidade (KP)—que todas as verdades são cognoscíveis por alguém em algum momento:

$$\forall p(p \rightarrow \Diamond Kp). \tag{KP}$$

E suponha que, coletivamente, não somos oniscientes, e que há uma verdade que não é conhecida.

$$\exists p(p \wedge \neg Kp). \tag{NonO}$$

Se essa asserção existencial é verdadeira, então uma instância dela também é:

$$p \wedge \neg Kp. \tag{1}$$

Agora considere a instância de KP substituindo a variável p em KP pela linha 1:

$$(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge \neg Kp). \tag{2}$$

Segue de forma trivial que é possível saber a conjunção expressa na linha 1:

$$\Diamond K(p \wedge \neg Kp) \tag{3}$$

Todavia, pode ser mostrado de forma independente que é impossível conhecer essa conjunção. Portanto, a linha 3 é falsa.

O resultado independente pressupõe dois princípios epistêmicos muito modestos: primeiramente, conhecer uma conjunção implica conhecer cada um dos conjuntos; em segundo lugar, conhecimento implica verdade. Respectivamente,

$$K(p \wedge q) \vdash Kp \wedge Kq \tag{A}$$

$$Kp \vdash p \tag{B}$$

São também pressupostos dois modestos princípios modais: primeiro, que todos os teoremas são necessários; segundo, que necessariamente $\neg p$ implica que p é impossível. Respectivamente,

$$\textit{Se } \vdash p, \textit{ então } \vdash \Box p. \tag{C}$$

$$\textit{Se } \Box\neg p, \textit{ então } \neg\Diamond p. \tag{D}$$

Considere o resultado independente:

$K(p \wedge \neg Kp)$	hipótese [para redução ao absurdo]	(4)
$Kp \wedge K\neg Kp$	de 4, por (A)	(5)
$Kp \wedge \neg Kp$	de 5, aplicando (B) à direita	(6)
$\neg K(p \wedge \neg Kp)$	de 4–6, por redução ao absurdo, descarregando a assunção 4	(7)
$\Box\neg K(p \wedge \neg Kp)$	de 7, por (C)	(8)
$\neg\Diamond K(p \wedge \neg Kp)$	de 8, por (D)	(9)

A linha 9 contradiz a linha 3, então uma contradição é consequência de (KP) e (NonO). A adepta da posição de que todas as verdades são cognoscíveis deve negar que somos não-oniscientes:

$$\neg\exists p(p \wedge \neg Kp). \tag{10}$$

E segue-se disso que todas as verdades são conhecidas:

$$\forall p(p \rightarrow \neg Kp). \tag{11}$$

A aliada da visão de que todas as verdades são cognoscíveis (por alguém em algum momento) é forçada a admitir, de modo absurdo, que toda verdade é conhecida (por alguém em algum momento).

3. Revisões Lógicas

Nesta seção, inspecionamos a viabilidade de admitir o raciocínio de Fitch como inválido. Estaria o raciocínio epistêmico de Fitch em ordem? A lógica da cognoscibilidade é a lógica clássica? Ou mais diretamente: o princípio da cognoscibilidade carrega em si considerações especiais que justifiquem a revisão da lógica clássica? Se sim, essa revisão lógica invalida o raciocínio de Fitch? Se o raciocínio é inválido, existem paradoxos próximos que ameaçam o princípio da cognoscibilidade sem violar os padrões lógicos de maior relevância?

3.1 Revisão Epistêmica

O problema com o raciocínio de Fitch não é com nenhuma das inferências epistêmicas A ou B. Embora algumas pessoas tenham argumentado que conhecer uma conjunção não implica conhecer cada um dos argumentos da conjunção (Nozick 1981), Williamson (1993) e Jago (2010) mostraram que versões do paradoxo não demandam a suposição desta distributividade. E perguntas sobre a factividade de K podem ser neutralizadas de forma rápida, uma vez que paradoxos relacionados emergem substituindo o operador factivo “É sabido que” por um operador não-factivo, tal como “Acredita-se racionalmente que” (Mackie 1980:92; Edgington 1985:558–559; Tennant 1997:252–259; Wright 2000:357).

Aparecem em vários artigos discussões profundas e interessantes sobre operadores epistêmicos e/ou análogos temporais no contexto do paradoxo de Fitch. Burgess (2009) considera análogos temporais. Van Benthem (2004; 2009), van Ditmarsh et al. (2012), Berto et al. (em fase de publicação) e Holliday (2018) exploram o problema em estruturas epistêmicas dinâmicas. Palczewski (2007), Kelp and Pritchard (2009), Chase et al. (2018), e Heylen (no prelo) consideram noções não-factivas de conhecimento e cognoscibilidade. Linsky (2009), Paseau (2008), Jago (2010), Carrara et al. (2011), e Rosenblatt (2014) debatem a viabilidade de conhecimento sobre tipos.

3.2 Revisão Intuicionista

Williamson (1982) argumenta que a prova de Fitch não é uma refutação do antirrealismo, mas um motivo para que a antirrealista aceite a lógica intuicionista. A lógica intuicionista não valida a eliminação da dupla negação, graças a uma leitura verificacionista (ou construtivista) da negação e da quantificação existencial:

$$\neg\neg p \vdash p,$$

nem a seguinte regra de troca de quantificadores:

$$\neg\forall x P[x] \vdash \exists x \neg P[x].$$

Sem a eliminação da dupla negação, não é possível derivar a conclusão de Fitch de que "todas as verdades são conhecidas" (linha 11) de "não há uma verdade que é incognoscível" (linha 10). Considere a linha 10,

$$\neg\exists p(p \wedge \neg Kp).$$

Disso podemos derivar de forma intuicionista

$$\forall p \neg(p \wedge \neg Kp).$$

Note, todavia, que, sem a eliminação da dupla negação,

$$\neg(p \wedge \neg Kp)$$

não acarreta

$$p \rightarrow Kp.$$

Suponha

$$\neg(p \wedge \neg Kp),$$

e suponha p para a introdução de um condicional. E suponha $\neg Kp$, por redução ao absurdo. Nós podemos fazer a conjunção de p com $\neg Kp$ para obter

$$p \wedge \neg Kp.$$

Isso contradiz a nossa primeira hipótese. Então, por redução, $\neg\neg Kp$. Sem a eliminação da dupla negação, não podemos concluir Kp, e, dessa forma, não podemos introduzir o condicional

$$p \to Kp.$$

A intuicionista, todavia, está comprometida, pela introdução do condicional, com

$$p \to \neg\neg Kp.$$

Existe algum debate sobre se esta consequência é suficientemente problemática, mas a antirrealista intuicionista se consola com o fato de que não está comprometida com a asserção absurda de que todas as verdades são conhecidas. Aparecem em Murzi (2010; 2012), Murzi et al. (2009), e Zardini (2015) discussões muito interessantes sobre as esperanças e os sonhos do antirrealismo intuicionista neste contexto.

3.3 Problemas para a Revisão Intuicionista

Uma vez que o raciocínio de Fitch é intuicionisticamente válido por meio da linha 10, a antirrealista intuicionista deve aceitar que nenhuma verdade é desconhecida: $\neg\exists p(p \to Kp)$. Isso pode ser prejudicial o bastante, porque a antirrealista parece não poder dar credência ao truísmo de que nós não somos oniscientes (nem individualmente, nem coletivamente). Williamson responde que a antirrealista intuicionista pode expressar naturalmente a nossa não-onisciência como "nem todas as verdades são conhecidas":

$$\neg\forall p(p \to Kp). \tag{12}$$

Esta asserção é classicamente, mas não intuicionisticamente, equivalente à tese de não-onisciência:

$$\exists p(p \wedge \neg Kp).$$

Isso ocorre porque, na lógica intuicionista, a regra de troca de quantificadores, $\neg\forall x P[x] \vdash \exists x \neg P[x]$, não é válida de forma irrestrita. É relevante destacar que a expressão de não-onisciência na linha 12, $\neg\forall p(p \to Kp)$, é inconsistente somente de forma clássica (e não de forma intuicionista) com a linha 10, $\neg\exists p(p \wedge \neg Kp)$. Então a intuicionista antirrealista pode dar expressão consistentemente ao truísmo de que nós não somos oniscientes (com a linha 12), ao passo em que aceita a consequência intuicionista derivada na linha 10. Com efeito, a antirrealista admite tanto que nenhuma verdade é desconhecida, quanto que nem todas as verdades são conhecidas. A satisfatibilidade dessa asserção em termos intuicionistas é demonstrada por Williamson (1988, 1992).

3.4 A indecidibilidade do Paradoxo da Cognoscibilidade

Um problema mais profundo pode persistir para a intuicionista antirrealista. O paradoxo de Fitch se situa na suposição de que existem verdades desconhecidas. Mas considere a suposição intuicionisticamente mais fraca de que existem asserções não decididas, isto é, que alguma p, tal que p é desconhecida e $\neg p$ é desconhecida. Formalmente,

$$\exists p(\neg Kp \wedge \neg K\neg p). \tag{Und}$$

Se (Und) é verdadeiro, então também é verdadeira uma instância sua:

$$\neg Kp \wedge \neg K\neg p. \tag{i}$$

E note que a conclusão intuicionisticamente aceitável na linha 10, $\neg\exists p(p \wedge \neg Kp)$, é intuicionisticamente equivalente à asserção universal:

$$\forall p(\neg Kp \rightarrow \neg p). \tag{ii}$$

Derivando $\neg Kp \rightarrow \neg p$ e $\neg K\neg p \rightarrow \neg\neg p$ de (ii), e aplicando os dois operandos da conjunção (i), respectivamente, nos dá a contradição $\neg p \wedge \neg\neg p$. A antirrealista intuicionista é forçada a admitir que não há constatações não decididas:

$$\neg\exists p(\neg Kp \wedge \neg K\neg p). \tag{iii}$$

O argumento acima é dado por Percival (1990:185). Uma vez que é intuicionisticamente aceitável, ele tem a intenção de mostrar que a antirrealista intuicionista ainda está em apuros.

Em resposta, a antirrealista pode novamente utilizar a estratégia de Williamson para revisar a lógica e reconstruir uma expressão do truísmo epistêmico. Aceite apenas as consequências intuicionistas de KP (neste caso, de que não há asserções não decididas), e dê expressão ao truísmo sobre indecidibilidade ao afirmar que nem todas as asserções são decididas:

$$\neg\forall p(Kp \vee K\neg p). \tag{iv}$$

A reinterpretação da intuição de indecidibilidade na linha (iv) nos dá uma asserção que é classicamente, mas não intuicionisticamente, equivalente a (Und). E, dessa forma, é somente classicamente, e não intuicionisticamente, inconsistente com a inferência na linha (iii).

Paradoxos correlatos sobre a indecidibilidade da cognoscibilidade são discutidos em Wright (1987: 311), Williamson (1988: 426) e Brogaard e Salerno (2002: 146–148). Os paradoxos da indecidibilidade dão à antirrealista motivos adicionais para revisar a lógica clássica em favor da lógica intuicionista. Acompanhados de uma reinterpretação de nossas modestas intuições epistêmicas, o espaço lógico para o antirrealismo é retomado.

O que tudo isso sugere é que o antirrealismo intuicionista é coerente. Mas essa abordagem tem a motivação correta? Seria *ad hoc* a revisão da lógica clássica, ou a engenhosa reinterpretação de nossas intuições epistêmicas?

O suposto direito da antirrealista de abrir mão da lógica clássica em favor da lógica intuicionista tem sido defendido de forma independente. O argumento remonta a Dummett (1976, e alhures). Interpretações mais recentes do argumento antirrealista pela revisão lógica aparecem em Wright (1992: Chp. 2), Tennant (1997: Chp. 7), e Salerno (2000). Os detalhes sobre, e o sucesso ou fracasso dos argumentos pela revisão lógica são assunto para outro momento. Por enquanto, é suficiente indicar que a ameaça do paradoxo de Fitch não é a única motivação para a antirrealista favorecer uma lógica não-clássica.

E a reinterpretação de nossas intuições epistêmicas? Seria ela bem justificada? De acordo com Kvanvig (1995), ela não é. Por que pensaríamos que os tratamentos intuicionistas da não onisciência e da indecidibilidade são melhores do que os nossos tratamentos iniciais do senso comum? E como a antirrealista deve explicar a aparente trivialidade desses tratamentos do senso comum? Essas questões não foram respondidas.

Além disso, algumas das consequências intuicionistas de KP são por vezes consideradas suficientemente ruins. Ainda que "não há verdades incognoscíveis" ou "não há asserções indecidíveis" sejam intuicionisticamente toleráveis, o seguinte não parece ser o mesmo: "se p é desconhecida, então $\neg p$". Formalmente, $\neg Kp \rightarrow \neg p$. Esta asserção segue-se intuicionisticamente de $p \rightarrow \neg\neg Kp$, que nós já estabelecemos como uma consequência intuicionista de KP. Mas $\neg Kp \rightarrow \neg p$ parece ser falsa no discurso empírico. Por que o fato de que ninguém nunca saberá que p ser suficiente para a falsidade de p? Veja Percival (1990) e Williamson (1988) para discussões adicionais sobre esse problema, e sobre problemas relacionados, em torno da aplicação do antirrealismo intuicionista ao discurso empírico. DeVidi e Solomon (2001) discordam. Eles argumentam que as consequências intuicionistas não são inaceitáveis para alguém interessado em uma teoria epistêmica da verdade. De fato, elas são centrais para uma teoria epistêmica da verdade.

Por estes motivos, um apelo à lógica intuicionista por si próprio é geralmente encarado como insatisfatório em lidar com o paradoxo da cognoscibilidade. Exceções incluem Bermúdez (2009), Dummett (2009), Rasmussen (2009) e Maffezioli, Naibo & Negri (2013).

3.5 Revisão Paraconsistente

Outro desafio à lógica do paradoxo de Fitch é mencionado em Routley (1981) e defendido por Beall (2000): a ideia de que a lógica correta da cognoscibilidade é paraconsistente. Em uma lógica paraconsistente, contradições não trivializam a teoria, porque elas não "explodem". Isto é, em uma lógica paraconsistente, a inferência de $p \wedge \neg p$ para uma conclusão arbitrária r não é válida. Desta forma, algumas contradições são permitidas e consideradas possíveis.

Beall afirma que (1) a prova de Fitch assume a suposição de que, para todas as sentenças p, a contradição $Kp \wedge \neg Kp$ é impossível e que (2) temos evidência independente para pensar que $Kp \wedge \neg Kp$, para algum p. A evidência independente jaz no paradoxo do agente cognoscente (que não deve ser confundido com o paradoxo da cognoscibilidade). A versão relevante do paradoxo do agente cognoscente pode ser demonstrada ao considerar a seguinte sentença autorreferencial:

$$k \text{ é desconhecida.} \tag{k}$$

Assuma pelo argumento que k é conhecida. Então, assumindo que conhecimento implique verdade, k é verdadeira; mas k afirma que k é desconhecida. Então k é desconhecida. Consequentemente, k tanto é conhecida quanto desconhecida. Mas, desse modo, a nossa suposição (de que k é conhecida) é falsa, e de maneira demonstrável. E, dado que uma falsidade demonstrada é sabidamente falsa, segue que é conhecido que k é desconhecida. Isso é afirmar que é conhecido que k. Mas já mostramos que, se é conhecido que k, então k é tanto conhecida quanto desconhecida. Então está demonstrado que k é tanto conhecida quanto desconhecida. É demonstravelmente o caso que a descrição completa do nosso conhecimento inclui tanto Kk quanto $\neg Kk$. Esse é o paradoxo do agente cognoscente.

Beall sugere que o paradoxo nos dá alguma evidência independente para pensar que $Kp \wedge \neg Kp$, para algum p, tal que a descrição completa do conhecimento humano tem a propriedade interessante de ser inconsistente. Com uma lógica paraconsistente, poder-se-ia aceitá-lo sem incorrer em trivialidade. E dessa forma é sugerido que se tome o caminho da paraconsistência e se aceite $Kp \wedge \neg Kp$ como uma consequência real do princípio da cognoscibilidade. Beall conclui afirmando que o raciocínio de Fitch, sem uma devida resposta ao agente cognoscente, é ineficaz contra o princípio da cognoscibilidade, uma vez que o raciocínio de Fitch supostamente depende da suposição de que, para todo p, é impossível que $Kp \wedge \neg Kp$.

Note que a nossa apresentação do raciocínio de Fitch não faz menção explícita à suposição de que $Kp \wedge \neg Kp$ é impossível. Então aqui tentamos indicar exatamente onde o raciocínio de Fitch falha na abordagem acima. Afirma-se na linha 9 (na primeira seção deste verbete) que $K(p \wedge \neg Kp)$ é impossível. É claro que $K(p \wedge \neg Kp)$ implica a contradição $Kp \wedge \neg Kp$. Dessa forma, se o raciocínio é que $K(p \wedge \neg Kp)$ é impossível porque contradições são impossíveis, então Beall estaria atacando diretamente o argumento aqui apresentado. Note, todavia, que há uma diferença sutil no argumento, que segue da seguinte forma. $K(p \wedge \neg Kp)$ implica a contradição $Kp \wedge \neg Kp$. Então, por redução ao absurdo, $K(p \wedge \neg Kp)$ é falso. Por necessitação, segue que $K(p \wedge \neg Kp)$ é necessariamente falso. A depender da lógica paraconsistente, a paraconsistentista pode objetar ao uso de *reductio*, ou pode objetar a outras inferências. A afirmação de que $K(p \wedge \neg Kp)$ é impossível (na linha 9) é inferida dessa afirmação de que $K(p \wedge \neg Kp)$ é necessariamente falsa. Isso pode preocupar a paraconsistentista. Pela ótica de alguém convivendo com contradições, pode não se seguir que uma asserção inconsistente é impossível mesmo que seja necessariamente falsa. Afinal, nessa abordagem uma asserção necessariamente falsa pode ser tanto verdadeira quanto falsa em algum mundo, caso em que a asserção é tanto necessariamente falsa quanto possível. Se isto estiver correto, então a inferência de $\Box\neg p$ para $\neg\Diamond p$ tem contraexemplos e não pode ser empregada para inferir $\neg\Diamond K(p \wedge \neg Kp)$ de $\Box\neg K(p \wedge \neg Kp)$.

Os insights de Beall dependem de uma série de coisas: (1) a força da evidência independente para contradições epistêmicas verdadeiras; (2) a adequação das resoluções propostas ao paradoxo do agente cognoscente, (3) a dúvida sobre se o raciocínio de Fitch é ineficaz sem uma solução ao agente cognoscente, e (4) uma interpretação para $\Box$ e $\Diamond$ que invalida a inferência relevante (de $\Box\neg p$ para $\neg\Diamond p$) enquanto permanece fiel ao papel desempenhado por $\Diamond$ no princípio da cognoscibilidade. Nós deixamos esses problemas para um debate posterior; mas compare a Wansing (2002), onde uma lógica modal construtiva paraconsistente com uma negação forte é proposta para bloquear o paradoxo.

Desenvolvimentos mais recentes da abordagem paraconsistentes aparecem em Beall (2009) e Priest (2009).

4. Restrições Semânticas

O restante das propostas são estratégias de restrição. Elas reinterpretam KP restringindo o seu quantificador universal. Na prática, as estratégias de restrição invalidam o raciocínio de Fitch ao proibir as instâncias de KP que levam ao paradoxo. Nesta seção, nós examinamos as razões semânticas para restringir o quantificador universal em KP.

4.1 Situações e Operações Rígidas

Edgington (1985) oferece um diagnóstico situacional do paradoxo de Fitch. Ela afirma que o problema jaz no fracasso em distinguir entre "saber em uma situação que p" e "saber que p é o caso em uma situação". No último caso, a situação é (ao menos parcialmente) aquela sobre a qual é o conhecimento. No caso anterior, a situação é aquela em que o conhecimento é entretido. Por exemplo, eu poderia saber na minha situação atual que eu estaria sofrendo em uma situação contrafactual onde meu dente é arrancado. Notavelmente, a situação em que o conhecimento é entretido pode ser diferente da situação que é tema do conhecimento. Em uma situação em que meu dente *não* é arrancado, eu posso saber coisas que são sobre uma situação em que meu dente é puxado.

O que são situações? O exemplo acima parece sugerir que situações são mundos. Situações, todavia, podem ser menos completas que mundos; isto é, elas não precisam ter valores de verdade fixados para asserções que são irrelevantes para o contexto. Considere um exemplo por Linstöm: eu posso saber em uma dada situação perceptual s que João (um dos participantes de um jogo de cartas) tem a melhor mão e que nenhum dos outros participantes sabe disso. Neste caso, meu conhecimento é de uma situação s^*, o jogo de cartas, mas meu conhecimento é adquirido em uma diferente situação s, minha situação perceptual. A situação s^* não somente é determinada, mas a informação relevante é limitada pelo contexto do jogo de cartas. E s é fixada e limitada pelo contexto da situação perceptual. Edgington prefere falar de situações em vez de mundos, porque o conhecimento de situações não-atuais, diferentemente do conhecimento de mundos não-atuais, não requer conhecimento com uma quantidade infinita de detalhes.

Explicitando a distinção da teoria da situação entre “saber em” e “saber sobre”, nós podemos reinterpretar o princípio da cognoscibilidade: para cada asserção p e situação s, se p é verdadeiro em s então existe uma situação s^* em que é sabido que p é verdadeiro em s. Edgington requer da cognoscibilidade uma tese menos geral: se p é verdadeiro em uma situação atual s, então existe uma situação possível s^* em que é sabido que p é verdadeiro em s. Chame isto de E-cognoscibilidade ou EKP:

$$Ap \to \Diamond KAp, \tag{EKP}$$

onde A é o operador de atualidade que pode ser lido como “em alguma situação atual”, e $\Diamond$ é o operador de possibilidade que pode ser lido como “em alguma situação possível”.

Como vemos, EKP restringe o princípio da cognoscibilidade para verdades atuais, dizendo que p é atualmente verdadeiro somente se existe uma situação possível em que é sabido que p é atualmente verdadeiro.

A sugestão importante é esta. Assim como pode existir conhecimento atual do que é contrafactualmente o caso, pode haver conhecimento contrafactual do que é atualmente o caso. De fato, em vista da prova de Fitch, EKP requer a existência de conhecimento não-atual. Vejamos por quê.

Verdades atuais da forma $p \wedge \neg Kp$ teriam de ser E-cognoscíveis, mas $p \wedge \neg Kp$ não pode ser atualmente conhecido como sendo atualmente o caso. O raciocínio aqui é análogo ao raciocínio de Fitch.

A lição é esta. Visto que, para algum p, $p \wedge \neg Kp$ é atualmente o caso, E-cognoscibilidade nos compromete a um conhecimento possível de que $p \wedge \neg Kp$ é possivelmente o caso. Uma vez que esse conhecimento não pode ser atual, a E-cognoscibilidade requer conhecimento não atual do que é atualmente o caso. E-cognoscibilidade então nega a seguinte suposição: dada uma asserção p, se é conhecido que p em s, então em s é sabido que p. Pela análise de Edgington, é exatamente essa suposição implícita que descarrilha o raciocínio de Fitch. O paradoxo é bloqueado sem ela.

4.2 Problemas para as Situações

Uma vez que o operador de atualidade designa *rigidamente* as situações, os valores de verdade das asserções da forma Ap não vão variar entre as situações possíveis. "Ap" implica "em toda situação Ap". Logo, como Edgington sabe, se Ap então é necessário que Ap. Isso por si mesmo impõe um problema a EKP. A crítica é que a abordagem de Edgington não é suficientemente geral. Qualquer um que endossaria o princípio da cognoscibilidade provavelmente pensaria que ele vale para todas as verdades, e não apenas as verdades necessárias que envolvem o operador de atualidade. EKP parece ser uma tese muito limitada que falha na especificação de uma restrição epistêmica sobre a verdade contingente (Williamson 1987a).

Outras críticas emergem quando tentamos afirmar algo informativo sobre o que constitui conhecimento não atual sobre o que atualmente é o caso. Se existe um tal conhecimento não-atual, existe pensamento não atual sobre uma situação atual. Então o pensador não atual de alguma forma tem o conceito de uma situação atual. Mas como é possível para um pensador não atual ter um conceito que é especificamente sobre situações neste mundo atual? Não será suficiente para a pensadora expressar o pensamento "atualmente p", uma vez que "atualmente" vai designar rigidamente somente situações em seu próprio mundo. Além disso, uma vez que não existe nenhum elo causal entre o mundo atual w_1 e o mundo não atual relevante w_2, não é claro como um pensamento não atual em w_2 pode ser unicamente sobre w_1 (Williamson, 1987a: 257-258). Logo, não é claro como pode haver conhecimento não atual sobre o que é atualmente o caso.

É claro que o conhecimento atual sobre o não-atual não é melhor em especificar mundos. O problema especial para a pensadora não atual é que o conteúdo do seu pensamento deve ser precisamente o conteúdo que apreendemos quando consideramos a verdade de Ap. Estando no mundo atual, somos capazes de especificar esse mundo de forma a distingui-lo de todos os outros. Quando consideramos a verdade de Ap, nosso contexto fixa o conteúdo de A especificamente. Então se realmente é Ap que é cognoscível pelo agente cognoscente não atual, então deve ser Ap aquilo que ele apreende—isto é, deve ser o mesmo conceito que nós apreendemos. Como isso é possível é precisamente o problema.

Críticas adicionais, e relacionadas à proposta de Edgington, podem ser encontradas em Wright (1987), Williamson (1987b,2000) e Percival (1991). Desenvolvimentos formais desta proposta, incluindo pontos que tratam de algumas dessas preocupações aparecem em Rabinowicz e Segerberg (1994), Lindström (1997), Rückert (2003), Edgington (2010), Fara (2010), Proietti e Sandu (2010), e Schlöder (no prelo).

4.3 Falácias Modais e Asserções Não-rígidas

Kvanvig (1995) acusa Fitch de uma falácia modal. A falácia é uma substituição ilícita em um contexto modal. Considere uma falácia modal familiar. Para todas as pessoas x, existe um mundo possível em que x não é o inventor das lentes bifocais. Mesmo Benjamin Franklin, o verdadeiro inventor das bifocais, pode não as ter inventado. Logo, existe um mundo possível em que o inventor das bifocais não é o inventor das bifocais. Nós representamos o argumento de maneira formal. Permita que nossos quantificadores quantifiquem sobre pessoas, e seja "i" o designador não-rígido de "o inventor das bifocais". Considere o argumento:

$$\forall x \Diamond \neg (x = i)$$

Logo,

$$\Diamond \neg (i = i)$$

Embora qualquer um possa não ter sido o inventor das lentes bifocais, disso não se segue (e de fato é falso) que é possível que o inventor das bifocais não seja idêntico ao inventor das bifocais. Afinal, é necessário que o inventor das bifocais seja o inventor das bifocais.

A lição é que nós não podemos substituir de forma irrestrita em contextos modais. Substituir em contextos modais, nós podemos dizer, é permitido somente se os termos substituidores são designadores rígidos. No caso do resultado de Fitch, nossos termos são sentenças. O princípio da cognoscibilidade, $\forall p(p \rightarrow \Diamond Kp)$, aparentemente nos permite substituir p por qualquer sentença. Note, todavia, que nosso quantificador tem um escopo amplo sobre $\Diamond$. Nós esperaríamos que as lições da lógica modal quantificada sejam válidas em lógica modal quantificada proposicional. Se é assim, então nós não podemos substituir p por qualquer outra asserção que não seja um designador rígido.

Segundo o diagnóstico de Kvanvig, o problema com a linha argumentativa de Fitch é que, quando ele substituiu p por $p \wedge \neg Kp$ em KP (na linha 2 do resultado), ele não parou para determinar se $p \wedge \neg Kp$ era uma asserção rígida. Kvanvig mantém que $p \wedge \neg Kp$ não é rígido. Então o resultado de Fitch é falacioso devido a uma substituição ilícita em contexto modal. Nós podemos, todavia, reinterpretar $p \wedge \neg Kp$ como rígido. Quando o fazemos, o paradoxo se evapora.

Kranvig propõe que expressões quantificadas são não-rígidas. O motivo que ele dá é que quantificadores designam diferentes objetos em diferentes mundos possíveis. “Todos no curso de lógica de João devem fazer a prova final” é sobre diferentes estudantes em diferentes mundos possíveis. Se Susie tivesse feito o curso, a expressão seria sobre ela, mas ela decidiu não fazer o curso, então atualmente não é sobre ela. Kp é uma abreviação para “é sabido por alguém em algum tempo que p”. Então, Kp é implicitamente quantificado. Implicitamente lê-se $\exists x \exists t(Kxpt)$, que afirma que existe um ser x em um tempo t tal que x sabe que p no tempo t. Da mesma forma, nessa abordagem, a expressão quantificada que Kp abrevia é não-rígida. $\exists x \exists t(Kxpt)$ é sobre diferentes seres e instantes em diferentes contextos modais. Por exemplo, a expressão $\exists x \exists t(Kxpt)$ é sobre coisas e instantes atuais. Mas no contexto modal (por exemplo, $\Diamond \exists x \exists t(Kxpt)$), a expressão é sobre seres e instantes possíveis. Ela diz “existe um mundo possível em que existe um ser x e um instante t tal que x sabe que p no instante t”.

Agora considere a instância relevante da tese de não onisciência de Fitch: $p \wedge \neg Kp$. De forma não abreviada, isso seria lido como $p \wedge \neg \exists x \exists t(Kxpt)$, que diz que p é verdadeiro, mas ninguém nunca sabe que p. A expressão quantificada é, nesta visão, um designador não-rígido. Enunciada no mundo atual, é sobre seres e instantes atuais. Entretanto, assim é argumentado, dentro do escopo de um operador de possibilidade, a designação varia entre ser sobre seres e instantes possíveis. Quando Fitch substituiu p pela verdadeira conjunção ($p \wedge \neg \exists x \exists t(Kxpt)$) no princípio da cognoscibilidade, ele substituiu p por um designador não-rígido, alterando assim a referência da conjunção e cometendo uma falácia modal.

De forma alternativa, sugere Kvanvig, nós podemos caracterizar Kp rigidamente de forma a dizer “existe um ser atual x e um instante atual t tal que é sabido por x no instante t que p”. Uma vez que esta expressão designa rigidamente (isto é, faz referência ao mundo atual independentemente do contexto modal em que ela aparece), ela pode substituir p no princípio da cognoscibilidade. A conjunção reinterpretada não muda a sua designação quando subsumida sob o escopo de $\Diamond$. Além disso, nesta leitura da conjunção, o paradoxo se dissolve. É possível saber que a conjunção reinterpretada é verdadeira. Não há contradição em supor que exista algum ser possível em algum instante possível que sabe que p é uma proposição verdadeira mas nunca sabida por um ser atual em um instante atual. O paradoxo se dissolve.

Discussões adicionais sobre falácias modais e asserções não-rígidas podem ser encontradas em Brogaard e Salerno (2008) e Kennedy (2014).

4.4 Problemas para a Não-rigidez

Williamson (2000b) defende a linha de argumentação de Fitch contra a acusação de Kvanvig. Ele sugere que Kvanvig não está justificado em pensar que a conjunção de Fitch $p \wedge \neg\exists x \exists t(Kxpt)$ não designa rigidamente. A razão que Williamson dá é a seguinte. Uma expressão é não-rígida se, quando proferida em um contexto fixo, ela varia sua referência conforme as circunstâncias em que é avaliada. Mas Kvanving não dá uma razão convincente para pensar que a conjunção de Fitch, como proferida em um contexto fixado, varia sua referência dessa forma. No máximo, Kvanvig mostrou que a conjunção varia sua referência quando o contexto sofre variação, uma vez que seu argumento é que uma sentença quantificada, quando proferida em mundos diferentes, será sobre objetos distintos. Pensar que isso é suficiente para não-rigidez, Williamson argumenta, é confundir não-rigidez com indexicalidade. Por exemplo, "eu estou cansado" é sobre mim mesmo, e continua sendo sobre mim mesmo quando eu avalio o seu valor de verdade em circunstâncias contrafactuais. A sentença pode ter sido falsa. Tivesse eu dormido o suficiente, eu não estaria cansado. Enunciado em um contexto fixo, "eu" designa rigidamente, ainda que seja um indexical. Isso é assim ainda que, tivesse sido proferido em um contexto diferente por outrem, teria sido sobre alguém que não eu. De forma análoga, ainda que expressões quantificacionais sejam indexicais, não se segue que elas sejam não-rígidas. Deste modo, ainda que a conjunção de Fitch seja uma expressão indexical, não nos foi dada uma razão para considerá-la não-rígida. Se isso estiver correto, então nós não temos motivos para pensar que Fitch cometeu a falácia modal em questão.

Kvanvig (2006) responde e desenvolve outros temas interessantes no *Paradoxo da Cognoscibilidade*, que é a única monografia até a presente data dedicada ao tópico.

5. Restrições Sintáticas

As estratégias de restrição supramencionadas envolviam motivos semânticos para limitar a quantificação universal. Nesses casos, KP estava restrito à luz de considerações sobre situações, mundos possíveis ou designação rígida. Outro tipo de estratégia de restrição é sintática. Ela limita o escopo da quantificação universal àquelas fórmulas que têm uma certa forma lógica ou se localizam em uma certa relação de demonstrabilidade. De forma mais geral,

$$p \rightarrow \Diamond Kp, \text{ onde } p \text{ tem a propriedade lógica } F.$$

F então deveria ser uma propriedade lógica que restringe o quantificador conforme algum conjunto de princípios.

5.1 Asserções Cartesianas

Tennant (1997) foca na propriedade de ser cartesiana: uma asserção p é cartesiana se, e somente se, Kp não é demonstravelmente inconsistente. De acordo com a definição, ele restringe o princípio da cognoscibilidade às asserções cartesianas. Chame isso de princípio restrito da cognoscibilidade (T-cognoscibilidade), ou TKP:

$$p \to \Diamond Kp, \text{ onde } p \text{ é cartesiano.} \tag{TKP}$$

Note que o princípio da T-cognoscibilidade é livre dos paradoxos que nós discutimos. É livre do paradoxo de Fitch e do paradoxo de indecidibilidade correlato, uma vez que ambos os resultados substituem a variável em $p \to \Diamond Kp$ pela problemática conjunção de Fitch, $p \wedge \neg Kp$, nos dando $(p \wedge \neg Kp) \to \Diamond K(p \wedge \neg Kp)$. Isto é, eles requerem que $p \wedge \neg Kp$ seja cognoscível caso verdadeira (linha 2 do resultado de Fitch). Mas $p \wedge \neg Kp$ não é cartesiana, já que $K(p \wedge \neg Kp)$ é demonstravelmente inconsistente (acarretando a contradição na linha 6 do resultado de Fitch). Com efeito, TKP oferece uma das restrições mais tolerantes necessárias para proibir a substituição incômoda. Isso porque proíbe somente a substituição daquelas asserções que são logicamente impossíveis de saber.

5.2 Asserções Básicas

Dummett (2001) concorda que o erro da teórica da cognoscibilidade jaz em fornecer um princípio da cognoscibilidade abrangente, ao invés de restrito. E ele concorda que a restrição deveria ser sintática. Dummett restringe o princípio da cognoscibilidade a asserções “básicas” e caracteriza a verdade indutivamente a partir daí. Para Dummett,

$$\begin{aligned}
p &\text{ sse } \Diamond Kp, \text{ onde } p \text{ é básico.}\\
p \text{ e } q &\text{ sse } p \wedge q;\\
p \text{ ou } q &\text{ sse } p \vee q;\\
\text{Se } p \text{ então } q &\text{ sse } p \rightarrow q;\\
\text{Não é o caso que } p &\text{ sse } \neg p;\\
F[\text{Algo}] &\text{ sse } \exists x F[x];\\
F[\text{Tudo}] &\text{ sse } \forall x F[x].
\end{aligned}$$

onde a constante lógica no lado direito de cada cláusula bicondicional é compreendida como sujeita às leis das lógicas intuicionistas.

O princípio da cognoscibilidade de Dummett ou DKP, como o de Tennant, não é ameaçado pelos paradoxos da cognoscibilidade, e pelo mesmo motivo. Ele restringe a classe de asserções que são sujeitas à cognoscibilidade. No caso de Dummett, a conjunção problemática de Fitch, $p \wedge \neg Kp$, sendo composta, e assim não básica, não pode substituir a variável em $p \rightarrow \Diamond Kp$. O paradoxo é, assim, evitado.

5.3 Problemas para as Restrições Sintáticas

Os méritos relativos dessas duas restrições são sopesados por Tennant (2002). A restrição de Tennant é a menos exigente entre as duas, uma vez que ela barra somente a substituição daquelas asserções que são logicamente incognoscíveis; e, dessa forma, somente aquelas asserções que são responsáveis pelos paradoxos. A restrição de Dummett, em comparação, barra não somente a restrição dessas proposições, mas também a substituição de proposições logicamente complexas que são claramente cognoscíveis. Tennant também indica que, se o princípio da cognoscibilidade é a motivação antirrealista primária para revisar a lógica clássica, então restringir aquele princípio a asserções básicas pode erodir argumentos contra um tratamento clássico de asserções complexas.

As principais objeções às estratégias restritivas classificam-se em dois campos. No primeiro campo, encontramos a acusação de que uma dada restrição sintática ao princípio da cognoscibilidade não é justificada por um princípio. No segundo campo surgem formulações de paradoxos à la Fitch que não são evitados por essas restrições sintáticas sobre a verdade cognoscível.

Do primeiro campo, Hand e Kvanvig (1999) argumentam que TKP não foi restrito de uma maneira coerente com os princípios—e não foi dado nenhum bom motivo, além da ameaça de paradoxo, para restringir o princípio às asserções cartesianas. Uma asserção análoga foi feita com relação à DKP de Dummett. Tennant (2001b) responde a Hand e Kvanvig com uma discussão geral sobre a admissibilidade das restrições na prática de análise conceitual e clarificação filosófica. Ao traçar analogias entre sua própria restrição e outras que são claramente admissíveis, ele defende que a restrição cartesiana não é *ad hoc*. Ele também indica que TKP, e não o KP irrestrito, serve como o tema de disputa mais interessante entre a realista semântica e a antirrealista. A realista acredita que é possível que a verdade seja incognoscível em princípio. O argumento de Fitch, na melhor das hipóteses, mostra-nos que há incognoscibilidade estrutural, isto é, incognoscibilidade que é uma função somente de considerações lógicas. Mas ainda existe uma forma mais substancial de incognoscibilidade, por exemplo, a incognoscibilidade que é uma função da transcendência de reconhecimento do conteúdo não-lógico? Uma realista que recrimina a natureza *ad hoc* de TKP (ou DKP) falha em se engajar com a teórica da cognoscibilidade no cerne do debate sobre realismo.

Outras reclamações de que a estratégia de restrição de Tennant não é justificada por um conjunto de princípios aparecem em DeVidi & Kenyon (2003), e Hand (2003). Hand oferece uma forma de restringir a cognoscibilidade conforme princípios.

Essas preocupações podem ser dispensadas ao notar versões do paradoxo que não violam as restrições propostas ao princípio da cognoscibilidade. Williamson (2000a) considera o seguinte paradoxo. Seja p a sentença decidível “existe um fragmento de cerâmica romana naquele lugar”. Seja n tal que designa rigidamente o número de livros atualmente sobre a minha mesa agora. Seja E o predicado “é par”. Williamson constrói a conjunção

$$p \wedge (Kp \rightarrow En),$$

que consideramos uma sentença cartesiana. Sabê-la aparentemente não implica uma contradição. Se ele está certo, nós podemos aplicá-la a TKP, resultando em

$$(p \wedge (Kp \rightarrow En)) \rightarrow \Diamond K(p \wedge (Kp \rightarrow En)) \qquad (1.)$$

Adicionalmente, se p é verdadeiro e Kp é falso, então a conjunção de Williamson é verdadeira. Deste modo,

$$(p \wedge \neg Kp) \rightarrow (p \wedge (Kp \rightarrow En)) \tag{2.}$$

As linhas (1) e (2) acarretam

$$(p \wedge \neg Kp) \rightarrow \Diamond K(p \wedge (Kp \rightarrow En)) \tag{3.}$$

Aceitando os modestos recursos epistêmicos encontrados na linha de argumentação de Fitch, pode-se provar o seguinte teorema:

$$K(p \wedge (Kp \rightarrow En)) \rightarrow En. \tag{4.}$$

Eis o porquê. Uma conjunção só é conhecida se suas conjuntas são conhecidas. Dessa forma, se $K(p \wedge (Kp \rightarrow En))$, então Kp. E somente verdades podem ser conhecidas. Então, se $K(p \wedge (Kp \rightarrow En))$, então $(Kp \rightarrow En)$. É claro, Kp e $Kp \rightarrow En$ de forma conjunta acarretam En. Então o teorema 4 é válido, se os recursos epistêmicos de Fitch são válidos. Agora, 4 é um teorema e, desse modo, é verdadeiro em todos os mundos possíveis. Deste modo, seu consequente é possível se o seu antecedente é possível:

$$\Diamond K(p \wedge (Kp \rightarrow En)) \rightarrow \Diamond En. \tag{5.}$$

Das linhas (3) e (5) podemos derivar

$$(p \wedge \neg Kp) \rightarrow \Diamond En. \tag{6.}$$

Uma vez que n designa rigidamente, não é contingente se n é par. Segue da linha (6) que

$$(p \wedge \neg Kp) \rightarrow En. \tag{7.}$$

Um argumento análogo substituindo “par” por “ímpar” nos dá

$$(p \wedge \neg Kp) \rightarrow \neg En. \tag{8.}$$

Mas então nós temos uma contradição sobre TKP e a conjunção de Fitch, $p \wedge \neg Kp$. O resultado envolve as substituições de p por $p \wedge (Kp \rightarrow En)$ e $p \wedge (Kp \rightarrow \neg En)$ em TKP, mas Williamson argumenta que nenhuma viola a restrição cartesiana. Retornamos ao paradoxo.

Tennant (2001a) discorda da asserção de Williamson de que $p \wedge (Kp \rightarrow En)$ é cartesiana. No caso em que n é ímpar, En expressa uma falsidade necessária (por exemplo, "13 é par"). Mas então a linha 4 nos diz que $K(p \wedge (Kp \rightarrow En))$ implica algo que é necessariamente falso. E se a falsidade de "13 é par" é uma questão de necessidade *lógica*, então $p \wedge (Kp \rightarrow En)$ não pode ser consistentemente conhecido, e logo não é cartesiana. Logo, quando n é ímpar, a primeira parte da prova de Williamson (envolvendo o predicado "é ímpar") viola, na verdade, a restrição cartesiana. Em contraste, a conjunção de Williamson é cartesiana quando En é verdadeira. Mas, analogamente, se a verdade de En é uma questão de necessidade lógica, então $p \wedge (Kp \rightarrow \neg En)$ não pode ser consistentemente conhecida e, portanto, não é cartesiana. Logo, quando n é par, a segunda parte da prova de Williamson (envolvendo o predicado "é ímpar") viola a restrição cartesiana. De todo modo, Tennant argumenta, Williamson não mostrou que TKP é um tratamento inadequado do paradoxo de Fitch.

O debate continua em Williamson (2009) e Tennant (2010).

Brogaard & Salerno (2002) desenvolvem outros paradoxos à la Fitch contra as estratégias restritivas. Note que o princípio de cognoscibilidade de Dummett é um bicondicional: $p \leftrightarrow \Diamond Kp$, onde p é básico. Tennant (2002) concorda que o princípio da cognoscibilidade deveria preservar a natureza factiva de $\Diamond K$. Então Brogaard & Salerno começam com o princípio da cognoscibilidade mais forte a seguir:

$$p \leftrightarrow \Diamond Kp, \text{ onde } p \text{ satisfaz a condição sintática relevante.} \qquad \text{(SKP)}$$

Além disso, afora uma discussão mais aprofundada da lógica de K, não é implausível que a teórica intuicionista da cognoscibilidade deseje validar o princípio KK:

$$\Box(Kp \rightarrow KKp). \qquad \text{(KK)}$$

O princípio diz que, necessariamente, se p é conhecido, então é conhecido que p é conhecido. Um outro recurso é utilizado, qual seja, o princípio de fechamento que diz que o antecedente de um condicional necessário é possível somente se o consequente é possível.

Se estes comprometimentos nos são concedidos, é possível derivar o resultado de Fitch sem violar a restrição cartesiana de Tennant:

1.	$p \wedge \neg Kp$	Hipótese (conjunção de Fitch)
2.	$Kp \rightarrow KKp$	de KK
3.	$p \rightarrow \Diamond Kp$	de SKP (esquerda para direita)
4.	$\Diamond Kp$	de 1 e 3
5.	$\Diamond KKp$	de 4 e 2, por fechamento
6.	$\Diamond KKp \rightarrow Kp$	de SKP (direita para esquerda)
7.	Kp	de 5 e 6
8.	$Kp \wedge \neg Kp$	de 1 e 7

SKP é aplicado nas linhas 3 e 6 a p e Kp, respectivamente. E estes substituendos não violam a restrição cartesiana. Nem Kp, nem KKp são auto-contraditórios. Todavia, a antirrealista é forçada, de maneira absurda, a admitir que nenhuma verdade é desconhecida.

Pode-se argumentar que esse resultado ameaça também o princípio da cognoscibilidade restrita de Dummett. Mas isso depende de se aplicamos o princípio apenas às asserções básicas. p é básica, mas a caracterização de Dummett da verdade subdetermina o status de Kp. Talvez ela seja básica, uma vez que Kp não é verofuncionalmente complexa. Todavia, o problema não pode ser resolvido sem uma discussão sobre K.

Brogaard & Salerno demonstram outros paradoxos contra as estratégias de restrição. Esses resultados adicionais não pressupõem um comprometimento com o princípio KK. Eles dependem, em última análise, da interpretação de $\Diamond$ da teórica da cognoscibilidade. Quando $\Diamond$ é uma possibilidade metafísica, ou governada por uma lógica ao menos tão forte quanto $S4$, o princípio da cognoscibilidade forte (devidamente restrito, e tomado como uma tese necessária) implica que não há verdades desconhecidas. Quando $\Diamond$ é possibilidade epistêmica, e o princípio da cognoscibilidade é encarado como uma tese necessária que é conhecida, o princípio da cognoscibilidade implica que, necessariamente, não há asserções indecididas. Diferentemente dos paradoxos da indecidibilidade de Wright (1987), Williamson (1988) e Percival (1990), o raciocínio de Brogaard e Salerno não viola a restrição cartesiana de Tennant. Uma resposta a Brogaard e Salerno aparece em Rosenkranz (2004). Discussões adicionais sobre a restrição cartesiana podem ser encontradas em Broogard & Salerno (2006, 2008). Tennant (2009) é uma defesa e um desenvolvimento adicionais da estratégia cartesiana. Uma defesa da estratégia da restrição pode ser encontrada em Fischer (2013).

Muito do que foi escrito sobre o paradoxo da cognoscibilidade aparece na forma de tentativas de expressar a forma relevante do antirrealismo sem o paradoxo. Propostas incluem Chalmers (2012), Dummett (2009), Edgington (2010), Fara (2010), Hand (2009,2010), Jenkins (2005), Kelp & Pritchard (2009), Linsky (2009), Hudson (2009), Restall (2009), Tennant (2009), Alexander (2013), Dean & Kurokawa (2010), Proietti (2016).

Chalmers (2002, 2012: chap. 2), por exemplo, defende a ideia que, dada suficiente informação qualitativa sobre o mundo, nós poderíamos em princípio saber o valor de verdade de qualquer asserção. Mais especificamente, a sua *tese do escrutínio* diz que, se D é uma descrição qualitativa completa do mundo, então, para todo T, é conhecido *a priori* que D implica (materialmente) T. Notoriamente, o paradoxo da cognoscibilidade não ameaça a asserção de que conjunções de Fitch verdadeiras são deriváveis *a priori* de uma descrição completa do mundo.

Dummett vê $\forall p(p \rightarrow \neg\neg Kp)$ como a melhor expressão da sua forma de antirrealismo e aceita as suas consequências intuicionistas de braços abertos. Edgington defende seu princípio da cognoscibilidade (a saber, se atualmente p, então é possível saber que atualmente p) ao argumentar por algum tipo de cognoscibilidade transmundana—especificamente naqueles casos em que o conhecedor meramente possível compartilha o histórico causal relevante com o mundo atual. Hand defende o antirrealismo ao indicar a distinção entre tipos de verificações e seus tokens de performances, e argumenta que a existência de um tipo de verificação não implica que ela seja performável. A lição aqui é que a antirrealista deve pensar sobre a verdade não tanto em termos da performabilidade de procedimentos de verificação, e mais em termos da existência de tipos de verificação. E Linsky arregimenta os princípios epistêmicos e o raciocínio com uma teoria dos tipos. O debate acerca da melhor caracterização de um antirrealismo semântico irá muito além do escopo deste verbete. Quanto à prova da cognoscibilidade em si, segue não havendo consenso sobre se ela está errada, e onde estaria errada.

Discussões de nicho sobre o paradoxo que não cabem em nenhuma dessas seções incluem o novo paradoxo da felicidade em Salerno (2018); o argumento de Kvanvig (2010) de que o paradoxo ameaça o cristianismo em si, dada a sua doutrina da encarnação de Cristo; e o argumento de Cresto (2017) de que o princípio da cognoscibilidade levanta dúvidas sobre o princípio da reflexão como um requerimento da racionalidade.

Referências Bibliográficas

S. Alexander. An axiomatic version of fitch's paradox. *Synthese*, 190:2015–2020, 2013.

J. C. Beall. Fitch's proof, verificationism, and the knower paradox. *Australasian Journal of Philosophy*, 78:241–247, 2000.

J. C. Beall. Knowability and possible epistemic oddities. In *Proceedings of the conference or edited volume edited by Salerno (2009)*, pages 105–125. Unknown, 2009.

J. Bermüdez. Truth, indefinite extensibility, and fitch's paradox. In *Proceedings of the conference or edited volume edited by Salerno (2009)*, pages 76–90. Unknown, 2009.

F. Berto and P. Hawke. Knowability relative to information. *Mind*, 2018. doi: 10.1093/mind/fzy045. First online, 25 October 2018.

B. Brogaard. Knowability and a modal closure principle. *American Philosophical Quarterly*, 43:261–270, 2006.

B. Brogaard. Knowability, possibility and paradox. In *V. Hendricks and D. Pritchard (eds.), New Waves in Epistemology*. Palgrave Macmillan, New York, 2008.

B. Brogaard. On keeping blue swans and unknowable facts at bay: a case study on fitch's paradox. *In Salerno (ed.) 2009*, pages 241–251, 2009.

B. Brogaard and J. Salerno. Clues to the paradoxes of knowability: Reply to dummett and tennant. *Analysis*, 62:143–150, 2002.

O. Bueno. Fitch's paradox and the philosophy of mathematics. In *Proceedings of the conference or edited volume edited by Salerno (2009)*, pages 252–280. 2009.

J. Burgess. Can truth out? In *Edited Volume*, pages 147–162. Salerno, 2009.

D. J. Chalmers. Actuality and knowability. *Analysis*, 71(3):411–419, 2011.

D. J. Chalmers. *Constructing the World*. Oxford University Press, Oxford, 2012.

David J. Chalmers. Does conceivability entail possibility? In *Gendler, T. and Hawthorne, J. (eds.), Conceivability and Possibility*. Oxford University Press, Oxford, 2002.

J. Chase and P. Rush. Factivity, consistency and knowability. *Synthese*, 195:899–918, 2018.

A. Church. Referee reports on fitch's 'a definition of value'. In *Proceedings of Salerno (ed.) 2009*, pages 13–20. Saleno (2009), 2009.

A. Costa-Leite. Fusions of modal logics and fitch's paradox. *Croatian Journal of Philosophy*, 6:281–290, 2006.

C. Cozzo. What we can learn from the paradox of knowability. *Topoi*, 13:71–78, 1994.

E. Cresto. Lost in translation: Unknowable propositions in probabilistic frameworks.

Synthese, 194:3955–3977, 2017.

W. Dean and H. Kurokawa. From the knowability paradox to the existence of proofs. *Synthese*, 176:177–225, 2010.

D. DeVidi and T. Kenyon. Analogues of knowability. *Australasian Journal of Philosophy*, 81 (4):481–495, 2003.

D. DeVidi and G. Solomon. Knowability and intuitionistic logic. *Philosophia*, 28:319–334, 2001.

M. Dummett. Truth. *Proceedings of the Aristotelian Society*, 59:141–162, 1959.

M. Dummett. The philosophical basis of intuitionistic logic. In *H. Rose and J. Shepherdson*. Amsterdam: North Holland, 1975.

M. Fara. Knowability and the capacity to know. *Synthese*, 173:53–73, 2010.

M. Fischer. Some remarks on restricting the knowability principle. *Synthese*, 190:63–88, 2013.

F. Fitch. A logical analysis of some value concepts. *The Journal of Symbolic Logic*, 28: 135–142, 1963a.

F. Fitch. A logical analysis of some value concepts. *The Journal of Symbolic Logic*, pages 21–28, 1963b.

M. Hand. Knowability and epistemic truth. *Australasian Journal of Philosophy*, 81(2):216–228, 2003.

M. Hand. Performance and paradox. *Salerno (ed.) 2009*, pages 283–301, 2009.

M. Hand. Antirealism and universal knowability. *Synthese*, 173:25–39, 2010.

M. Hand and J. Kvanvig. Tennant on knowability. *Australasian Journal of Philosophy*, 77: 422–428, 1999.

W. D. Hart. The epistemology of abstract objects: Access and inference. *Proceedings of the Aristotelian Society*, 53:153–165, 1979.

W. D. Hart. Invincible ignorance. *Salerno (ed.) 2009*, pages 283–301, 2009.

W. D. Hart and C. McGinn. Knowledge and necessity. *Journal of Philosophical Logic*, 5: 205–208, 1976.

J. Heylen. Factive knowability and the problem of possible omniscience. *Philosophical Studies*.

W. Holliday. Knowledge, time, and paradox: Introducing sequential epistemic logic. In *H. van Ditmarsch and G. Sandu (eds.), Jaakko Hintikka on Knowledge and Game-Theoretical Semantics*, pages 363–394. Springer, Berlin, 2018.

R. Hudson. Faint-hearted anti-realism and knowability. *Philosophia*, 37:511, 2009.

M. Jago. Closure on knowability. *Analysis*, 70:648–659, 2010.

C. Jenkins. Realism and independence. *American Philosophical Quarterly*, 42:199–209, 2005.

C. Jenkins. The mystery of the disappearing diamond. *Salerno (ed.) 2009*, pages 302–319, 2009.

C. Kelp and D. Pritchard. Two deflationary approaches to fitch-style reasoning. *Salerno (ed.) 2009*, pages 324–338, 2009.

N. Kennedy. Defending the possibility of knowledge. *Journal of Philosophical Logic*, 43: 579–601, 2014.

J. Kvanvig. The knowability paradox and the prospects for anti-realism. *Noûs*, 29:481–499, 1995.

J. Kvanvig. *The Knowability Paradox*. Oxford University Press, Oxford, 2006.

J. Kvanvig. Restriction strategies for knowability: Some lessons in false hope. In *Salerno (ed.) 2009*, pages 205–222. 2009.

J. Kvanvig. The incarnation and the knowability paradox. *Synthese*, 173:89–105, 2010.

S. Lindström. Situations, truth and knowability: A situation-theoretic analysis of a paradox of fitch. In E. Ejerthed and S. Lindström, editors, *Logic, Action and Cognition: Essays in Philosophical Logic*, pages 183–210. Kluwer Academic Publishers, Dordrecht, 1997.

B. Linsky. Logical types in arguments about knowability and belief. *Salerno (ed.) 2009*, pages 163–179, 2009.

J. L. Mackie. Truth and knowability. *Analysis*, 40:90–92, 1980.

P. Maffezioli, A. Naibo, and S. Negri. The church-fitch knowability paradox in the light of structural proof theory. *Synthese*, 190:2677–2716, 2013.

J. Melia. Anti-realism untouched. *Mind*, 100:341–342, 1991.

J. Murzi. Knowability and bivalence: intuitionistic solutions to the paradox of knowability. *Philosophical Analysis*, 149:269–281, 2010.

J. Murzi. Manifestability and epistemic truth. *Topoi*, 31:17–26, 2012.

R. Palczewski. Distributed knowability and fitch's paradox. *Studia Logica*, 86:455–478, 2007.

A. Paseau. Fitch's argument and typing knowledge. *Notre Dame Journal of Formal Logic*, 49:153–176, 2008.

P. Percival. Fitch and intuitionistic knowability. *Analysis*, 50:182–187, 1990.

P. Percival. Knowability, actuality and the metaphysics of context-dependence. *Australasian*

Journal of Philosophy, 69:82–97, 1991.

G. Priest. Beyond the limits of knowledge. *Salerno (ed.) 2009*, pages 93–104, 2009.

C. Proietti. The fitch-church paradox and first order modal logic. *Erkenntnis*, 81:87–104, 2016.

C. Proietti and G. Sandu. Fitch's paradox and ceteris paribus modalities. *Synthese*, 173: 75–87, 2010.

W. Rabinowicz and K. Segerberg. Actual truth, possible knowledge. *Topoi*, 13:101–115, 1994.

S. Rasmussen. The paradox of knowability and the mapping objection. *Salerno (ed.) 2009*, pages 53–75, 2009.

S. Rasmussen and J. Ravnkilde. Realism and logic. *Synthese*, 52:379–437, 1982.

G. Restall. Not every truth can be known (at least, not all at once). *Salerno (ed.) 2009*, pages 339–354, 2009.

L. Rosenblatt. The knowability argument and the syntactic type-theoretic approach. *Theoria*, 80:201–221, 2014.

S. Rosenkranz. Fitch back in action again? *Analysis*, 64(1):67–71, 2004.

R. Routley. Necessary limits to knowledge: Unknowable truths. In *Essays in Scientific Philosophy. Dedicated to Paul Weingartner*, pages 93–115. Come-Verlag, Bad Reichenhall, 1981.

H. Rückert. *A Solution to Fitch's Paradox of Knowability*. Kluwer Academic Publishers, 2003.

J. Salerno. Revising the logic of logical revision. *Philosophical Studies*, 99:211–227, 2000.

J. Salerno, editor. *New Essays on the Knowability Paradox*. Oxford University Press, 2009a.

J. Salerno. Knowability noir: 1945–1963. In *Salerno (ed.) 2009*, pages 29–48. 2009b.

W. K. San. Knowability and a new paradox of happiness. In *H. van Ditmarsch and G. Sandu (eds.), Jaakko Hintikka on Knowledge and Game-Theoretical Semantics*, pages 457–474. Springer, 2018.

J. Schlöder. Counterfactual knowability revisited. *Synthese*, 198(2):1123–1137, 2014.

A. Stephenson. Kant, the paradox of knowability, and the meaning of 'experience'. *Philosophers' Imprint*, 15(27):available online, 2015.

N. Tennant. *The Taming of the True*. Oxford University Press, 1997.

N. Tennant. Is every truth knowable? reply to williamson. *Ratio*, XIV:263–280, 2001a.

N. Tennant. Is every truth knowable? reply to hand and kvanvig. *Australasian Journal of Philosophy*, 79:107–113, 2001b.

N. Tennant. Victor vanquished. *Analysis*, 62:135–142, 2002.

N. Tennant. Revamping the restriction strategy. In *Salerno (ed.) 2009*, pages 223–238. 2009.

N. Tennant. Williamson's woes. *Synthese*, 173:9–23, 2010.

J. van Benthem. What one may come to know. *Analysis*, 64(2):95–105, 2004.

J. van Benthem. Actions that make us know. In *Proceedings of the conference or edited volume edited by Salerno (2009)*, pages 129–146. Unknown, 2009.

H. van Ditmarsch, W. van der Hoek, and P. Iliev. Everything is knowable—how to get to know whether a proposition is true. *Theoria*, 78:93–114, 2012.

H. Wansing. Diamonds are a philosopher's best friend: The knowability paradox and modal epistemic relevance logic. *Journal of Philosophical Logic*, 31(6):591–612, 2002.

T. Williamson. Intuitionism disproved? *Analysis*, 42:203–207, 1982.

T. Williamson. On the paradox of knowability. *Mind*, 1987.

(II) Paradoxo de Sorites[1]

Título Original: Sorites Paradox
Autores: Dominic Hyde e Diana Raffman
Tradução: Daniel Alves da Silva Lopes Diniz
Revisão: Alan R. Antezana

O paradoxo de sorites se originou em um antigo quebra-cabeça que parece ser gerado por termos vagos, ou seja, termos com limites de aplicação indefinidos ("borrados" ou "difusos"). "Calvo", "amontoado", "alto", "velho", e "azul" são exemplos paradigmáticos de termos vagos: não há uma linha clara que separe as pessoas que são calvas das que não são, ou objetos azuis de verdes (portanto, não azuis), ou pessoas velhas de pessoas de meia-idade (portanto, não velhas). Como o predicado "amontoado" tem limites indefinidos, parece que nenhum grão de trigo em particular pode fazer a diferença entre um número de grãos que constitui um amontoado e um número que não constitui. Portanto, já que um grão de trigo não forma um amontoado, segue-se que dois grãos também não formam; se dois não formam, então três não formam; e assim por diante. Esse raciocínio leva à conclusão absurda de que nenhum número de grãos de trigo forma um amontoado.

A mesma forma de raciocínio é familiar à vida cotidiana. Dorothy Edgington observa (Edgington, 1996, p. 296):

> Existe o paradoxo do amanhã: a tarefa desagradável que precisa ser feita,

[1]HYDE, Dominic; RAFFMAN, Diana, "Sorites Paradox", In: ZALTA, E. N. (ed.). The Stanford Encyclopedia of Philosophy (Summer 2018 Edition). Stanford, CA: The Metaphysics Research Lab, 2018. Disponível em: https://plato.stanford.edu/archives/sum2018/entries/sorites-paradox/.
A seguir está a tradução da entrada sobre o Paradoxo de Sorites de Dominic Hyde e Diana Raffman na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/sum2018/entries/sorites-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/sorites-paradox/. Agradecemos ao Prof. Dr. Edward N. Zalta pela permissão para traduzir e publicar esta entrada.

mas tal que é sempre indiferente se ela é feita hoje ou amanhã; o paradoxo da pessoa em dieta: eu não me importo com a diferença para o meu peso que um chocolate fará.

O quebra-cabeça pode ser expresso como um argumento do modo mais simples usando *modus ponens*:

- 1 grão de trigo não forma um amontoado.
- Se 1 grão não forma um amontoado, então 2 grãos não formam um amontoado.
- Se 2 grãos não formam um amontoado, então 3 grãos não formam um amontoado.

…

- Se 999.999 grãos não formam um amontoado, então 1 milhão de grãos não formam um amontoado.

Portanto,

- 1 milhão de grãos não formam um amontoado.

O argumento é um paradoxo porque um raciocínio aparentemente impecável a partir de premissas aparentemente impecáveis gera uma falsidade. A partir da premissa de que um milhão de grãos formam um amontoado, o argumento pode igualmente ir na direção oposta: se um milhão de grãos formam um amontoado, então um milhão de grãos menos um formam um amontoado, se um milhão de grãos menos um formam um amontoado, então um milhão de grãos menos dois formam um amontoado, etc. Segue-se de forma absurda que até mesmo um único grão forma um amontoado. Logo, o raciocínio sorítico parece mostrar tanto que nenhum número de grãos forma um amontoado quanto que qualquer número de grãos forma um amontoado.

Que conclusão devemos tirar desse resultado inconveniente? Há algo errado com o argumento paradoxal, ou o uso de predicados vagos realmente leva ao absurdo?[2] Em parte porque usamos essas palavras comuns exitosamente o tempo todo, e normalmente não

[2]Outras classes de palavras—por exemplo, verbos, advérbios, nomes, e mesmo indexicais—também parecem ser soríticas, mas os predicados receberam mais atenção na literatura. Confira, por exemplo, Ellis (2004) para uma discussão relevante.

incorremos em absurdos como esses acima, a maioria dos teóricos da vagueza supõe que o paradoxo é solucionável, ou seja, que o argumento paradoxal é defeituoso e podemos descobrir o defeito. No que se segue, consideramos algumas das principais tentativas de resolvê-lo.

1. O sorites na história

O filósofo megárico Eubulides (século IV a.C.) é geralmente creditado pela primeira formulação do quebra-cabeça (o nome "sorites" deriva da palavra grega *soros*, que significa "amontoado"). Embora não saibamos suas motivações para introduzi-lo (junto com vários outros quebra-cabeças lendários), o paradoxo foi mais tarde usado por filósofos gregos como uma arma dialética, mais notavelmente pelos céticos contra as alegações de conhecimento dos estoicos.

Curiosamente, o paradoxo atraiu pouco interesse subsequente até o final do século XIX. Filósofos marxistas na tradição neo-hegeliana, como Plekhanov (1908, [1937], p. 114), citaram o paradoxo como evidência tanto do fracasso da lógica "costumeira", quanto da utilidade da "lógica da contradição". Dessa forma, alguns marxistas buscaram estabelecer o triunfo da dialética. Enquanto isso, na filosofia anglo-americana, a lógica formal recuperou seu lugar central, e sua formalização clássica não deixou espaço para a vagueza da linguagem natural. A vagueza e o paradoxo associado foram vistos como excedendo o escopo da lógica e, portanto, não representando nenhum desafio a ela. No entanto, desde o fim das doutrinas de linguagem ideal da segunda metade do século XX (ver §3.1), o interesse nas idiossincrasias da linguagem natural, incluindo sua vagueza, aumentou muito.

2. Diferentes formulações do paradoxo

Pelo menos três condições devem ser satisfeitas para que um argumento seja uma instância do paradoxo de sorites. (1) Deve ser possível construir uma série de sorites para o predicado em questão, a saber, uma ordenação finita de valores em uma dimensão decisiva para a aplicação do predicado. Uma série de sorites para "alto" é uma ordenação na dimensão de altura (uma ordenação de alturas); para "velho", uma ordenação na dimensão de idade (uma ordenação de idades), e assim por diante. (2) Valores adjacentes na série devem ser apenas incrementalmente diferentes, isto é, ou indiscrimináveis, ou apenas ligeiramente diferentes. Uma diferença incremental deveria garantir que, se um predicado vago aplica-se a um objeto de um par de vizinhos, aplica-se igualmente ao outro (como em Wright—por exemplo, Wright (1975)—a propriedade de se aplicar ao longo de diferenças incrementais em uma dimensão decisiva é frequentemente chamada de tolerância de um termo vago). (3) O predicado deve ser verdadeiro para o primeiro valor na série e falso para o último.

O paradoxo é frequentemente apresentado na forma condicional discutida acima. Mais formalmente: seja Φ um predicado sorítico e seja α_n (onde n é um número natural) a representação de um valor em uma série de sorites para Φ. Então o paradoxo pode ser representado da forma mais simples deste modo, usando *modus ponens*:

$$\text{Sorites condicional:}\quad \begin{array}{l} \Phi_{\alpha_1} \\ \text{Se } \Phi_{\alpha_1} \text{ então } \Phi_{\alpha_2} \\ \text{Se } \Phi_{\alpha_2} \text{ então } \Phi_{\alpha_3} \\ \text{etc.} \\ \text{Se } \Phi_{\alpha_{n-1}} \text{ então } \Phi_{\alpha_n} \\ \hline \Phi_{\alpha_n} \;\; \text{(onde } n \text{ pode ser arbitrariamente grande)} \end{array}$$

Uma formulação diferente do paradoxo substitui o conjunto de premissas condicionais por uma generalização universal e procede por indução matemática. Seja n uma variável sobre os números naturais, e seja $\forall n(\ldots n \ldots)$ a afirmação de que todo número n satisfaz a condição $(\ldots n \ldots)$. Além disso, representemos a afirmação "para todo n, se α_n é Φ então α_{n+1} é Φ" como "$\forall n(\Phi_{\alpha_n} \rightarrow \Phi_{\alpha_{n+1}})$".

$$\text{Sorites de indução matemática:}\quad \begin{array}{l} \Phi\alpha_1 \\ \forall n(\Phi\alpha_n \rightarrow \Phi\alpha_{n+1}) \\ \hline \forall n(\Phi\alpha_n) \end{array}$$

Por exemplo, como um homem com 1 fio de cabelo na cabeça é calvo, e como, para qualquer número n de fios de cabelo, se um homem com n fios de cabelo é calvo, então também o é um homem com $n+1$ fios de cabelo, então todo número n é tal que um homem com n fios de cabelo na cabeça é calvo.

Outra versão do quebra-cabeça é a variante de forma indutiva. Sabemos que uma série de sorites para "calvo" contém alguns números de cabelos tais que homens com aqueles números de cabelo não são calvos. Pelo princípio de menor número (equivalente ao princípio de indução matemática), deve existir um número mínimo, digamos $i+1$, tal que um homem com $i + 1$ cabelos na cabeça não é calvo. Já que um homem com 1 cabelo na cabeça é calvo, segue-se que $i + 1$ deve ser maior que 1. Assim, a série contém um número de cabelos n $(= i)$ tal que um homem com n cabelos é calvo, ao passo que um homem com $n + 1$ cabelos não o é. Seja $\exists(\ldots n \ldots)$ a asserção de que um número n satisfaz a condição $(\ldots n \ldots)$. Esquematizamos então esse último raciocínio desta forma:

$$\text{Sorites de traçamento de linha:}\quad \begin{array}{l} \Phi\alpha_1 \\ \sim\forall n(\Phi\alpha_n) \\ \hline \exists n \geq 1(\Phi\alpha_n \,\&\, {\sim}\Phi\alpha_{n+1}) \end{array}$$

As formas de desenho de linha e indutiva do quebra-cabeça ilustram bem o impasse sorítico; aparentemente, usuários competentes de "calvo" devem, e não devem, traçar uma linha na série. Por conveniência, no que se segue, a maioria dos exemplos é colocada em termos das formas condicional ou indutiva do paradoxo. Certamente, uma resolução adequada do sorites presumivelmente precisará desarmar todas as suas versões.

Devemos mencionar também uma versão informal do paradoxo, conhecida como "sorites de marcha forçada" (Horgan (1994a); Soames (1999)). Aqui ele é descrito em termos de classificações hipotéticas que seriam feitas por uma falante competente percorrendo uma série de sorites passo a passo. Uma falante competente deve dizer que um único grão de trigo não forma um amontoado; mas se isso está correto, então ela também deve dizer que dois grãos não formam um amontoado; e que três grãos não formam; e assim por diante, até ela dizer que um milhão de grãos, por exemplo, não formam um amontoado. Como veremos, o paradoxo da marcha forçada desempenha um papel importante em vários tratamentos do paradoxo.

Vale a pena notar que a definição popular de vagueza em termos de soriticalidade (por exemplo, Wright (1976); Bueno and Colyvan (2012)) pode muito bem estar incorreta. Se o sorites é uma falácia solucionável, como acredita a maioria das teóricas da vagueza, então a vagueza não é, afinal, uma fonte de paradoxo. Talvez alguém diga que, mesmo depois que o diagnóstico correto do quebra-cabeça tenha sido descoberto, o argumento permanecerá um paradoxo porque ainda *parecerá* consistir em um raciocínio incontestável a partir de premissas verdadeiras para uma conclusão falsa. Mas essa visão torna a vagueza uma propriedade contingente demais; dado tudo que sabemos, uma vez que tenhamos descoberto a solução adequada para o quebra-cabeça, a premissa maior não parecerá verdadeira mais. Pode parecer verdadeira aos não iniciados, mas esta também seria uma maneira duvidosa de definir a vagueza—a saber, como a propriedade de gerar um argumento que anteriormente parecia, ou parece aos não iniciados, paradoxal. Bueno and Colyvan (2012) dizem que "um predicado é vago apenas no caso de poder ser empregado para gerar um argumento de sorites". Mas o que "pode ser empregado" significa aqui? Se um argumento de sorites é uma falácia, um predicado vago não pode ser *corretamente* empregado nele. O critério seria que um predicado vago é um termo que, quando empregado incorretamente, temporariamente parece (aos não iniciados?) gerar um paradoxo de sorites (Raffman, 2014, p. 18–19)?

O mais provável é que a soriticalidade seja uma característica ilusória de palavras como "velho' e "rico"; sua vagueza é real. Se isso estiver correto, então a soriticalidade aparente pode ser melhor vista como um sintoma temporário de vagueza, ou talvez como um elemento de sua "caracterização superficial" (Smith (2008), por exemplo, p. 132; veja o terceiro capítulo de Smith para uma visão perspicaz sobre o que é vagueza).

3. Respostas ao paradoxo

Como com qualquer paradoxo, quatro tipos gerais de respostas parecem estar disponíveis. Alguém poderia:

1. Negar que a lógica aplica-se a expressões soríticas.

Alternativamente, alguém poderia aceitar que o paradoxo é um argumento legítimo ao qual a lógica se aplica, mas negar sua corretude ao

2. Rejeitar alguma(s) premissa(s), ou

3. Negar que ele seja válido.

   A resposta mais drástica seria

4. Acolher o paradoxo e concluir que expressões vagas são ou incoerentes ou vazias.

No que se segue, consideramos os principais tratamentos filosóficos do sorites e as formas com que eles usaram essas estratégias para desfazer o quebra-cabeça.

3.1 Abordagens de linguagem ideal

Comprometidos como estavam com doutrinas de linguagem ideal, não é surpreendente encontrar Frege e Russell (veja as entradas sobre Gottlob Frege e Bertrand Russell) buscando uma resposta do tipo (1) (por exemplo, Frege (1903); Russell (1923)). Diz-se que um atributo-chave da linguagem ideal é sua precisão; portanto, a vagueza da linguagem natural, incluindo todas as expressões soríticas, é um defeito a ser eliminado. Se isso estiver correto, então, ao contrário do que muitas teóricas acreditam, os termos soríticos não podem ser coligidos para desafiar a lógica clássica. A lógica simplesmente não se aplica a eles. Ecoando essa resposta, Quine (1981, p. 31–37) argumenta que, embora a eliminação de termos vagos possa incorrer em algum custo para as formas comuns de falar, é um custo que vale a pena pagar na medida em que nos permite preservar a "doce simplicidade" da lógica clássica.

No entanto, com o fim das doutrinas de linguagem ideal e o subsequente renascimento do interesse pela linguagem comum, a vagueza não é mais considerada uma propriedade superficial ou facilmente dispensável. Se a lógica tem alguma força, ela precisa se aplicar à linguagem natural como ela é; expressões soríticas são inevitáveis e o paradoxo deve ser encarado de frente. Respostas do tipo (2) fazem exatamente isso e são a família mais comum de respostas. A lógica é vista como aplicável à linguagem natural, em particular ao argumento paradoxal, e este último é diagnosticado como baseado em uma premissa falha.

3.2 A teoria epistemicista

A maioria das teóricas da vagueza concebe a vagueza como um fenômeno semântico, de alguma forma oriundo dos significados de palavras como "alto" e "velho". Como veremos, teorias semânticas tipicamente introduzem lógicas e/ou semânticas especiais não clássicas para resolver o paradoxo (e acomodar o fenômeno de casos limítrofes). Em contraste, as epistemicistas pensam que a vagueza é apenas uma forma de ignorância: termos vagos têm limites nítidos cujas localizações nos estão ocultas. Amontoados são, de fato, nitidamente separadas dos não-amontoados, e alturas altas são nitidamente separadas das alturas médias, mas não podemos descobrir onde essas divisões estão (ver, por exemplo, Sorensen (1988, 2001); Williamson (1994, 1995, 2000); Fara (2000, 2008); Rescher (2009)). Nessa visão, o paradoxo sorites é eliminado imediatamente: a premissa maior, ou uma das premissas condicionais, é simplesmente falsa. E a bivalência é preservada: qualquer aplicação de uma expressão vaga é ou verdadeira ou falsa, embora nem sempre possamos saber qual.

Quais fatos sobre o mundo, ou a linguagem natural, ou falantes competentes, poderiam servir para fixar limites nítidos para palavras vagas? De acordo com Williamson (por exemplo, Williamson (1995, p. 184)), o significado sobrevém ao uso; em outras palavras, as localizações dos limites nítidos de uma expressão vaga são uma função das disposições dos falantes de usá-la como usam. Na medida em que o uso de uma expressão vaga varia ao longo do tempo, seus limites podem ser instáveis. Claro, não podemos conhecer a totalidade dessas disposições, e não conhecemos a função relevante; e nossa ignorância desses fatores interdita um caminho para o conhecimento das localizações dos limites da expressão.

Outra rota possível para o conhecimento das localizações dos limites é interditada pelo fato de que nosso conhecimento da aplicação de uma expressão vaga é inexato. O conhecimento inexato é governado por princípios de margem de erro, a saber, princípios da forma "se x e y diferem incrementalmente em uma dimensão decisiva e x é sabidamente Φ (velho, azul, etc.), então y é Φ".[3] Por exemplo, quando o conhecimento é inexato, podemos saber de um objeto azul que ele é azul somente se objetos cujas cores são incrementalmente diferentes também forem azuis—portanto, somente em casos claros. Em contraste, na região limítrofe ou "penumbral" de uma série de sorites para "azul", onde o limite se encontra, um certo tom de azul é apenas incrementalmente diferente de (e na verdade pode parecer o mesmo que) um tom que não é azul; e não podemos saber onde se encontra essa diferença. Consequentemente, se classificarmos o tom anterior como azul, essa classificação está correta por sorte e, portanto, não constitui conhecimento. Na suposição plausível de que ver que um certo x é azul é suficiente para saber que x é azul, segue-se que algumas coisas azuis são tais que não podemos ver que são azuis, mesmo sob condições ideais de visão.

As virtudes e o apelo da teoria epistêmica são significativos, e ela angariou sua parcela de apoiadores. Ao mesmo tempo, essa visão pode ser difícil de aceitar. Até mesmo suas proponentes admitem que o epistemicismo é intuitivamente implausível, e parece multiplicar mistérios. Como uma primeira aproximação, a epistemicista diz que

> as expressões vagas têm limites nítidos e desconhecidos que são fixados por uma função desconhecida de seus padrões de uso incognoscíveis (ou seja,

[3]Williamson (1992, p. 161) afirma que, ao contrário de princípios de tolerância, princípios de margem de erro não geram um paradoxo.

> não totalmente conhecíveis).

No entanto, parece que a função também deve ser incognoscível, não apenas desconhecida; pois como poderíamos reconhecê-la se nos deparássemos com ela? Como poderíamos dizer se tínhamos obtido a função correta, senão determinando se ela gera os limites corretos como seus valores? Se isso estiver certo, então a tese da epistemicista deve ser, na verdade, que

> termos vagos têm limites nítidos e desconhecidos que são fixados por uma função incognoscível de seus padrões de uso incognoscíveis (Raffman, 2014, p. 10).

Certamente, explicações são fornecidas para nossa ignorância irremediável nesses casos: por exemplo, não podemos saber onde estão os limites nítidos porque nosso conhecimento é inexato, e não podemos saber o padrão total do uso da expressão porque "os dados são infinitos" (Williamson, 1995, p. 184–185), e assim por diante. Não obstante, o epistemicismo pode soar como uma história do tipo "é assim porque é assim".[4] Veja §5.1 para uma discussão mais aprofundada sobre Williamson.

Graff Fara defende uma linha diferente de epistemicismo (Fara, 2000, 2008). Åkerman and Greenough (2010, p. 277) observam que a perspectiva dela

> é uma forma de epistemicismo em que predicados vagos traçam fronteiras nítidas e bivalentes.[5] Ao contrário do epistemicismo de Sorensen (1988) e Williamson (1994, 1995), no entanto, é constitutivo da vagueza que a fronteira possa se deslocar em função de mudanças [nos interesses dos falantes].[6]

---

[4] A defesa do epistemicismo de Williamson apoia-se fortemente na afirmação de que outras resoluções do paradoxo são insustentáveis, mas alguns dos seus principais argumentos negativos são vistos como petições de princípio; ver, por exemplo, Wright (1995, p. 135–8) e Raffman (2014, p. 95–96), para discussão.

[5] Fara (2000, p. 70) escreve: "nas conversas em que temos oportunidade de usar expressões vagas como "alto" […] será simplesmente um fato bruto que haverá uma altura mínima […] da qual é verdade dizer que [é alta]… Qualquer quantidade menor [de altura] simplesmente não pode ser a mesma para quaisquer propósitos que estejam em vigor".

[6] Em Fara (2008), Graff Fara explica que a caracterização de sua visão em Fara (2000) era enganosa. Na verdade, os limites mudam em função de mudanças em "o que melhor satisfaz nossos interesses—dado nosso interesse permanente em eficiência resultante de […] uma mudança no que está sendo ativamente considerado" (Fara, 2008, p. 328). Stanley (2003) manifesta dúvidas quanto à ideia de que predicados vagos expressam propriedades relativas a interesses; Graff Fara responde em Fara (2008).

Esse limite nítido é incognoscível entre outras coisas porque está constantemente se movendo em uma série de sorites, mudando de localização conforme os interesses do falante tal que ele nunca o encontra (Fara, 2008, p. 328). Como Stanley (2003, p. 269) coloca,

> quando procuramos por [uma] fronteira da extensão de [uma expressão vaga] em sua penumbra, nossa própria procura tem o efeito de mudar a [extensão] da expressão vaga de modo que a fronteira não esteja onde estamos procurando.

Assim, nunca podemos descobrir onde está a fronteira, e cada premissa condicional parece verdadeira enquanto estamos considerando-na (o papel da relatividade de interesse no relato de Graff Fara é discutido mais adiante em §3.3.4). A retenção da lógica clássica e da bivalência é supostamente uma das principais vantagens da abordagem epistêmica em relação às outras perspectivas (por exemplo, Williamson (1992, p. 162)). Com efeito, como é amplamente assumido que a bivalência implica fronteiras nítidas, muitas teóricas da vagueza acreditam que, para todos os efeitos, o epistemicismo é a única teoria que pode empregar uma semântica bivalente (por exemplo, Rosenkranz (2003); Keefe (2000)).[7] Em particular, elas acreditam que nenhuma teoria semântica da vagueza pode ser clássica. Desenvolvimentos subsequentes lançam dúvidas sobre essa visão, no entanto; veja §3.3.5.

3.3 Abordagens semânticas

Como indicado acima, a vagueza é geralmente considerada uma característica semântica da linguagem. E se é uma característica semântica, sua lógica e/ou semântica não podem ser clássicas, ou assim diz o pensamento padrão. Começando na parte final do século XX, uma série de lógicas e semânticas não clássicas foram desenvolvidas para expressões vagas, cada uma avançando sua própria solução do paradoxo de sorites. O tamanho da inovação lógica proposta varia.

A maioria das teorias semânticas da vagueza e tratamentos semânticos do sorites concebe a aplicação de um termo vago como indeterminada em uma certa gama de casos. Especificamente, em uma série de sorites para o predicado vago "Φ", diz-se que é indeterminado—não há "nenhuma questão de fatos"—qual valor é o último valor que é Φ. Geralmente se pensa que a indeterminação manifesta-se na posse do predicado de (possíveis) casos limítrofes. Os casos limítrofes são concebidos de várias maneiras: como nem

[7] Mas veja a referência a Burns (1991) em §3.3.4.

definitivamente (ou determinadamente) Φ, nem definitivamente não Φ, ou como tais que a sentença "x é Φ" não é nem verdadeira nem falsa, nem super-verdadeira nem super-falsa, ou nem verdadeira até o grau 1 nem falsa até o grau 1, por exemplo.[8] A ideia em comum parece ser que as regiões de casos limítrofes em uma série de sorites para "Φ" constituem os limites borrados do predicado; e que, como a série contém esses valores indeterminados, a premissa principal (ou uma ou mais premissas condicionais) do paradoxo é ou menos que verdadeira, ou simplesmente falsa. A seguir, analisamos alguns dos principais tratamentos semânticos do paradoxo.

3.3.1 Supervaluacionismo

Em consonância com um princípio de menor mutilação, uma abordagem adapta a semântica de supervaloração de van Fraassen (1966) ao paradoxo sorites, e à vagueza de forma mais geral (por exemplo, Fine (1975); Keefe (2000)). Como resultado, ela endossa uma lógica não bivalente que, pelo menos à primeira vista, retém a relação de consequência clássica e as leis clássicas, ao passo que admite lacunas de valor de verdade. Nessa visão, o desafio colocado pelo paradoxo de sorites pode ser enfrentado pela revisão lógica apenas na metateoria, e uma resposta do tipo (2) é defendida. Diferentemente da concepção epistêmica de vagueza, uma concepção semântica tratará como real a aparente indeterminação semântica de predicados vagos. Casos limítrofes são valores aos quais o predicado nem definitivamente se aplica nem definitivamente não se aplica, onde "definitivamente" recebe uma análise semântica ao invés de epistêmica. A extensão positiva de um predicado é dada por aqueles valores aos quais o predicado definitivamente se aplica, a extensão negativa por aqueles valores aos quais o predicado definitivamente não se aplica, e os casos restantes (penumbrais) são valores aos quais o predicado nem definitivamente se aplica, nem definitivamente não se aplica. Consistentemente com uma visão da vagueza como uma deficiência semântica (por exemplo, Fine (1975)) ou como indecisão semântica (por exemplo, Lewis (1986)), as supervaluacionistas definem uma noção de "superverdade" ("superfalsidade") como o estado de ser verdadeiro (falso) independentemente de como a deficiência ou indecisão semântica é resolvida ou precisificada, ou seja, verdadeiro (falso) em cada precisificação do predicado. Aplicar o predicado a algo em sua extensão positiva resulta em

[8] As epistemicistas representam casos limítrofes da maneira usual, como nem definitivamente Φ nem definitivamente não Φ, mas eles interpretam o operador de definição epistemicamente: casos limítrofes não são nem cognoscivelmente Φ nem cognoscivelmente não Φ.

uma sentença superverdadeira, enquanto aplicá-lo a algo em sua extensão negativa resulta em uma sentença superfalsa. Tornar equivalentes superverdade e verdade *simpliciter*, e superfalsidade e falsidade *simpliciter*, resulta então em uma lógica não bivalente com casos limítrofes que originam lacunas de valor de verdade.

Com validade então definida da maneira usual como preservação da verdade (*simpliciter*), a explicação supervaloracionista de validade coincide com a validade clássica. Em particular, tratando leis como argumentos com zero premissas, o supervaloracionismo preserva todas as leis clássicas. Assim, apesar de seu abandono da bivalência, o supervaluacionismo valida a lei do terceiro excluído. Por exemplo, independentemente da vagueza de "amontoado", é logicamente verdadeiro para qualquer número de grãos de trigo que ele ou forma, ou não forma, um amontoado. Como consequência, a semântica de supervaloração não é verofuncional. Ela permite instâncias de disjunções verdadeiras, nenhuma das quais é (super)verdadeira. A conjunção e o condicional exibem características não clássicas análogas.

Como todas as formas assumidas pelo paradoxo sorites são classicamente válidas, elas também são supervaluacionisticamente válidas. A conclusão da forma condicional usando *modus ponens* é evitada ao notarmos que alguma premissa condicional falha em ser verdadeira; embora, reconhecidamente, nenhuma seja falsa. O sorites condicional é válido, mas incorreto. Mais revelador é o diagnóstico da versão que emprega uma premissa maior universal. Esta versão também é considerada incorreta devido à falha de uma das premissas—a premissa universal. O condicional universalmente quantificado não é verdadeiro; na verdade, é falso. Embora não haja nenhuma premissa condicional que seja falsa, é, no entanto, verdade segundo a teoria da supervaloração que algum condicional é falso. Ou seja, é verdade que algum n é tal que não é o caso de que se Φ_{α_n}, então $\Phi_{\alpha_{n+1}}$ (onde "Φ" é sorítico em relação aos sujeitos da forma α_n).

Como a semântica de supervaloração admite que a falsidade de "$\forall n(\Phi_{\alpha_n} \rightarrow \Phi_{\alpha_{n+1}})$" é logicamente equivalente à verdade de "$\exists n(\Phi_{\alpha_n} \;\&\; {\sim}\Phi_{\alpha_{n+1}})$", a forma de linha traçada do sorites é correta: é supervaluacionisticamente válida, pois classicamente válida, e suas premissas são incontestavelmente verdadeiras. O que a semântica de supervaloração alega fornecer é uma descrição formal de como, ao contrário das aparências, tal conclusão poderia ser verdadeira; é verdadeira porque é verdadeira não importa como se resolva a indeterminação do termo vago envolvido (ou seja, o predicado sorítico).

Dessa forma, os paradoxos de sorites são considerados resolvidos. Com a vagueza vista como um fenômeno semântico, a semântica clássica não é mais apropriada como semântica de uma linguagem vaga, e a semântica de supervaloração é proposta em seu lugar. Uma preocupação imediata diante dessa solução é o fato de que ela, em última análise, trata a indução matemática e as formas de linha traçada do sorites da mesma maneira que a teoria epistêmica, conservadora logicamente, o faz. Somos forçados a aceitar a verdade sabidamente contraintuitiva de "$\exists n(\Phi_{\alpha_n} \& \sim\Phi_{\alpha_{n+1}})$" que parece postular a existência de uma fronteira nítida, embora a existência de tal fronteira seja exatamente o que a teoria semântica da vagueza pretende negar. Supervaluacionistas respondem negando que a conclusão do sorites de linha traçada expresse a existência de uma fronteira nítida. Embora comprometidos com a afirmação expressa por a precisão semântica é capturada devidamente apenas pela expressão e isso é claramente negado pela teoria supervaluacionista. Embora seja verdade que existe algum ponto de corte, não existe um ponto em particular em relação ao qual é verdade que ele é o ponto de corte. Uma vez que apenas o último tipo de ponto de corte é considerado um limite nítido, nenhum compromisso é feito com tal limite sobre o qual somos ignorantes (em oposição à teórica epistemicista).

Com essa explicação, no entanto, surgem dúvidas quanto à adequação da lógica. Não apenas (b) deve ser encarado apropriadamente como representando a precisão semântica de "Φ", mas também devemos estar preparados para admitir que algumas afirmações existenciais podem ser verdadeiras sem ter nenhuma instância verdadeira, bloqueando assim qualquer inferência de (a) para (b). Assim como a falha do princípio metateórico da bivalência em conjunto com a retenção da lei do terceiro excluído compromete a supervaluacionista com a presença de disjunções verdadeiras sem disjuntos verdadeiros, da mesma forma devemos permitir um comportamento não padrão análogo na teoria de quantificação da lógica. Com efeito, os aspectos contraintuitivos da teoria epistemicista são evitados apenas mediante um custo para outras intuições.

Neste ponto, a supervaluacionista pode procurar explicar essas anomalias semânticas mostrando como elas são exigidas por uma compreensão adequada do fenômeno subjacente da vagueza. Mais precisamente, a sugestão é que uma visão da vagueza como meramente semântica, não refletindo nenhum fenômeno subjacente de vagueza metafísica (ou seja, uma visão da vagueza como meramente representacional) pode embasar uma abordagem supervaluacionista. Fine (1975) parece promover essa visão representacional ao defender a lei do terceiro excluído, por exemplo, e Varzi (2001), entre outros, também defendem o supervaluacionismo dessa forma. Se bem-sucedida, tal defesa também forneceria uma justificativa virtuosa da ligação comum *de facto* entre a teoria do supervaluacionismo e uma visão representacional da vagueza. Se é essa a explicação a ser buscada, então a maquinaria formal do supervaluacionismo resolve o paradoxo apenas em conjunção com uma negação da vagueza metafísica. O debate metafísico está em andamento. Keefe (2000), por outro lado, opta por uma defesa pragmática arriscada: embora contraintuitivas, as anomalias semânticas que afligem o supervaluacionismo devem ser aceitas porque integram uma teoria que, no geral, se sai melhor do que qualquer outra; nenhuma defesa adicional é necessária.

Williamson (1994) aponta dois problemas adicionais que aparentemente afligem a explicação supervaluacionista. Primeiro, inferências clássicas como demonstração condicional, dilema construtivo e *reductio ad absurdum* não vigoram mais em uma linguagem estendida para expressar vagueza pela adição de um operador de determinação "D" ou similar. A lógica da linguagem estendida é decididamente não clássica. Dummett (1975) oferece uma definição alternativa de validade que não se depara com esse problema, mas Williamson levanta outras objeções a ela. No entanto, Graff Fara [2003] mostra que, se fortalecermos a noção de consequência obtendo consequência penumbral, obtemos falhas desses princípios mesmo na ausência de um operador de determinação. Segundo, surgem problemas também com relação ao fenômeno da vagueza de ordem superior. Ao acomodar a vagueza de ordem superior, a supervaluacionista deve admitir que o conceito de verdade que propõe (a saber, superverdade), carece de propriedades que usualmente se pensa que a verdade tem. Ao contrário das alegações das supervaluacionistas, então, verdade não é superverdade (veja Keefe (2000) para uma refutação).

3.3.2 Parentes do supervaluacionismo

Algumas críticas ao supervaluacionismo são feitas a partir de posições mais próximas da própria perspectiva da supervaluacionista, compartilhando algumas de suas percepções centrais, mas também abandonando outras. Embora concordem com a resposta do tipo (2) defendida pelas supervaluacionistas, Burgess and Humberstone (1987) discordam da retenção (que é muito discutida), por parte da teoria, da lei do terceiro excluído, adotando em vez disso uma variante da lógica supervaluacionista que abandona a lei clássica perante aparentes contraexemplos fornecidos pela vagueza. Para uma discussão e crítica a partir de uma perspectiva supervaluacionista, confira o capítulo 7 de Keefe (2000).

Outra variante do supervaluacionismo é a "lógica discussiva" paraconsistente de Jaśkowski (veja a entrada sobre lógica paraconsistente) que endossa uma resposta do tipo (3) ao sorites condicional. Uma década antes de Mehlberg (1958) propor pela primeira vez o que era, com efeito, um tratamento supervaluacionista da vagueza, um aluno de Łukasiewicz (veja a entrada), Stanisław Jaśkowski, publicou uma descrição de uma lógica que ele propusera como uma lógica dos conceitos vagos. Ela foi, na verdade, o primeiro sistema formal de lógica paraconsistente. Curiosamente, tanto Mehlberg quanto Jaśkowski eram alunos da Escola de filosofia de Lvóv-Varsóvia (veja a entrada) onde Łukasiewicz era professor. Abordagens paraconsistentes ao paradoxo de sorites vinham sendo defendidas por marxistas por algum tempo, com predicações de casos limítrofes fornecendo exemplos paradigmáticos de situações dialéticas. O paradoxo era comumente citado como evidência da inadequação da lógica clássica; mas foi somente com o trabalho pioneiro de Jaśkowski que a proposta recebeu explicação formal. Essa lógica, às vezes agora chamada de "subvaluacionismo" para enfatizar sua dualidade com o supervaluacionismo, que é mais familiar, representa a indeterminação semântica postulada como sobredeterminação semântica, em vez da subdeterminação típica das respostas de lacuna de valor de verdade ao fenômeno da vagueza. Ao argumentar por uma semântica supervaluacionista para vagueza, Fine (1975) observou que a abordagem de excesso de valor de verdade (subvaluacionista) pode ser alcançada por uma simples reinterpretação da abordagem de lacuna de valor de verdade defendida nela. Para mais informações sobre esse sistema e uma defesa limitada dele, veja Hyde (1997). Para críticas, veja o capítulo 7 de Keefe (2000), e Beall and Colyvan (2001).

3.3.3 Teorias de grau e multivaloradas

Em contraste com as lógicas não verofuncionais delineadas acima, várias lógicas não clássicas verofuncionais foram propostas, e em particular, lógicas multivaloradas (veja a entrada sobre lógica multivaloradas). Novamente, a vagueza é vista como um fenômeno propriamente semântico, com as indeterminações concomitantes fornecendo casos de subdeterminação ou sobredeterminação semântica, mas com a verofuncionalidade preservada. As abordagens variam quanto ao número de valores de verdade não clássicos considerados apropriados para modelar a vagueza e desfazer o paradoxo de sorites.

Uma proposta inicial, desenvolvida pela primeira vez em Halldén (1949) e Körner (1960) e reformulada em Tye (1994), usa uma lógica de três valores. A motivação para tal lógica é semelhante à da supervaluacionista. Assim como um predicado vago divide objetos em extensão positiva, extensão negativa e penumbra, as sentenças vagas podem ser divididas em verdadeiras, falsas, e indeterminadas. Ao contrário da semântica de supervaloração, no entanto, os conectivos são todos definidos verofuncionalmente. Embora Halldén tenha proposto as tabelas trivaloradas fracas de Kleene, as tabelas trivaloradas fortes de Kleene prevaleceram como a escolha preferida. Para as tabelas relevantes, veja o apêndice de Haack (1974). Uma variação recente, nesse tema, é Field (2003), que complementa as tabelas fortes de Kleene com um condicional não verofuncional aprimorado e distingue a semântica trivalorada da abordagem comum de lacuna de valor de verdade.

A resposta específica ao paradoxo de sorites depende ainda, então, da definição de validade adotada. Uma generalização comum do conceito de validade à lógica multivalorada envolve a designação de certos valores. Uma sentença vale (ou é asserível) em uma interpretação multivalorada somente se ela recebe um valor designado. A validade pode então ser definida como a necessária preservação de valor designado. Na lógica clássica, é claro, somente a verdade é designada e, portanto, o conceito generalizado se reduz ao conceito clássico de necessária preservação da verdade. Há então duas escolhas não triviais: seja o conjunto de valores designados {verdadeiro} ou {verdadeiro, indeterminado}. A primeira proposta, defendida por Körner e por Tye, resulta em uma resposta de tipo (2) ao paradoxo. A última proposta resulta em uma lógica paraconsistente e produz uma resposta de tipo (3) (veja a seção sobre sistemas multivalorados na entrada sobre lógica paraconsistente). Quando associada às tabelas fortes de Kleene, resulta no sistema paraconsistente LP, proposto em outro lugar para lidar com o paradoxo do mentiroso e oferecido como uma lógica da vagueza em Weber (2010).

Enquanto algumas são motivados a adotar as abordagens trivaloradas anteriores por sua verofuncionalidade, outras consideram as consequências inaceitáveis. Aquelas que, por exemplo, acham argumentos supervaluacionistas para leis clássicas plausíveis rejeitarão asserções de terceiro excluído às vezes não sendo totalmente verdadeiras, ou contradições às vezes não sendo totalmente falsas, como pode ser o caso em tais sistemas. Uma preocupação adicional com tais abordagens, aplicável também ao supervaloracionismo, é que a invocada divisão tripartite de sentenças parece enfrentar objeções semelhantes àquelas que levaram ao abandono da divisão bipartida efetuada pela lógica clássica bivalorada. Devido ao fenômeno da vagueza de ordem superior (em particular vagueza de segunda ordem), não parece haver mais motivos para supor que exista uma fronteira nítida entre as sentenças verdadeiras e as indeterminadas ou as sentenças indeterminadas e as sentenças falsas do que havia para supor que exista uma fronteira nítida entre as sentenças verdadeiras e as falsas. O fenômeno da vagueza que orienta o paradoxo de sorites não sugere duas fronteiras nítidas mais do que sugere uma fronteira. Conceitos vagos parecem ser conceitos sem quaisquer limites. Nenhum número finito de divisões parece adequado. Tye (1994) busca evitar essas dificuldades empregando uma metalinguagem vaga; Sainsbury (1990) propõe que termos vagos são "sem limites", e que pertencer à extensão de um predicado vago é mais como ser atraído por um polo magnético do que como se encaixar em um nicho (como a sabedoria convencional poderia sugerir).

Goguen (1969) e Zadeh (1975), por outro lado, sugerem substituir a lógica clássica de dois valores por uma de infinitos valores de verdade. Hyde (2008) também adota essa abordagem, embora a semântica de valores infinitos seja considerada um dispositivo puramente formal e não um compromisso com graus de verdade (veja Cook (2002) a esse respeito). Lógicas de valores infinito ou difusas (veja a entrada sobre lógica difusa) também foram, no entanto, promovidas precisamente por seu reconhecimento de graus de verdade. Assim como a calvície apresenta-se em graus, argumenta-se que também é graduada a verdade de sentenças que predicam calvície das coisas. O fato de João ser mais calvo do que José reflete-se na sentença "John é calvo" ter um grau de verdade maior do que "José é calvo". Smith (2008) defende uma lógica difusa exatamente por esse motivo.

Lógicas de valores infinitos são então desenvolvidas para resolver o paradoxo de sorites de várias maneiras. Como em todas as lógicas multivaloradas, os conectivos e a validade podem ser definidos de várias maneiras, dando origem a uma série de lógicas distintas. Uma proposta padrão prossegue por meio da semântica verofuncional, de valores contínuos, de Łukasiewicz (ver o apêndice de Haack (1974)). Assim como no caso trivalorado, o tipo de resposta oferecida ao paradoxo também depende crucialmente da definição de validade. Onde a validade é definida como preservação de valor designado e apenas o valor máximo é designado, o sorites condicional admite uma resposta de tipo (2), como em Hyde (2008). No entanto, restabelecer a validade das leis clássicas nessa abordagem geral exigiria tornar designado mais do que o valor máximo, e uma resposta de tipo (3) se segue. Em contraste, Machina (1976) sugere definir validade como preservação do menor grau de verdade possuído por qualquer uma das premissas do argumento. Nessa abordagem, os sorites condicionais são inválidos e, portanto, novamente ocorre uma resposta de tipo (3). Edgington (1996) expõe uma teoria de graus não verofuncional, particularmente distinta, que preserva o princípio da bivalência e a lógica clássica. Nessa abordagem, a forma condicional do sorites é válida e uma resposta de tipo (2) é defendida. Smith (2008) combina uma teoria de grau não bivalente e não verofuncional com a lógica clássica por meio de uma definição distinta de validade. A abordagem singular de Smith fornece outra resposta de tipo (2) ao paradoxo.

Assim como nas abordagens trivaloradas, uma série de problemas acomete as abordagens de valores infinitos para a vagueza. Primeiro, onde a infinidade de valores semânticos é encarada como modelando graus de verdade, a própria ideia de um grau de verdade precisa de explicação. Segundo, se valores de verdade numéricos são usados, alguma justificativa parece necessária para as atribuições particulares de valor de verdade. Terceiro, as implicações completas de abandonar a bem compreendida teoria clássica em favor de uma teoria de grau precisam ser explicadas antes que uma avaliação adequada de seu mérito possa ser feita. Sobre esses pontos, veja o capítulo 2 de Sainsbury (1995), e o capítulo 4 de Keefe (2000). Para uma defesa estendida, veja o capítulo 5 de Smith (2008). Além disso, está longe de ser claro se tal abordagem evita com sucesso problemas de vagueza de ordem superior. E a suposição de um conjunto de verdade totalmente ordenado é excessivamente simples. Nem todas as sentenças de linguagem natural são comparáveis quanto a sua verdade. Devido à natureza multidimensional de um conceito como vermelhidão, podemos ser incapazes de dizer de duas manchas avermelhadas que diferem em matiz, brilho, ou saturação de cor, se uma é mais vermelha que a outra. Sobre os últimos pontos, veja o capítulo 4 de Williamson (1994). Smith (2008, capítulo 6) argumenta que o chamado problema da vagueza de ordem superior é, na verdade, um fenômeno distinto, e propõe uma resposta distinta.

Smith defende a visão que ele chama de plurivaluacionismo difuso, misturando elementos de teorias de grau e do supervaluacionismo. A semântica plurivaluacionista diverge da supervaluacionista ao atribuir a cada predicado vago múltiplas extensões clássicas precisas ("interpretações aceitáveis") e ao abandonar a noção semântica de superverdade. Ele substitui superverdade por "apenas um nível de conversa" governado pela instrução "diga que uma frase é simplesmente verdadeira se for verdadeira em todas as interpretações aceitáveis" (Smith, 2008, p. 109–110). Smith (2008, p. 110) escreve:

> A plurivaluacionista nos dirá que "esta folha é vermelha" e "esta folha não é vermelha" não podem ser ditas nem simplesmente verdadeiras nem simplesmente falsas, ao passo que "esta folha é vermelha ou não vermelha" pode ser dita simplesmente verdadeira… Não temos violação da verofuncionalidade porque não há nível de fato semântico no qual […] uma disjunção recebe o valor Verdadeiro, enquanto nenhum de seus componentes o é. Pois os únicos fatos semânticos são os fatos sobre o que está acontecendo em cada interpre-

> tação aceitável—e estes são inteiramente clássicos (portanto, verofuncionais). O que temos é apenas um nível de conversa sobreposto a esses fatos semânticos. A conversa soa não verofuncional, mas é na verdade epifenomenal… Ela não descreve literalmente uma realidade semântica não funcional da verdade.

Embora "verdade simples" possa diferir significativamente da superverdade, a plurivaluacionista endossa a noção de que as propriedades definidas em múltiplas avaliações (interpretações) desempenham um papel significativo no comportamento verbal de falantes competentes.

3.3.4 Contextualismo e seus parentes

Insatisfeito com abordagens multivaloradas e supervaluacionistas, Kamp (1981) introduziu uma solução contextualista para o paradoxo. Focando-se na forma indutiva do sorites, Kamp sustentou que toda instanciação da premissa maior é verdadeira em seu contexto individual, onde um contexto consiste das sentenças (contendo o predicado dado) previamente aceitas como verdadeiras. Para Kamp, um contexto é apenas um conjunto de sentenças. Cada instância é verdadeira porque seu antecedente deve ser adicionado ao contexto operativo antes que seu consequente seja avaliado, e os valores adjacentes referidos no antecedente e consequente são apenas incrementalmente diferentes. Em uma semântica clássica, a premissa maior universal seria então verdadeira também; mas Kamp adota uma definição não clássica ditando que a premissa universal é verdadeira em contextos (i) onde suas instâncias são verdadeiras e (ii) que permanecem coerentes quando a premissa universal é ela própria adicionada. O problema é que adicionar essa premissa produz um contexto incoerente que "atribui valores de verdade opostos à mesma frase" (Kamp, 1981, p. 252). Portanto, a premissa é falsa, apesar de todas as suas instâncias serem verdadeiras. A relatividade contextual dessa visão é intuitivamente atraente, e é isenta da necessidade de explicar por que cada instância da premissa universal parece verdadeira quando pelo menos uma precisa ser falsa. Ao mesmo tempo, a semântica não padrão para o quantificador universal pode parecer contraintuitiva.

Inspirada por Kamp, uma abordagem contextualista subsequente (Raffman, 1994) afirma que a premissa maior do paradoxo é falsa, mas parece verdadeira por ao menos duas razões. Primeiro, nós a confundimos com a afirmação verdadeira de que se α_i é Φ então α_{i+1} é Φ quando os dois valores são considerados juntos, dois a dois. A afirmação sobre pares, embora verdadeira, não autoriza a conclusão paradoxal, que faz referência apenas a um valor considerado individualmente. A segunda razão é uma hipótese, a saber, que a premissa principal pode ser falsa, embora parecendo verdadeira, porque a falante que executa uma marcha forçada sofre uma mudança característica em suas disposições verbais no momento de mudar de "Φ" para "não-Φ". Essa mudança disposicional constitui uma mudança de contexto (semelhante a uma mudança de Gestalt) que permite que as extensões coordenadas de "Φ" e de "não-Φ" mudem de modo que os valores α_i e α_{i+1} adjacentes ao ponto de troca sejam agora ambos classificados como não-Φ; em particular, α_i é classificada como Φ antes da troca, e como não-Φ depois. Assim, nenhum predicado é aplicado de forma a distinguir entre os dois valores relativos ao mesmo contexto, e assim a falante é capaz de mudar de "Φ" para "não-Φ" sem cruzar uma fronteira. A premissa maior parece verdadeira porque falhamos em perceber que a verdade pode ser assegurada para todas as suas instâncias conjuntamente apenas mediante equivocação de contexto.

Essa última visão foi criticada por (entre outras coisas) aplicar-se apenas ao paradoxo da marcha forçada, em oposição ao sorites propriamente dito; o sorites diz respeito a uma série de valores (propriedades como cores, alturas, idades, etc.) em abstrato, independentemente de qualquer coisa relativa às disposições verbais ou comportamentais de falantes. Para colocar a crítica de outra forma, a descrição de Raffman pode explicar por que a premissa maior do sorites de marcha forçada parece verdadeira, mas não toca o paradoxo propriamente dito. Na medida em que suas soluções frequentemente envolvem um elemento dinâmico, outros tratamentos contextualistas do paradoxo também podem ser vulneráveis a essa objeção.

Soames (1999, 2002) sustenta que termos vagos são sensíveis ao contexto da mesma forma que expressões indexicais. Stanley (2003) objeta que, se Soames estiver certo, então um diagnóstico do paradoxo como equivocando um parâmetro contextual implícito é interditado porque indexicais não admitem interpretação variável em elipse de sintagma verbal. Considere a afirmação "José está cansado agora e Joana também". Tanto a primeira quanto a segunda ocorrência (implícitas) do indexical "agora" devem receber a mesma interpretação: José e Joana estão cansados ao mesmo tempo. Como resultado dessa fixidez de interpretação, versões do paradoxo de sorites que empregam essas elipses não estão abertas à solução contextualista, mesmo na presença do tipo relevante de variação contextual. Stanley (2003, p. 272) dá o seguinte exemplo:

> Se aquilo$_1$ é um amontoado então aquilo$_2$ também é, e se aquilo$_2$ é, então aquilo$_3$ é, e se aquilo$_3$ é, então aquilo$_4$ é, …, e então aquilo$_n$ é…

onde “aquilo$_n$” refere-se ao enésimo elemento de uma série de sorites para “amontoado”. Se “amontoado” é indexical, como propõe Soames, não há como supor que sua extensão muda de conjunto para conjunto na formulação de Stanley. Defendendo a contextualista, Raffman (2005) responde negando que termos vagos sejam indexicais. Ela argumenta que, na elipse do sintagma verbal, termos vagos devem ser entendidos no modelo de “aquele elefante é grande, e aquela pulga também”. Aqui a extensão de “grande” varia entre os dois conjuntos, apesar da elipse (Ludlow, 1989).

Embora Graff Fara defenda uma solução epistêmica para o paradoxo, ela propõe uma explicação contextualista dinâmica para o apelo intuitivo da(s) premissa(s) condicional(ais). Em sua visão, predicados vagos expressam propriedades relativas a interesses, no sentido de que suas extensões são determinadas pelo que conta como significativo para uma falante em um momento. As premissas do paradoxo parecem verdadeiras porque um falante realizando uma marcha forçada tem um “interesse permanente em eficiência que [o] faz evitar fazer discriminações que são muito custosas” (Fara, 2008, p. 327–328). Por exemplo, para qualquer par de alturas adjacentes, incrementalmente diferentes em uma série de sorites para “alto”: quando o falante está considerando ativamente o par, de modo que a similaridade entre as duas alturas é saliente, o custo de discriminar entre elas supera os benefícios:

> Suponha que meu propósito principal seja escolher uma cerejeira para o quintal. Uma discriminação entre duas cerejeiras que são muito semelhantes em altura será muito custosa, dado meu interesse em eficiência. Mas a discriminação será ainda mais custosa quando eu estiver ativamente considerando as duas árvores como opções vigentes para meu propósito (Fara, 2008, p. 328).

Devido ao alto custo, as duas alturas em questão serão tratadas como “iguais para os propósitos presentes”, e, se uma das árvores for alta, a outra também será. Com efeito, a aplicação de “alta” é governada por uma forma de tolerância relativa a interesses. Presumivelmente, o interesse do falante deve invariavelmente ser em custo sempre que ele considera valores salientemente semelhantes; em particular, seu interesse não pode ser substituído por um diferente, como, digamos, um interesse na localização da fronteira nítida.

Alguns contextualizadores, como Burns (1991), fazem uso da ideia de que os limites nítidos de um predicado nunca estão onde se está olhando, para defender uma análise puramente pragmática do paradoxo de sorites que deixa a semântica e a lógica clássicas intactas; outros veem consequências para a lógica e a semântica e defendem uma abordagem não clássica. O livro de Shapiro (2006) desenvolve uma teoria contextualista dinâmica empregando uma variante distinta da lógica e da semântica supervaluacionistas para fornecer uma solução de tipo (2) para o paradoxo. Soames (1999) apela à sensitividade a contexto para defender uma lógica trivalorada de predicados vagos, postulando limites entre os exemplares determinados, os não exemplares determinados, e os casos limítrofes. Juntamente com a semântica forte trivalorada de Kleene, esse contextualismo não clássico nega a verdade da premissa maior universalmente quantificada do paradoxo, ao passo em que também nega sua falsidade. Tappenden (1993) sugere uma abordagem trivalorada similar que apela ao contexto para explicar a aparente verdade da premissa universalmente quantificada, mas seu uso da noção de contexto aqui difere sutilmente daquele de Kamp e Soames. O sorites condicional também admite solução. Aceitando as condições de verdade trivaloradas padrão para o quantificador universal, Soames (1999) considera que o sorites condicional tem alguma premissa condicional que não é verdadeira.

Para críticas e uma taxonomia útil de diferentes variedades de contextualismo (com atenção especial à distinção entre perspectivas de "mudança de extensão" e "mudança de fronteira"), veja Åkerman (2009) e Åkerman and Greenough (2010).

3.3.5 A teoria de intervalos múltiplos

A teoria de intervalos múltiplos ("multi-intervalos") é uma teoria semântica da vagueza que pretende manter a lógica clássica e a bivalência.[9] Aqui, a vagueza de uma expressão consiste em ela ter múltiplas formas de ser aplicada, igualmente permissíveis e arbitrariamente diferentes, em relação a um dado contexto (Raffman, 2014, capítulo 4). Em uma série de sorites, a vagueza de um termo reflete-se em ele ter múltiplos lugares, igualmente permissíveis e arbitrariamente diferentes, para parar de ser aplicado. Qualquer teoria adequada da vagueza deve reconhecer a existência de lugares de parada permissíveis em uma série sorites, uma vez que usuários competentes de um termo vago são obrigados a parar de aplicá-lo antes do fim. Por exemplo, em uma série de sorites de idades procedendo de uma idade claramente avançada de 90 para uma idade claramente de meia-idade (não idosa, portanto) de 50, digamos para os americanos em 2018—torne o contexto tão detalhado quanto quiser—os falantes podem parar de aplicar "velho" aos 70, ou aos 67, ou 65,

[9]A bivalência é preservada em grande medida por meio de uma análise incompatibilista de casos limítrofes que os define em termos de uma oposição entre predicados incompatíveis como "velho" e "de meia-idade" em vez de contraditórios como "velho" e "não velho" (capítulo 2 de Raffman (2014)).

ou 63,5, etc.[10] Diferentes falantes pararão em idades diferentes, e o mesmo falante parará em idades diferentes em ocasiões diferentes. Qualquer ponto de parada específico na série é arbitrário, portanto sem força legislativa; falantes não podem justificadamente acusar umas às outras de erro quando param em lugares diferentes. Em contraste, um limite seria legislativo; as falantes que falhassem em parar de aplicar "velho" nos limites deste estariam cometendo erros. A distinção entre limites e pontos de parada permissíveis é uma pedra angular da abordagem multi-intervalos.

Diz-se que essa multiplicidade de aplicações reflete-se na semântica do predicado na forma de múltiplos intervalos de aplicação. Um intervalo de aplicação é apenas uma representação abstrata, na semântica, de uma maneira permissível de aplicar o predicado. Mais formalmente: um intervalo é um conjunto de valores (por exemplo, idades) a cujas instanciações o predicado pode ser competentemente aplicado. Em uma série de 90 a 50, um intervalo de aplicação de "velho" pode conter as idades de 90 a 70, outro, de 90 a 65, outro, de 90 a 63,5, e assim por diante; e as idades nesses vários intervalos (por exemplo, 90 a 70) serão instanciadas por pessoas diferentes em mundos diferentes.

De acordo com a visão de múltiplos intervalos, uma sentença aplicando um termo vago a um determinado valor é verdadeira em relação a cada um dos seus intervalos que contêm esse valor, e falsa em relação a cada um dos outros. Alguns intervalos de "velho", "limítrofe", e "meia-idade" são exibidos na figura (II).1. Note que cada predicado tem alguns intervalos que se sobrepõem a alguns intervalos dos outros dois. A figura indica que, para uma pessoa de 63 anos, a sentença "x é velho" é verdadeira em relação ao terceiro, ao quarto, e ao quinto intervalo de "velho", e falsa em relação ao primeiro e ao segundo. A sentença "x é de meia-idade" é verdadeira em relação ao primeiro e ao segundo intervalo de "meia-idade", e falsa em relação a cada um dos outros. A sentença "x é meio velho" é verdadeira em relação a cada intervalo de "quase velho", exceto o quarto, e falsa em relação ao último.

Raffman alerta contra duas confusões em potencial. (1) Intervalos de aplicação não são precisificações (Raffman, 2014, p. 102–3). Para ver porque, observe que na visão multi-intervalos, o predicado "limítrofe" tem intervalos de aplicação como qualquer outro termo vago; intervalos de aplicação de "limítrofe" contêm valores limítrofes. Em contraste, por sua natureza, as precisificações não contêm valores limítrofes. Segundo, um intervalo de (por exemplo) "velho" contém apenas idades avançadas, enquanto uma precisificação de "velho"

[10] Aqui, um contexto é definido por coordenadas como dimensões decisivas, classes de comparação, e categorias contrastivas; também pode incluir propósitos, interesses das falantes, e padrões de aplicação, entre outros.

```
90 85 80 78 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45
V                                                                                            MI
90 ------------------------------ 70
90 ------------------------------------------------ 65
90 ------------------------------------------------------------------------------------- 55         "velho"
90 ------------------------------------------------------------------ 60
90 ------------------------------------------------------- 63
                                        69 --------------------------------------------------------------- 50
                                                      64 ------------------------------------------ 50   "de meia-
                                                                   59 -------------------------- 50       idade"
                                                                                 54 --------- 50
                                                            62 ------------------------------------- 50
                                        69 ------------------- 62
                                                      64 --------- 60
                                                 66 ---------------------------------- 55               "limítrofe"
                                                                 60 ------ 57
                                            68 ------------------------------------- 56
```

Figura (II).1: Alguns intervalos de aplicação de “velho”, “meia-idade” e “velho [meia-idade] limítrofe”.

contém idades avançadas e idades não avançadas (de meia-idade, por exemplo). Portanto, um intervalo contém apenas um local de parada permissível, enquanto uma precisificação contém um limite nítido. Consequentemente, em terceiro lugar, enquanto um intervalo especifica uma maneira pela qual falantes competentes podem realmente aplicar o termo, um falante que aplicasse “antigo” conforme uma precisificação estaria aplicando-o (talvez erroneamente) como se ele tivesse limites nítidos. Em quarto lugar, a visão de múltiplos intervalos não contém nenhum análogo de superverdade; a verdade comum é a verdade relativa a um único intervalo. (2) Intervalos de aplicação não são (aspectos de) contextos. Entre outras coisas, enquanto os falantes normalmente estão (ou podem estar) cientes do contexto ao qual estão relativizando, e podem escolher um determinado contexto por uma certa razão, eles não escolhem (não podem escolher) os intervalos aos quais relativizarão suas aplicações de um termo vago. Em vez disso, os falantes simplesmente escolhem como classificarão um determinado valor, e essa classificação é relativizada—automaticamente, por assim dizer, em virtude da semântica do termo—em relação a cada um de seus intervalos que contém o valor em questão. A relativização a intervalos não é algo que falantes fazem. Nesse âmbito, vale notar que, enquanto os tratamentos contextualistas do sorites são tipicamente pareados com um tipo distinto de semântica para termos vagos, por exemplo, uma semântica epistemicista, supervaluacionista ou trivalorada, a solução multi-intervalo emprega uma semântica multi-intervalo própria.

Na visão multi-intervalos, diz-se que o sorites é resolvido porque, sob pena de equivocação de intervalos, cada linha no paradoxo deve ter seu valor de verdade relacionado aos mesmos intervalos de aplicação de “velho”. E como cada intervalo contém uma última idade—um ponto de parada permissível—a premissa maior do paradoxo é falsa em relação a cada intervalo do predicado para qualquer contexto.

A teórica de multi-intervalos levanta a hipótese de que a premissa maior do paradoxo parece verdadeira porque a confundimos com duas regras pragmáticas para o uso de palavras vagas (Raffman, 2014, p. 172–175):

1. Para qualquer termo vago “Φ”: se α_n e α_{n+1} são apenas incrementalmente diferentes em uma dimensão decisiva, então qualquer aplicação diferencial do predicado entre eles, ou seja, qualquer aplicação de “Φ” a um, mas não ao outro, deve ser *arbitrária* (isto é: arbitrária ao invés de *ilícita*).

2. Para qualquer termo vago “Φ”: se α_n e α_{n+1} são apenas incrementalmente diferentes em uma dimensão decisiva, então, se α_n, então α_{n+1}, na medida em que α_n e α_n são considerados *dois a dois*.[11]

Dos seus aspectos discutidos aqui, a abordagem multi-intervalo foi criticada mais proeminentemente por seu comprometimento com um relativismo extremo sobre a verdade. Oponentes objetam que uma coisa é relativizar a verdade a mundos possíveis, e a fatores contextuais como falantes, tempos, localizações espaciais, classes de comparação, interesses e propósitos de falantes, apostas, e padrões de avaliação; e outra bem diferente é relativizar a verdade a fatores que variam mesmo depois que todos esses parâmetros contextuais foram fixados. A relatividade extremamente refinada proposta pela teórica multi-intervalo parece esticar a noção de verdade até o ponto de ruptura. Além disso, surgem questões sobre a vagueza (de ordem superior) do próprio predicado “intervalo de aplicação”; e não está claro se falantes seguem uma regra como (I) acima. Veja Åkerman (2014); Égré (2015); Sainsbury (2015); Scharp (2015); Caie (2015) para essas e outras críticas; e Raffman (2015) para

[11] O suporte para (II) é extrapolado de experimentos de Raffman e colegas sobre o uso de predicados de cor (capítulo 5 de Raffman (2014)). Alguém poderia questionar, é claro, se o uso de um termo não observacional como “velho” ou “rico” produziria os mesmos resultados.

algumas respostas. Embora não haja espaço para revisá-los aqui, devemos observar que teóricas da vagueza fizeram uma variedade de estudos empíricos reveladores investigando o uso de termos vagos por falantes comuns. Só a título de exemplo, veja Égré (2009); Ripley (2011); Alxatib and Pelletier (2011); Serchuk et al. (2011); Huang (2012, 2013); Égré et al. (2013).

3.4 Aceitando o paradoxo

Várias filósofas endossaram uma resposta de tipo (4), tirando a conclusão radical de que o paradoxo é insolúvel; estamos simplesmente presos com ele. A questão então é o que o paradoxo mostra. Dummett (1975), por exemplo, sustenta que predicados observacionais vagos cuja aplicação é supostamente governada por uma relação de indiscriminabilidade não transitiva são incoerentes. Essa visão parece fatal para a noção familiar de um tom determinado de cor (veja, por exemplo, Jackson (1975); Wright (1975); Peacocke (1992); Fara (2001); Mills (2002); Hellie (2005); Chuard (2007) para discussão).

Uma resposta de tipo (4) diferente sustenta que, ao contrário das aparências, os paradoxos de sorites condicionais são corretos. Por exemplo, é verdade, afinal, que nenhum número de grãos de trigo forma um amontoado. No entanto, tal visão imediatamente depara-se com problemas porque os paradoxos vêm em pares. Como observado acima, há versões negativas e positivas do quebra-cabeça dependendo de se o predicado de sorites é negado. Aceitar todos os argumentos de sorites como corretos requer assentir à afirmação adicional de que, já que um grão de trigo forma um amontoado, qualquer número forma. Uma incoerência radical se segue, uma vez que há um compromisso com todo e qualquer número tanto formando quanto não formando um amontoado. Da mesma forma, todos são calvos e ninguém é; todos são ricos e ninguém é, e assim por diante.

O problema é que a corretude de qualquer sorites condicional positivo enfraquece a verdade da premissa incondicional da respectiva versão negativa, e vice-versa. A menos que se esteja preparado para aceitar uma pandemia de contradições na linguagem natural, nem todos os sorites podem ser corretos. Unger (1979) e Wheeler (1979) propõem uma aceitação mais restrita. Insatisfeitos com respostas de tipos (1) e (3), nós aceitamos a aplicabilidade e validade das normas clássicas de raciocínio. No entanto, a insatisfação com as respostas de tipo (2) consideradas até agora—rejeitar alguma premissa condicional—deixa em aberto a opção de rejeitar a premissa menor (incondicional) ou aceitá-la e, com ela, a corretude do paradoxo. O que é defendido é a corretude daqueles sorites que negam "monticidade",[12] calvície, hirsutismo, riqueza, pobreza, etc. de tudo—uma resposta de tipo (4)—e a respectiva falsidade da premissa incondicional de todas as respectivas variantes positivas do argumento—uma resposta de tipo (2). Termos como "amontoado", "calvo", "cabeludo", "rico", e "pobre" não se aplicam a nada. Para críticas, veja o capítulo 6 de Williamson (1994).

4. Unificação com o paradoxo do mentiroso

O paradoxo sorites tem sido tradicionalmente visto como não relacionado de nenhuma forma substancialmente interessante aos paradoxos semânticos, e relativos a teoria de conjuntos, de autorreferência. No entanto, McGee (1991) e Tappenden (1993) propuseram um tratamento unificado dos paradoxos do mentiroso e de sorites, com base em similaridades entre predicados vagos e o predicado da verdade. Mais recentemente, Field (2003, p. 262) fala sobre

> certa tentação de conectar [os paradoxos semânticos aos paradoxos da vagueza] ao ver os paradoxos semânticos como devidos a algo semelhante à vagueza ou indeterminação em conceitos semânticos como "verdadeiro".

Field (2008) desenvolve mais esse tema, embora seja dedicado principalmente a uma resolução dos paradoxos do mentiroso, de Curry, e outros. A abordagem de Field é por meio de uma lógica que abandona a lei do terceiro excluído.

Algumas veem a unificação como muito mais claramente indicada pelo suposto fato de que os paradoxos semânticos e de sorites são eles próprios "de um tipo". Assim, Colyvan (2009) aponta para uma série de maneiras pelas quais os paradoxos podem ser considerados de um tipo *per se*, concluindo que o mentiroso e o sorites são exemplares e, portanto, merecedores de uma solução similar. Priest (2010) endossa essa afirmação, argumentando que tanto os paradoxos de autorreferência quanto o de sorites têm uma estrutura subjacente comum, satisfação do que Priest chama de "esquema de fechamento". Partindo do pressuposto de que essa estrutura comum basta para justificar um tratamento semelhante, Priest defende uma resposta paraconsistente ao sorites, tendo defendido em outros lugares uma resposta paraconsistente aos paradoxos da autorreferência. De fato, assim como acontece com sentenças paradoxais, algumas frases vagas envolvendo casos limítrofes fornecerão exemplos de contradições verdadeiras, dialeteias.

---

[12] N.T.: a propriedade de ser um amontoado (de alguma coisa).

5. Lições filosóficas

Tendo considerado várias grandes famílias de respostas aos desafios lógicos e semânticos colocados pelo sorites, vale a pena refletir sobre algumas das questões filosóficas mais amplas que o problema levanta. Uma vez que o fenômeno profundamente intrigante da vagueza manifesta-se primeiramente e sobretudo como um fenômeno linguístico, não é surpreendente que as respostas se cruzem de várias maneiras com problemas relativos a significado, verdade e referência.

5.1 Significado como uso

Um desafio colocado contra a resposta da teórica epistemicista é que, nessa visão, a conexão comumente assumida entre significado e uso parece ser tensionada, se não completamente rompida (veja novamente §3.2). Embora o princípio de margem de erro discutido em Williamson (1994) possa servir para explicar como poderíamos ser ignorantes dos limites nítidos postulados, pode-se pensar que, uma vez que nosso uso de termos vagos não traça limites nítidos, ele não poderia contê-los, dada a conexão geralmente aceita entre significado e uso. Como Williamson (1994, p. 205) relata essa preocupação que outros podem ter, "a visão epistemicista de vagueza posiciona condições de verdade como flutuando inaceitavelmente livres de nossas disposições a concordar e discordar". Parece que tal visão deve abandonar a ideia de que nosso uso determina o significado.

Uma resposta óbvia é que a conexão entre significado e uso não é tão forte quanto se poderia supor. A natureza também pode às vezes desempenhar um papel na determinação do significado, por exemplo, no caso de termos de tipo natural; mas para um predicado como, digamos, "fino", é improvável que a natureza forneça o que nosso uso não fornece. Williamson responde ainda apontando que a tese de determinação em questão é realmente uma tese de superveniência—significado sobrevém ao uso—e essa tese pode ser aceita pelas epistemicistas. De fato, a epistemicista não pode dizer exatamente como o significado sobrevém ao uso, e não pode portanto calcular o significado ou as condições de verdade de uma aplicação de um termo vago a partir de fatos sobre o uso. No entanto, continua a resposta, essa incapacidade é algo que todas as teóricas devem aceitar. Supor que a teoria epistêmica deve atender essa demanda é depositar expectativas irracionais sobre a teoria (veja Williamson (1996) e Burgess (2001) para discussão adicional).

A tese da superveniência também é desafiada por considerações de simetria. Quando confrontada com um caso limítrofe de "fino", diz o argumento, uma usuária da linguagem não concordará nem com a aplicação do termo, nem com a aplicação de sua negação. Padrões de dissenso são simétricos de forma similar com relação às duas alegações.[13] E ainda assim, apesar dessa simetria no nível de uso, ela deve ser quebrada no nível de verdade e falsidade, onde um dos termos ou sua negação verdadeiramente aplica-se, conforme a teoria; uma ou outra asserção é verdadeira, e a outra é falsa. Se nossos padrões de uso deixam a questão incerta igualmente em qualquer direção, então como pode a verdade da questão ser resolvida sem arbitrariedade nem um rompimento da conexão entre significado e uso? A resposta, sugere Williamson, está no fato de que verdade e falsidade não são noções simétricas. A falsidade obtém-se na ausência de verdade, então onde há simetria no nível de uso, a falsidade reina. Se essa resposta é bem sucedida é debatido em Burgess (2001) e Weatherson (2003).

5.2 Verdade e o esquema T

Como já observado em conexão com o supervaluacionismo, teorias que abandonam a bivalência foram acusadas de ter que rejeitar a restrição tarskiana exigida sobre a verdade encapsulada em seu esquema T: "p" é verdadeiro se, e somente se, p. Diz-se que a rejeição da bivalência no contexto do esquema T leva ao absurdo (capítulo 7 de Williamson (1994); veja Wright (1995) para críticas). Essa acusação aplica-se de forma mais geral a qualquer teoria não bivalente de vagueza associada ao esquema T. Se validada, a pressão para abandonar a bivalência na presença de vagueza lançaria dúvidas sobre uma explicação deflacionária da verdade. Muitas acharão essa consequência desagradável. Field (2008), por exemplo, dedica-se a salvar essa explicação da verdade de uma série de paradoxos, e rejeita uma abordagem de lacunas de valor de verdade.

As supervaluacionistas responderam observando que, embora o esquema T não seja verdadeiro, uma tese de implicação mútua correspondente não é ameaçada: " 'p' é verdadeiro" implica e é implicado por "p". No entanto, a última afirmação é estritamente mais fraca do que a respectiva afirmação envolvendo o condicional de acordo com o supervaluacionismo, e podemos nos perguntar se o comprometimento mais fraco é suficiente para capturar o que importa sobre a verdade (ver o capítulo 8 de Keefe (2000)). Outras discordaram do argumento de Williamson apontando que, no contexto de abordagens não bivalentes

[13] Essas são, naturalmente, afirmações empíricas que precisariam ser testadas.

à vagueza, a negação pode ser definida de várias maneiras e que o argumento supõe uma rejeição da bivalência invocando uma leitura particularmente forte da negação. Williamson argumenta em resposta que, embora possa ser oferecida uma descrição apropriadamente fraca da negação, o suficiente para minar o argumento por uma aceitação geral da bivalência, no caso especial da vagueza, o fenômeno da vagueza de ordem superior fornece os materiais para reduzir ao absurdo, de forma semelhante, essa rejeição mais fraca. Ver Williamson (1994, 193f) e Pelletier and Stainton (2003) para discussão adicional.

A tese então é que, mesmo se houvesse um sentido em que a verdade fosse não bivalente e, ainda assim, satisfizesse o esquema T, disponibilizando assim uma explicação deflacionária, a natureza particular do problema colocado pela vagueza impede tal síntese. A profundidade do problema, como evidenciado pelo fenômeno da vagueza de ordem superior, mostra que ele não pode ser explicado por uma rejeição da bivalência apenas.

5.3 A imperscrutabilidade da referência

Tentativas de resolver o paradoxo de sorites também evidenciam questões de referência. Diferentemente das respostas epistemicistas ao sorites que postulam limites imperscrutáveis, o supervaluacionismo é frequentemente associado a uma abordagem semântica da vagueza aparentemente comprometida com a imperscrutabilidade da referência.

Considere um paradoxo de sorites usando o predicado "está no Everest" usando a série de discriminações milimétricas ao longo de uma linha do seu pico até o fundo do vale abaixo. O primeiro ponto (o cume) está claramente no Everest. O último (no vale) claramente não está. E não há um ponto claro no meio onde traçaríamos a fronteira nítida separando a montanha de seus arredores. A vagueza ou indeterminação que subjaz a esse paradoxo de sorites não é, nessa abordagem, resultado de limitações epistêmicas, nem resultado de indeterminação no próprio Everest, mas, em vez disso, surge como resultado da indeterminação em torno do que contar como o referente do termo. De acordo com a supervaluacionista, vagueza é uma questão de indecisão semântica, como frequentemente se coloca. No caso em questão, simplesmente não há fatos quanto a exatamente que porção da Terra está sendo referida. Há uma gama de candidatas admissíveis, todas com igual pretensão de ser o Everest, entre as quais simplesmente não decidimos, nem (parafraseando Lewis) alguém é estúpido o suficiente para tentar. Nesse caso, sobrepondo-se ao problema da multiplicidade (veja a entrada sobre o problema da multiplicidade), a teoria compromete-se com um termo singular "Everest", embora aparentemente como um sintagma denotativo, sem nenhum referente determinado único. Isso coaduna-se com a análise muito anterior de Russell de vagueza como “um-multiplicidade na denotação”.

Como aponta Keefe (2000, capítulo 7.1), o supervaluacionismo assim entendido não obstante torna verdadeira a alegação de que há apenas um (nitidamente delimitado) Monte Everest (reivindicando assim uma solução para o problema da multiplicidade, e para o paradoxo de sorites supramencionado, já que é verdade que há um ponto de corte nítido para se estar no Everest), mesmo que não haja nenhuma montanha (nitidamente delimitada) da qual seja verdade que é a coisa referida por “Everest” (e, portanto, nenhum ponto na montanha do qual possamos dizer que é verdadeiramente o ponto de corte). Há apenas um Everest, mas não há fatos sobre o que ele é.

Assim como nos problemas anteriores sobre o papel da quantificação existencial no supervaluacionismo, pode-se debater se essa é uma consequência a ser aceita ou uma consequência indesejada que enfraquece a teoria sendo avançada. É certamente surpreendente que a referência seja inescrutável dessa forma. Além disso, tais casos não são a exceção; dada a ubiquidade de termos singulares vagos, tais casos parecem ser a regra (ver Lewis (1993); McGee and McLaughlin (2000); Morreau (2002)).

Referências Bibliográficas

Jonas Åkerman. Contextualist theories of vagueness. *Philosophy Compass*, 7(7):470–480, 2009. doi: 10.1111/j.1747-9991.2012.00495.x.

Jonas Åkerman. Review of Diana Raffman’s Unruly Words: A study of vague language. Notre Dame Philosophical Reviews, March 2014. URL http://ndpr.nd.edu/news/unruly-words-a-study-of-vague-language-2.

Jonas Åkerman and Patrick Greenough. Hold the context fixed—vagueness still remains. In Richard Dietz and Sebastiano Moruzzi, editors, *Cuts and Clouds: Vagueness, Its Nature and Its Logic*, pages 275–288. Oxford University Press, 2010.

Sam Alxatib and Francis Jeffry Pelletier. The psychology of vagueness: Borderline cases and contradictions. *Mind and Language*, 26(3):287–326, 2011. doi: 10.1111/j.1468-0017.2011.01419.x.

J.C. Beall and Mark Colyvan. Heaps of gluts and Hyde-ing the sorites. *Mind*, 110(438): 401–408, 2001. doi: 10.1093/mind/110.438.401.

Susanne Bobzien. Higher-order vagueness, radical unclarity, and absolute agnosticism. *Phi-*

losophers Imprint, 10(10):1–30, 2010.

Inga Bones and Diana Raffman. The contextualist solution to the sorites paradox. In S. Oms and E. Zardini, editors, *The Sorites Paradox*. Cambridge University Press, no prelo.

Otávio Bueno and Mark Colyvan. Just what is vagueness? *Ratio*, 25(1):19–33, 2012. doi: 10.1111/j.1467-9329.2011.00513.x.

J. Burgess. Vagueness, epistemicism, and response-dependence. *Australasian Journal of Philosophy*, 79(4):507–524, 2001. doi: 10.1080/713659306.

J.A. Burgess and I.L. Humberstone. Natural deduction rules for a logic of vagueness. *Erkenntnis*, 27(2):197–229, 1987. doi: 10.1007/BF00175369.

Linda Claire Burns. *Vagueness: An Investigation into Natural Languages and the Sorites Paradox*. Kluwer, Dordrecht, 1991.

Michael Caie. Review of Diana Raffman, unruly words. *Philosophical Review*, 124(3):415–419, 2015. doi: 10.1215/00318108-2895389.

Philippe Chuard. Indiscriminable shades and demonstrative concepts. *Australasian Journal of Philosophy*, 85(2):277–306, 2007. doi: 10.1080/00048400701343143.

Mark Colyvan. Vagueness and truth. In Heather Dyke, editor, *From Truth to Reality: New Essays in Logic and Metaphysics*, pages 29–40. Routledge, Londres, 2009.

Roy T. Cook. Vagueness and mathematical precision. *Mind*, 111(442):225–247, 2002. doi: 10.1007/BF00175369.

Richard Dietz and Sebastiano Moruzzi, editors. *Cuts and Clouds: Vagueness, Its Nature and Its Logic*. Oxford University Press, Oxford, 2010. doi: 10.1093/acprof:oso/9780199570386.001.0001.

Michael Dummett. Wang's paradox. *Synthese*, 30:301–324, 1975. doi: 10.1007/BF00485048. Reimpresso em Keefe and Smith (1996, p. 99–118).

Dorothy Edgington. Vagueness by degrees. In Rosanna Keefe and Peter Smith, editors, *Vagueness: A reader*, pages 294–316. MIT Press, Cambridge, 1996. doi: 10.7551/mitpress/7064.003.0018. URL https://doi.org/10.7551/mitpress/7064.003.0018.

Paul Égré. Soritical series and fisher series. In Hannes Leitgeb and Alexander Hieke, editors, *Reduction. Between the Mind and the Brain*, Publications of the Austrian Ludwig Wittgenstein Society, New Series 12, pages 91–115. Ontos-Verlag, 2009.

Paul Égré. Perceptual ambiguity and the sorites. In Rick Nouwen, Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, editors, *Vagueness and Communication: New Essays on Context-Sensitivity and Vagueness*, pages 64–90. Springer-Verlag, 2011. doi:

10.1007/978-3-642-18446-8_5.

Paul Égré. Borderline cases, incompatibilism and plurivaluationism. *Philosophy and Phenomenological Research*, 90(2):457–466, 2015. doi: 10.1111/phpr.12178.

Paul Égré, Vincent de Gardelle, and David Ripley. Vagueness and order effects: Evidence for enhanced contrast in a task of colour categorisation. *Journal of Logic, Language and Information*, 22(4):391–420, 2013. doi: 10.1007/s10849-013-9183-7.

Jonathan Ellis. Context, indexicals, and the sorites. *Analysis*, 64(4):362–364, 2004. doi: 10.1093/analys/64.4.362.

Delia Graff Fara. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, 28(1):45–81, 2000. ISSN 02762080, 2154154X. URL https://www.jstor.org/stable/43154331.

Delia Graff Fara. Phenomenal continua and the sorites. *Mind*, 110(440):905–936, 2001.

Delia Graff Fara. Profiling interest-relativity. *Analysis*, 68(4):326–335, 2008. doi: 10.1093/mind/110.440.905.

Hartry Field. No fact of the matter. *Australasian Journal of Philosophy*, 81(4):457–480, 2003. doi: 10.1080/713659756.

Hartry Field. *Saving Truth From Paradox*. Oxford University Press, Oxford, 2008. doi: 10.1093/acprof:oso/9780199230747.001.0001.

Kit Fine. Vagueness, truth and logic. *Synthese*, 30(3–4):265–300, 1975. doi: 10.1007/BF00485047.

Gottlob Frege. *Grundgesetze der Arithmetik*, volume 2. Verlag Hermann Pohle, Jena, 1903. Excertos traduzidos de Peter Geach and Max Black (eds.), 1960, Translations from the Philosophical Writings of Gottlob Frege, second edition, Oxford: Blackwell.

Peter Geach and Max Black, editors. *Translations from the Philosophical Writings of Gottlob Frege*. Blackwell, Oxford, segunda edition, 1960. Contém excertos de Frege, 1903.

J.A. Goguen. The logic of inexact concepts. *Synthese*, 19(3–4):325–373, 1969. doi: 10.1007/BF00485654.

Susan Haack. *Deviant Logic*. Cambridge University Press, Cambridge, 1974.

Sören Halldén. *The Logic of Nonsense*. Uppsala Universitets Arsskrift, Uppsala, 1949.

C. L. Hardin. Phenomenal colors and sorites. *Noûs*, 22(2):213–234, 1988. doi: 10.2307/2215860.

Benj Hellie. Noise and perceptual indiscriminability. *Mind*, 114(455):481–508, 2005. doi: 10.1093/mind/fzi481.

Terence Horgan. Robust vagueness and the forced-march sorites paradox. *Philosophical Perspectives*, 8:159–188, 1994a. doi: 10.2307/2214169.

Terence Horgan, editor. *Spindel Conference 1994: Vagueness. Special Issue*, volume 33. 1994b.

Minyao Huang. Manifestation of vagueness in language use: Two experiments. *Cambridge Occasional Papers in Linguistics*, 6:63–94, 2012. [Available online].

Minyao Huang. Tolerance effects in categorization with vague predicates. *Pragmatics & Cognition*, 21(2):340–358, 2013. doi: 10.1075/pc.21.2.05hua.

Dominic Hyde. From heaps and gaps to heaps of gluts. *Mind*, 106(424):440–460, 1997. doi: 10.1093/mind/106.424.641.

Dominic Hyde. *Vagueness, Logic and Ontology*. Ashgate New Critical Thinking in Philosophy. Ashgate, Aldershot, 2008.

Dominic Hyde. The sorites paradox. In *Ronzitti 2011*, pages 1–18. 2011. doi: 10.1007/978-94-007-0375-9_1.

Frank Jackson. *Perception*. Cambridge University Press, Cambridge, 1975.

Stanislaw Jaśkowski. Propositional calculus for contradictory deductive systems. *Studia Logica*, 24:143–157, 1948. doi: 10.1007/BF02134311. Translated by O. Wojtasiewicz. Originally published in 1948 in Polish in Studia Scientarium Torunensis, Sec. A, Vol. I, No. 5, II: 55–77.

Hans Kamp. The paradox of the heap. In Uwe Mönnich, editor, *Aspects of Philosophical Logic*, pages 225–277. Reidel, Dordrecht, 1981. doi: 10.1007/978-94-009-8384-7_8.

Rosanna Keefe. *Theories of Vagueness*. Cambridge University Press, Cambridge, 2000.

Rosanna Keefe. Context, vagueness, and the sorites. In J.C. Beall, editor, *Liars and Heaps: New Essays on Paradox*, pages 73–83. Oxford University Press, New York, 2003.

Rosanna Keefe. Vagueness without context change. *Mind*, 116(462):275–292, 2007. doi: 10.1093/mind/fzm275.

Rosanna Keefe and Peter Smith, editors. *Vagueness: A Reader*. MIT Press, Cambridge, MA, 1996.

Stephan Körner. *The Philosophy of Mathematics*. Hutchinson, London, 1960.

David Lewis. *On the Plurality of Worlds*. Basil Blackwell, Oxford, 1986.

David Lewis. Many but almost one. In Keith Campbell, John Bacon, and Lloyd Reinhardt, editors, *Ontology, Causality, and Mind: Essays on the Philosophy of D. M. Armstrong*, pages 23–38. Cambridge University Press, Cambridge, 1993.

Peter Ludlow. Implicit comparison classes. *Linguistics and Philosophy*, 12(4):519–533, 1989. doi: 10.1007/BF00632474.

Kenton F. Machina. Truth, belief and vagueness. *Journal of Philosophical Logic*, 5(1):47–78, 1976. doi: 10.1007/BF00263657.

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

Vann McGee and Brian P. McLaughlin. The lessons of the many. *Philosophical Topics*, 28 (1):129–151, 2000. doi: 10.5840/philtopics200028120.

Henryk Mehlberg. *The Reach of Science*. University of Toronto Press, Toronto, 1958.

Eugene Mills. Fallibility and the phenomenal sorites. *Noûs*, 36(3):384–407, 2002. doi: 10.1111/1468-0068.00377.

Michael Morreau. What vague objects are like. *The Journal of Philosophy*, 99(7):333–361, 2002. doi: 10.2307/3655512.

Rick Nouwen, Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, editors. *Vagueness in Communication*. Lecture Notes in Artificial Intelligence 6517. Springer-Verlag, Berlin, 2011.

Christopher Peacocke. *A Study of Concepts*. MIT Press, Cambridge, MA, 1992.

F. J. Pelletier and R. J. Stainton. On "the denial of bivalence is absurd". *Australasian Journal of Philosophy*, 81(3):369–382, 2003. doi: 10.1080/713659705.

Georgiĭ Valentinovich Plekhanov. *Fundamental Problems of Marxism (Osnovye voprosy marksizma)*. Lawrence and Wishart, London, 1908. Translated 1929 (from the 1928 edition), New York: International Publishers.

Graham Priest. Sorites and identity. *Logique et Analyse*, 34(135–136):293–296, 1991.

Graham Priest. Inclosures, vagueness and self-reference. *Notre Dame Journal of Formal Logic*, 51(1):69–84, 2010. doi: 10.1215/00294527-2010-005.

W. V. Quine. What price bivalence? *The Journal of Philosophy*, 78(2):90–95, 1981. doi: 10.2307/2025901.

Diana Raffman. Vagueness without paradox. *Philosophical Review*, 103(1):41–74, 1994. doi: 10.2307/2185872.

Diana Raffman. How to understand contextualism about vagueness: Reply to stanley. *Analysis*, 65(3):244–248, 2005. doi: 10.1093/analys/65.3.244.

Diana Raffman. *Unruly Words: A Study of Vague Language*. Oxford University Press, Oxford, 2014. doi: 10.1093/acprof:oso/9780199915101.001.0001.

Diana Raffman. Responses to discussants. *Philosophy and Phenomenological Research*, 90(2):483–501, 2015.

Nicholas Rescher. *Unknowability: An Enquiry into the Limits of Knowledge*. Lexington Books, New York, 2009. Chapter 7.

David Ripley. Contradictions at the borders. In Rick Nouwen, Robert van Rooij, Uli Sauerland, and Hans-Christian Schmitz, editors, *Vagueness in Communication*, pages 169–188. Springer-Verlag, Berlin, 2011. doi: 10.1007/978-3-642-18446-8_10.

Teresa Robertson. On Soames's solution to the sorites paradox. *Analysis*, 60(4):328–334, 2000. doi: 10.1093/analys/60.4.328.

Giuseppina Ronzitti, editor. *Vagueness: A Guide*. Springer, Dordrecht, 2011. doi: 10.1007/978-94-007-0375-9.

S. Rosenkranz. Wright on vagueness and agnosticism. *Mind*, 112(447):449–464, 2003.

Bertrand Russell. Vagueness. *The Australasian Journal of Psychology and Philosophy*, 1 (2):84–92, 1923. doi: 10.1080/00048402308540623. Reimpresso em Keefe and Smith (1996, p. 61–68).

R. Mark Sainsbury. *Paradoxes*. Cambridge University Press, Cambridge, 2nd edition, 1995.

R. Mark Sainsbury. Lessons for vagueness in scrambled sorites. *Metaphysica*, 14(2):225–237, 2013. doi: 10.1007/s12133-013-0123-4.

R. Mark Sainsbury. Vagueness and semantic methodology. *Philosophy and Phenomenological Research*, 90(2):475–482, 2015. doi: 10.1111/phpr.12174.

R. Mark Sainsbury and Timothy Williamson. Sorites. In Bob Hale and Crispin Wright, editors, *A Companion to the Philosophy of Language*, chapter 18. Blackwell, Oxford, 1995.

Kevin Scharp. Tolerance and the multi-range view of vagueness. *Philosophy and Phenomenological Research*, 90(2):467–474, 2015. doi: 10.1111/phpr.12173.

Phil Serchuk, Ian Hargreaves, and Richard Zach. Vagueness, logic and use: Four experimental studies on vagueness. *Mind and Language*, 26(5):540–573, 2011. doi: 10.1111/j.1468-0017.2011.01430.x.

Stewart Shapiro. *Vagueness in Context*. Oxford University Press, Oxford, 2006. doi: 10.1093/acprof:oso/9780199280391.001.0001.

Nicholas J. J. Smith. *Vagueness and Degrees of Truth*. Oxford University Press, Oxford, 2008. doi: 10.1093/acprof:oso/9780199233007.001.0001.

Scott Soames. *Understanding Truth*. Oxford University Press, New York, 1999. doi: 10.1093/0195123352.001.0001.

Scott Soames. Replies. *Philosophy and Phenomenological Research*, 65(2):429–452, 2002. doi: 10.1111/j.1933-1592.2002.tb00215.x.

Roy A. Sorensen. *Blindspots*. Clarendon Press, Oxford, 1988.

Roy A. Sorensen. *Vagueness and Contradiction*. Oxford University Press, New York, 2001.

Jason Stanley. Context, interest-relativity, and the sorites. *Analysis*, 63(4):269–280, 2003. doi: 10.1093/analys/63.4.269.

Jamie Tappenden. The liar and sorites paradoxes: Towards a unified treatment. *The Journal of Philosophy*, 90(11):551–577, 1993. doi: 10.2307/2940846.

Michael Tye. Sorites paradoxes and the semantics of vagueness. *Philosophical Perspectives*, 8:189–206, 1994. doi: 10.2307/2214170. Reprinted in Keefe and Smith 1996: 281–293.

Peter Unger. There are no ordinary things. *Synthese*, 41(2):117–154, 1979. doi: 10.1007/BF00869568.

Kees Van Deemter. Generating referring expressions that involve gradable properties. *Computational Linguistics*, 32(2):195–222, 2005. doi: 10.1162/coli.2006.32.2.195.

Bas C. van Fraassen. Singular terms, truth-value gaps, and free logic. *The Journal of Philosophy*, 63(7):481–495, 1966. doi: 10.2307/2024549.

Achille C. Varzi. Vagueness, logic and ontology. *The Dialogue. Yearbooks for Philosophical Hermeneutics*, 1:135–154, 2001.

Brian Weatherson. Epistemicism, parasites, and vague names. *Australasian Journal of Philosophy*, 81(2):276–279, 2003. doi: 10.1093/ajp/jag209.

Zach Weber. A paraconsistent model of vagueness. *Mind*, 119(476):1025–1045, 2010. doi: 10.1093/mind/fzq071.

Zach Weber and Mark Colyvan. A topological sorites. *The Journal of Philosophy*, 107(6): 311–325, 2010. doi: 10.5840/jphil2010107624.

Samuel C. Wheeler. On that which is not. *Synthese*, 41(2):155–194, 1979. doi: 10.1007/BF00869569.

Timothy Williamson. Vagueness and ignorance. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 66:145–177, 1992. doi: 10.1093/aristoteliansupp/66.1.145.

Timothy Williamson. *Vagueness*. Routledge, London, 1994.

Timothy Williamson. Definiteness and knowability. *The Southern Journal of Philosophy*, 33 (5):171–191, 1995.

Timothy Williamson. What makes it a heap? *Erkenntnis*, 44(3):327–339, 1996. doi: 10.1007/BF00167662.

Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000. doi: 10.1093/019925656X.001.0001.

Timothy Williamson. Soames on vagueness. *Philosophy and Phenomenological Research*, 65(2):422–428, 2002. doi: 10.1111/j.1933-1592.2002.tb00214.x.

Crispin Wright. On the coherence of vague predicates. *Synthese*, 30(3–4):325–365, 1975. doi: 10.1007/BF00485049.

Crispin Wright. Language mastery and the sorites paradox. In Gareth Evans and John McDowell, editors, *Truth and Meaning*, pages 523–549. Oxford University Press, Oxford, 1976.

Crispin Wright. The epistemic conception of vagueness. *The Southern Journal of Philosophy*, 33(5):133–159, 1995.

Crispin Wright. The illusion of higher-order vagueness. In Richard Dietz and Sebastiano Moruzzi, editors, *Cuts and Clouds: Vagueness, its Nature, and its Logic*, pages 523–549. 2010.

L. A. Zadeh. Fuzzy logic and approximate reasoning. *Synthese*, 30(3–4):407–428, 1975. doi: 10.1007/BF00485052.

(III) Paradoxos Epistêmicos[1]

Título Original: Epistemic Paradoxes
Autor: Roy Sorensen
Tradução: Renato Semaniuc Valvassori
Revisão: Rafael dos Santos Ongaratto

Paradoxos epistêmicos são enigmas que se voltam para o conceito de conhecimento ("*episteme*" é a palavra grega para conhecimento). Tipicamente, há respostas conflitantes e igualmente bem qualificadas a essas questões (ou pseudo-questões), o que leva diretamente a inconsistências. No longo prazo, esses enigmas nos instigam e nos orientam a corrigir ao menos algum erro profundo – quando não diretamente acerca do conhecimento, então no que concerne conceitos a ele aparentados, tais como justificação, crença racional e evidência.

Tais correções são do interesse de epistemólogos. Historiadores datam a origem da epistemologia quando da aparição dos céticos. Como fica evidente a partir da leitura dos diálogos de Platão em que figura Sócrates, paradoxos epistêmicos têm sido discutidos há dois mil e quinhentos anos. Dada a dificuldade que apresentam, alguns desses enigmas ainda podem muito bem ser alvo de discussão pelos próximos dois mil e quinhentos anos.

---

[1]SORENSEN, Roy, "Epistemic Paradoxes", In: ZALTA, E. N.; NODELMAN, Uri (eds.). Stanford Encyclopedia of Philosophy. Winter Edition. Stanford, CA: The Metaphysics Research Lab, 2023. Disponível em: https://plato.stanford.edu/archives/win2023/entries/epistemic-paradoxes/. Acesso em: 21 de mar. 2022.
A seguir está a tradução da entrada sobre Paradoxos Epistêmicos de Roy Sorensen na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/win2023/entries/epistemic-paradoxes/. 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/epistemic-paradoxes/. Gostaríamos de agradecer aos editores da Stanford Encyclopedia of Philosophy, principalmente o Prof. Dr. Edward Zalta e Uri Nodelman, por concederem permissão para traduzir e publicar esta entrada.

1. O Paradoxo da Prova Surpresa

Uma professora anuncia que haverá uma prova surpresa semana que vem. Será uma prova surpresa na medida em que os estudantes serão incapazes de saber com antecedência em que dia a prova será aplicada. Um estudante, contudo, objeta apontando que essa é uma impossibilidade: "A aula ocorre às segundas, quartas e sextas-feiras. Se a prova for aplicada na sexta-feira, então quinta-feira serei capaz de prever que a prova será na sexta-feira. Não seria, portanto, surpresa. Poderia a prova ser aplicada na quarta-feira? Não, pois na terça-feira eu saberia que a prova não seria aplicada na sexta-feira (dado o raciocínio anterior) e saberia que a prova não ocorreu na segunda (graças à minha memória). Portanto, na terça-feira eu poderia prever que a prova seria aplicada na quarta-feira e, assim, essa prova já não mais seria surpresa. A prova surpresa poderia ocorrer na segunda-feira? No domingo, as conclusões eliminatórias alcançadas já estariam disponíveis para mim e, consequentemente, eu saberia que a prova deveria ser aplicada na segunda-feira. Por conseguinte, uma prova na segunda-feira não poderia ser caracterizada como surpresa. Portanto, é impossível ocorrer uma prova surpresa".

Seria a professora capaz de cumprir aquilo que anuncia? Encontramo-nos numa encruzilhada. Por um lado, há o argumento da eliminação oferecido pelo estudante. (Para uma formalização, ver Holliday 2017). Por outro lado, o senso comum nos diz que provas surpresa são possíveis mesmo quando estamos em posição de saber que em algum momento elas ocorrerão. Qualquer uma das alternativas poderia ser vista como decisiva, não fosse o peso da resposta rival. Temos, assim, um paradoxo. Mas um paradoxo de que natureza? "Prova Surpresa" é um paradoxo definido em termos do que pode ser conhecido. Mais especificamente, uma prova é considerada surpresa se, e somente se, os estudantes não forem capazes de saber com antecedência em que dia a prova ocorrerá. Portanto, o enigma da prova surpresa é considerado um paradoxo *epistêmico*.

Paradoxos são mais do que surpresas edificantes. Eis uma surpresa edificante que não apresenta um paradoxo. A professora de estatística anuncia que realizará questionários aleatoriamente: "Haverá aulas todos os dias. A cada dia eu começarei uma nova aula rolando um dado. Quando o resultado da rolagem for seis, eu imediatamente aplicarei um questionário. Hoje, uma segunda-feira, obtive seis na rolagem. Portanto, vocês receberão um questionário. A última questão de seu questionário é: "Em qual dos dias subsequentes há a maior chance de ocorrer o próximo questionário aleatório?". A maior parte das pessoas responde que há a mesma probabilidade de ser aplicado um questionário aleatório em cada um dos dias subsequentes. Mas a resposta correta é: *Amanhã* (Terça-feira).

Fatos incontroversos sobre probabilidade revelam o erro incorrido e estabelecem a resposta correta. Para que o próximo questionário seja aplicado quarta-feira, deve haver uma *conjunção* de dois eventos: não ter um questionário aplicado terça-feira (com ▯ de chances disso ocorrer) *e* a aplicação do questionário quarta-feira (com ▯ de chances). A probabilidade de ocorrer a aplicação do questionário em cada dia subsequente vai tornando-se cada vez menor (seria impressionante se o próximo questionário só fosse aplicado dali a cem dias!). A questão não é se haverá uma rolagem resultando em seis em qualquer dos dias especificamente, mas quando o *próximo* seis será obtido. A obtenção do próximo seis é um evento que depende parcialmente do que ocorre no meio-tempo e parcialmente da rolagem do dia específico.

Esse enigma probabilístico é instrutivo e será retomado no decorrer deste verbete. Contudo, a existência de uma rápida e decisiva solução a ele mostra que apenas uma revisão moderada de nossas crenças iniciais foi necessária. Em contrapartida, quando nossas crenças profundas conflitam, as emendas propostas reverberam de forma imprevisível. Tentativas de se descartar uma crença profunda voltam-se contra nós, pois os problemas gerados ao nos distanciarmos de tais crenças demonstram sua centralidade. Frequentemente, a crença em questão acaba se tornando ainda mais arraigada: “Os problemas que merecem ser atacados provam seu valor ao revidar” (Hein 1966).

Um sinal da profundidade de uma crença (ou ao menos de desespero) é o fato de que comentadores começam a rejeitar regras de inferência bastante plausíveis. Há quem tenha defendido que a Prova Surpresa invalida a lei de bivalência, o princípio KK (de que se alguém está em posição de saber que *p*, então também está em posição de saber que sabe que *p*), e mesmo o princípio de fechamento (se alguém sabe que *p* e é capaz de deduzir competentemente *q* de *p*, então essa pessoa sabe que *q*) (Immerman 2017).

O paradoxo da prova surpresa também está ligado a problemas que não são evidentemente paradoxos – ou a problemas cuja caracterização enquanto sendo paradoxos é contestável. Considere o problema de Monty Hall. Há um prêmio por detrás de somente uma entre três portas. Após sua escolha, Monty Hall revelará o que há atrás de uma das portas não-premiadas. Ele dará a você, então, a opção de trocar sua porta escolhida pela outra porta fechada. Você deveria efetuar a troca para aumentar suas chances de ganhar o prêmio? Quando Marylin vos Savant respondeu que sim numa edição da revista *Parade* em 1990, muitos leitores — incluindo alguns acadêmicos — a repreenderam. A solução adequada ao problema foi fornecida por seu idealizador décadas antes e nunca chegou a ser esquecida ou efetivamente criticada.

O que torna o Problema de Monty Hall interessante é o fato de que *não* se trata de um paradoxo. Sempre houve consenso entre especialistas quanto à sua solução. Ainda assim, esse problema apresenta uma série de características psicológicas e sociológicas presentes em paradoxos. O Problema de Monty Hall é tão somente uma *ilusão cognitiva*. O *status* de paradoxo também é negado àqueles que só encontram *ironia* em previsões auto-derrotantes, e somente embaraço no "paradoxo da cognoscibilidade" (discutido abaixo). Tratar um problema como sendo um paradoxo tende a separá-lo das demais investigações que realizamos. Portanto, aqueles que querem se basear no resultado surpreendente negam que haja um paradoxo. No máximo ele concederá que *houve* um paradoxo. Paradoxos mortos são bons fertilizantes para a árvore do conhecimento.

Podemos aguardar futuros filósofos realizarem conexões históricas edificantes. O argumento de eliminação regressiva subjacente ao paradoxo da prova surpresa pode ser encontrado em contos populares alemães que datam de 1756 (Sorensen 2003a, 267). Talvez os estudiosos medievais tenham explorado esses argumentos escorregadios. Voltemo-nos, no entanto, aos comentários que presentemente podemos acessar.

1.1 Profecias auto-derrotantes e paradoxos pragmáticos

No século XX, a primeira reação pública ao paradoxo da prova surpresa endossou o argumento de eliminação defendido pelo estudante. D. J. O'Connor (1948) considerou o anúncio do professor como auto-derrotante. Se a professora não tivesse anunciado que haveria uma prova surpresa, ela de fato poderia ter aplicado uma prova surpresa. A lição de moral do paradoxo seria, portanto, que se você quiser aplicar uma prova surpresa, não anuncie sua intenção a seus alunos!

Mais precisamente, O'Connor comparou o anúncio do professor a proferimentos como "eu não estou falando agora". Ainda que consistentes, esses proferimentos "não poderiam ser verdadeiros em qualquer circunstância concebível" (O'Connor 1948, p. 358). L. Jonathan Cohen (1950) concordou com essa posição e classificou o anúncio da professora como sendo um paradoxo pragmático. Ele definiu um paradoxo pragmático como sendo uma declaração falseada por seu próprio proferimento. Assim, a professora teria ignorado o fato de que a maneira com a qual uma declaração é disseminada pode acabar condenando-a à falsidade.

A classificação de Cohen é excessivamente monolítica. É verdade que o anúncio da professora quanto à surpresa é comprometedor em ao menos um aspecto: os estudantes agora sabem que haverá uma prova. Mas esse aspecto comprometedor não é, por si só, suficiente para tornar a declaração auto-falseante. A *existência* de uma prova surpresa foi revelada, mas isso ainda deixa incerto em *qual* dia a prova ocorrerá. O anúncio de uma surpresa vindoura visa tornar uma ignorância desinformada em uma consciência da ignorância capaz de orientar a ação. Assim, um estudante que perde o anúncio não fica sabendo que haverá uma prova. Se nenhum aluno o informar da prova, o estudante em completa ignorância estará menos preparado para a prova do que seus colegas cientes de que não sabem o dia da prova.

Anúncios são feitos para servir a diferentes objetivos simultaneamente. A disputa entre a acurácia e a prestatividade torna possível que um anúncio seja auto-realizante por ser auto-derrotante. Imagine um meteorologista que avise o seguinte: "o tsunami da meia-noite causará fatalidade ao longo da costa". Tendo tido acesso ao aviso, amantes de espetáculos fazem uma viagem especialmente para presenciar a onda, e alguns se afogam. Por conseguinte, o anúncio do meteorologista foi bem sucedido enquanto predição na medida em que falha como aviso.

1.2 Determinismo Preditivo

Ao invés de encarar predições auto-derrotantes como uma demonstração da maneira com a qual a professora foi refutada, alguns filósofos interpretam que predições auto-derrotantes demonstram como o *estudante* é refutado. O argumento da eliminação sustentado pelo estudante incorpora previsões hipotéticas sobre qual dia em que a professora aplicará a prova. Não estaria o estudante, contudo, ignorando a capacidade e o desejo da professora de frustrar essas expectativas? Alguns teóricos de jogos sugerem que a professora seria capaz de derrotar essa estratégia ao escolher aleatoriamente o dia da prova.

Estudantes podem certamente ser mantidos na incerteza se a professora tiver a pretensão genuína de agir de maneira aleatória. Ela precisará preparar um questionário a cada dia e também terá de se preparar para a possibilidade de acabar aplicando seja muitos questionários, poucos questionários ou mesmo uma distribuição não representativa dos mesmos.

Se os custos dessa estratégia forem considerados onerosos, a professora pode optar por uma alternativa: no começo da semana, ela seleciona aleatoriamente um único dia. A identidade daquele dia será mantida em segredo. Como o estudante só saberá que a prova será em um daqueles dias, os alunos não conseguirão prever o dia da prova.

Esse plano é arriscado. Se, por mero acaso, o último dia acabar sendo selecionado para a aplicação da prova, essa deixará de ser surpresa. Isso porque, como no cenário original, o aluno tem conhecimento do anúncio da professora e consciência de que não ocorreu a prova nos dias anteriores. Logo, a professora deve excluir o último dia da seleção aleatória. Não obstante, poderia a professora acatar a seleção aleatória do penúltimo dia? Já fica claro o raciocínio que se segue.

Outra crítica à réplica do aluno ao raciocínio da professora adapta o experimento mental de Michael Scriven (1964). Para refutar o determinismo preditivo (a tese de que todos os eventos são previsíveis), Scriven supõe um agente "Previsor", o qual detém todas as informações, conhecimento das leis e capacidades de cálculo necessárias para prever as escolhas dos outros. Scriven também imagina o "Escapista", cuja principal motivação é escapar às previsões. Assim sendo, o Previsor deve ocultar sua previsão. O problema é que o Escapista tem acesso aos mesmos dados, leis e capacidade de cálculo que o Previsor. Logo, o Escapista pode duplicar o raciocínio do Previsor e, consequentemente, mesmo o Previsor ideal não é capaz de prever o Escapista. Assumamos que a professora é o Escapista e o aluno o Previsor. O Escapista deve vencer. Portanto, é possível fazer uma prova surpresa.

O argumento original de Scriven pressupõe que o Previsor e o Escapista podem ter simultaneamente todos os dados, leis e capacidade de cálculo necessários. David Lewis e Jane Richardson objetam:

> … a quantidade de cálculo necessária para permitir que o previsor conclua sua previsão depende da quantidade de cálculo feita pelo escapista, e a quantidade de cálculo necessária para permitir que o escapista termine de duplicar o cálculo do previsor depende da quantidade de cálculo feita pelo previsor. Scriven pressupõe que esses requisitos são compatíveis: ou seja, que há quantidades de cálculo disponíveis para o previsor e o escapista, de modo que cada um tenha o suficiente para terminar seus cálculos, dada a quantidade que o outro tem. (Lewis & Richardson 1966, pp. 70–71)

De acordo com Lewis e Richardson, Scriven equivoca-se ao dizer que "tanto o Previsor quanto o Escapista têm tempo suficiente para terminar seus cálculos". Essa frase, lida de certa maneira, produz uma verdade: contra qualquer Escapista, o Previsor pode terminar seus cálculos, e contra qualquer Previsor, o Escapista pode terminar seus cálculos. Entretanto, a premissa de compatibilidade requer a leitura falsa em que Previsor e Escapista podem ambos terminar um contra o outro.

Idealizar o professor e o aluno como na hipótese do Escapista e do Previsor não seria suficiente para desarmar o argumento de eliminação do aluno. Teríamos apenas formulado um enigma que pressupõe falsamente que os dois tipos de agentes são co-possíveis. Seria como perguntar: "Se Bill é mais inteligente do que qualquer outra pessoa e Hillary é mais inteligente do que qualquer outra pessoa, qual dos dois é o mais inteligente?"

O determinismo preditivo afirma que tudo é previsível. O determinismo metafísico afirma que há apenas uma maneira pela qual o futuro pode ser dado, a partir da maneira com a qual o passado ocorreu. Simon Laplace usou o determinismo metafísico como premissa para o determinismo preditivo. Ele argumentou que, como todo evento tem uma causa, uma descrição completa de qualquer estágio da história, combinada com as leis da natureza, implica o que acontece em qualquer outro estágio do universo. Scriven estava apenas desafiando o determinismo preditivo em seu experimento mental. A próxima abordagem desafia o determinismo metafísico.

1.2.1 O Problema da Presciência

O conhecimento prévio de uma ação parece incompatível com ela ser uma ação livre. Se eu sei que você terminará de ler este artigo amanhã, então você terá de terminá-lo amanhã (porque conhecimento implica verdade). Mas isso significa que você terminará o artigo mesmo que decida não fazê-lo — nada pode impedi-lo. Assim, se eu sei que você terminará de ler este artigo amanhã, você não é livre para agir diferentemente.

Talvez toda leitura que você faz seja compulsória. Se Deus existe, então Ele sabe tudo. Assim, a ameaça à liberdade torna-se aterradora para o teísta. O problema da presciência divina levanta a possibilidade de que o teísmo (em vez do ateísmo) impeça a livre escolha e, portanto, impeça que tenhamos qualquer responsabilidade moral.

Em resposta ao aparente conflito entre liberdade e presciência, os filósofos medievais negaram que proposições contingentes futuras tenham um valor de verdade. Eles consideraram estar estendendo uma solução que Aristóteles discute em *De Interpretatione* para o problema do fatalismo lógico. De acordo com essa abordagem que aceita uma lacuna no valor de verdade, "Você terminará este artigo amanhã" não é verdade agora. A previsão *se tornará* verdadeira amanhã. Um teísta moralmente sério pode concordar com o *Rubaiyat* de

Omar Khayyam:

> Dispõe o Eterno Escriba. E havendo escrito,
> A folha vira: e não há ciência ou devoção
> Que cancele uma Linha; e não há pranto aflito
> Que risque uma palavra! Ah, todo choro é vão!

[2]

A onisciência de Deus requer apenas que Ele conheça cada proposição verdadeira. Deus saberá que "Você terminará este artigo amanhã" assim que isso se tornar verdade — mas não antes disso.

A professora tem livre-arbítrio. Portanto, previsões sobre o que ela fará não são verdadeiras (antes da prova). O metafísico Paul Weiss (1952) conclui que o argumento do aluno falsamente assume que ele sabe que o anúncio é verdadeiro. De fato, o aluno só pode saber que o anúncio é verdadeiro depois que ele se torna verdadeiro — mas não com antecedência.

O lógico W. V. O. Quine (1953) concorda com a conclusão de Weiss de que o anúncio de uma prova surpresa pela professora falha em dar ao aluno o conhecimento de que haverá uma prova surpresa. No entanto, Quine rejeita o raciocínio de Weiss. Weiss viola a lei da bivalência (a qual afirma que toda proposição tem um valor de verdade, verdadeiro ou falso). Quine acredita que o enigma da prova surpresa não deve ser respondido abrindo-se mão da lógica clássica.

1.3 Suicídio Intelectual

Quine insiste que o argumento da eliminação oferecido pelo aluno é apenas uma *reductio ad absurdum* da suposição de que o aluno *sabe* que o anúncio é verdadeiro (ao invés de uma *reductio* do anúncio realizado pela professora). Ele aceita essa *reductio* epistêmica, mas rejeita a *reductio* metafísica. Dada a ignorância do aluno sobre o anúncio, Quine conclui que uma prova em qualquer dia seria imprevista. Ou seja, Quine aceita que o aluno não tem conhecimento prévio acerca do dia da prova, mas rejeita que não haja antecipadamente a verdade acerca de quando a prova será aplicada.

[2]N.T. No verbete é citada uma tradução realizada por Fitzgerald (1859) do poema originalmente escrito por Omar Khayyam. Escolhi, então, reproduzir uma tradução da versão de Fitzgerald realizada por Haddad (1972).

O senso comum sugere que os alunos são informados pelo anúncio. A professora está assumindo que o anúncio esclarecerá algo aos alunos. Ela parece certa em assumir que o anúncio dessa intenção produz o mesmo tipo de conhecimento que suas outras declarações de intenções (sobre quais tópicos serão selecionados para a aula, a escala de notas e assim por diante).

Há premissas céticas que poderiam levar Quine à conclusão de que os alunos não sabem que o anúncio é verdadeiro. Se ninguém pode saber nada sobre o futuro, como alegado pelo problema de indução de David Hume, então o aluno não pode saber que o anúncio da professora é verdadeiro (veja a entrada sobre *o problema da indução*). Mas negar todo o conhecimento do futuro para negar o conhecimento do aluno sobre o anúncio da professora é uma medida desproporcional e exacerbada. Não mate uma mosca com um canhão — a menos que seja uma mosca assassina e somente um canhão seja páreo!

Em escritos posteriores, Quine demonstra reservas gerais acerca do conceito de conhecimento. Uma de suas objeções favoritas é que "saber" é uma noção vaga. Se o conhecimento implica certeza, então muito pouco contará como conhecido. Quine infere que devemos igualar o conhecimento a uma crença verdadeira firmemente suportada. Perguntar quão firme essa crença deve ser é semelhante a perguntar quão grande algo tem que ser para ser considerado grande. Não há resposta para essa pergunta porque "grande" não tem o tipo de limite desfrutado por palavras precisas.

> Não há lugar na ciência para termos como "grande", por causa dessa falta de limite; mas há um lugar para a relação de ser-maior-que. Aqui vemos a retificação familiar e amplamente aplicável da vagueza: renunciar ao termo vago e apegar-se ao comparativo preciso. Mas isso não se aplica ao verbo "saber", mesmo gramaticalmente. Os verbos não têm flexões comparativas e superlativas... Acho que para propósitos científicos ou filosóficos o melhor que podemos fazer é desistir da noção de conhecimento, impraticável, e nos contentarmos com seus ingredientes isolados. Ainda podemos falar de uma crença como sendo verdadeira, e de uma crença como mais firme ou mais certa, para aquele que detém a crença, do que outra (1987, p. 109).

Quine está aludindo à generalização de Rudolf Carnap (1950) de que os cientistas substituem termos qualitativos (*alto*) por comparativos (*mais alto que*) e então substituem os comparativos por termos quantitativos (*ter n milímetros de altura*).

É verdade que alguns casos limítrofes de um termo qualitativo não são casos limítrofes para o comparativo correspondente. Mas o inverso também se aplica. Um homem alto que fica curvado pode ficar menos alto do que outro homem alto que não é tão comprido, mas tem uma postura melhor. Ambos os homens são claramente altos. Não está claro que "O homem mais comprido é mais alto". Termos qualitativos podem ser aplicados quando uma cota vaga é satisfeita sem a necessidade de se atentar aos detalhes. Apenas termos comparativos são atormentados por questões de desempate.

A ciência é sobre o que é o caso, em vez do que deveria ser o caso. Isso parece implicar que a ciência não nos diz no que devemos acreditar. A maneira tradicional de preencher a lacuna normativa é delegar questões de justificação aos epistemólogos. No entanto, Quine se sente desconfortável em delegar tal autoridade aos filósofos. Ele prefere a tese de que a psicologia é suficiente para lidar com as questões tradicionalmente abordadas pelos epistemólogos (ou pelo menos com as questões que ainda vale a pena tratar em uma Era da Ciência). Essa "epistemologia naturalista" parece implicar que "saber" e "justificado" são termos antiquados — tão vazios quanto "flogisto" ou "alma".

Aqueles dispostos a abandonar o conceito de conhecimento podem dissolver o paradoxo da prova surpresa. Mas para epistemólogos que veem esperança em respostas menos drásticas, isso é como usar uma bomba suicida para matar uma mosca.

Nosso homem-bomba, contudo, pode protestar que as moscas foram subestimadas. O eliminativismo epistêmico dissolve *todos* os paradoxos epistêmicos pois, de acordo com essa perspectiva, esses paradoxos são sintomas de um problema com o próprio conceito de conhecimento.

Observe que o eliminativista é mais radical que o cético. O cético acha que o conceito de conhecimento é coerente e bem definido — nós apenas ficamos aquém de ser conhecedores. Ele trata "Nenhum homem é um conhecedor" como "Nenhum homem é imortal" — não há nada de errado com o conceito de imortalidade, mas a biologia nos garante que nenhum homem é imortal. Essa ausência universal de conhecimento, não obstante, seria chocante. Mas o potencial de nos chocar não deve nos levar a matar o mensageiro (o cético) ou declarar ininteligível o vocabulário que compreende a mensagem (especificamente, a palavra "saber").

Ao contrário do mensageiro que nos diz "nenhum homem é imortal", o cético tem dificuldade em nos *dizer* "não há conhecimento". De acordo com Sexto Empírico, uma afirmação expressa a crença de que se sabe o que é afirmado (*Outlines of Pyrrhonism*, I, 3, 226). Ele condena a *afirmação* “não há conhecimento” (embora não a proposição expressa pela afirmação) como ceticismo *dogmático*. Sexto prefere o agnosticismo sobre o conhecimento ao invés do ceticismo (entendido como “ateísmo” sobre o conhecimento). No entanto, é igualmente inconsistente afirmar “ninguém pode saber se algo é conhecido”, pois isso transmite a crença de que se sabe que ninguém pode saber se algo é conhecido.

O eliminativista tem dificuldades ainda mais severas em declarar sua posição do que o cético. Alguns eliminativistas descartam a ameaça de autoderrota traçando uma analogia: aqueles que negavam a existência de almas eram acusados de minar uma condição necessária para afirmar qualquer coisa. No entanto, a explicação do teórico da alma sobre tais condições necessárias não dá razões para que se negue que um cérebro saudável é suficiente para a realização de estados mentais.

Se o eliminativista pensa que asserções apenas impõem o objetivo de expressar uma verdade, então ele pode consistentemente afirmar que “saber” é um termo problemático. No entanto, um epistemólogo pode reviver a acusação de autoderrota ao mostrar que afirmações de fato requerem que o falante atribua conhecimento a si mesmo. Essa concepção de asserções baseada no conhecimento foi recentemente apoiada por pesquisas acerca do próximo paradoxo que abordaremos.

1.4 Loterias e o Paradoxo da Loteria

As loterias representam um problema para a teoria segundo a qual a alta probabilidade de que uma crença seja verdadeira é suficiente para o conhecimento. Dado que há um milhão de bilhetes e apenas um ganhador, a probabilidade de “Este bilhete é um bilhete perdedor” é muito alta. No entanto, relutamos em dizer que isso gera o conhecimento de que essa proposição é verdadeira.

Superamos essa relutância depois que o bilhete ganhador é anunciado. Agora o bilhete é sabidamente perdedor e jogado no lixo. Mas espere! O testemunho não fornece *certeza*. Nem a percepção ou a lembrança. Quando pressionados, admitimos que há uma pequena chance de que tenhamos visto mal o sorteio ou que o apresentador tenha lido mal o número vencedor ou que estejamos nos lembrando mal. Enquanto estamos nesse estado de espírito concessivo, somos capazes de abandonar nossa convicção de termos conhecimento. O cético silogiza a partir desta rendição: *Para qualquer proposição contingente, há uma declaração lotérica mais provável e que é desconhecida. Uma proposição conhecida não pode ser menos provável do que uma proposição desconhecida. Portanto, nenhuma proposição* *contingente é conhecida* (Hawthorne 2004). Isso é demais para aceitar! No entanto, as estatísticas do cético parecem impecáveis.

Esse paradoxo cético foi observado por Gilbert Harman (1968, p. 166), mas sua perspectiva sobre o papel da causação no conhecimento inferencial parecia resolver o problema (DeRose 2017, cap. 5). O paradoxo ainda bebê foi descartado como natimorto. Como o recém-chegado não recebeu o costumeiro batismo de atenção, os epistemólogos não notaram que a derrocada da teoria causal do conhecimento significava uma nova vida para o paradoxo da loteria de Harman.

As sugestões comuns do cético da probabilidade sobre como podemos estar enganados contrastam com as possibilidades extraordinárias evocadas pelo cético inspirado por René Descartes. O cético cartesiano tenta minar vastas faixas de conhecimento com uma única contra-explicação não testável à evidência (como a hipótese de que você está sonhando ou a hipótese de que um demônio maligno está enganando você). Essas alternativas abrangentes são projetadas para evitar qualquer refutação empírica. Já o cético probabilístico aponta para uma infinidade de contra-explicações simples. Todas são fáceis de testar: talvez você tenha confundido os dígitos de um número de telefone, talvez o agente de turismo tenha pensado que você queria voar para Moscou, Rússia, em vez de Moscou, Idaho, etc. Você pode verificar se há erros, mas qualquer verificação tem uma pequena chance de ter erros ela própria. Assim, sempre há algo a verificar, dado que os problemas não podem ser ignorados com base na improbabilidade.

Você pode verificar *qualquer um* desses possíveis erros, mas não pode verificar *todos* eles. Você não pode descartar essas simples possibilidades como ficção científica. Esses são exatamente os tipos de possibilidades que verificamos quando os planos dão errado. Por exemplo, você acha que sabe que tem o compromisso de encontrar uma possível empregadora para almoçar ao meio-dia. Quando ela não aparece no horário esperado, você para e pensa: Meu relógio está atrasado? Estou lembrando do restaurante certo? Poderia haver outro restaurante na cidade com o mesmo nome? Ela foi detida? Teria ela esquecido? Poderia ter havido uma falha de comunicação?

O ceticismo probabilístico remonta a Arcesilau, que assumiu a Academia duas gerações após a morte de Platão. Esse tipo moderado de ceticismo, relatado por Cícero (*Academica* 2.74, 1.46) sobre seus dias como aluno da Academia, permite a crença justificada. Muitos cientistas acham que devem atribuir apenas probabilidades, descartando a preocupação do epistemólogo com o conhecimento como antiquada.

Apesar do início precoce da teoria qualitativa da probabilidade, a teoria quantitativa não se desenvolveu até o estudo de Blaise Pascal sobre apostas no século XVII (Hacking 1975). Somente no século XVIII ela penetrou na indústria de seguros (embora as seguradoras tenham percebido que uma fortuna poderia ser feita calculando riscos com precisão). Apenas no século XIX a probabilidade deixou sua marca na física, e tão somente no século XX os probabilistas fizeram avanços importantes em relação a Arcesilau.

A maioria desses avanços filosóficos são reações ao uso da probabilidade por cientistas. No século XX, editores de periódicos científicos começaram a exigir que a hipótese dos autores fosse aceita somente quando fosse suficientemente provável — conforme medida por testes estatísticos. O limite para aceitação foi reconhecido como um tanto arbitrário, e também foi admitido que a regra de aceitação pode variar de acordo com os propósitos de alguém. Exigimos, por exemplo, uma probabilidade maior quando o custo de aceitar uma hipótese falsa é alto.

Em 1961, Henry Kyburg salientou que esta política entrava em conflito com um princípio de aglomeração: se você acredita racionalmente em *p* e racionalmente acredita em *q*, então você racionalmente acredita em *p* e *q*. Pequenas imagens de uma cena devem formar uma imagem maior da mesma cena. Se a crença racional pode ser baseada em uma regra de aceitação que requer apenas uma alta probabilidade, haverá crença racional em uma contradição! Para ver o porquê, suponha que a regra de aceitação permita a crença em qualquer proposição que tenha uma probabilidade de pelo menos 0,99. Dada uma loteria com 100 bilhetes e exatamente um ganhador, a probabilidade de "O bilhete n é perdedor" credencia a crença. Simbolizando proposições sobre um bilhete n ser perdedor como p_n, e "Eu acredito racionalmente" como B, segue-se uma crença numa contradição:

1. $B\neg(p_1 \wedge p_2 \wedge ... \wedge p_{100})$
   pela regra de aceitação probabilística.

2. $Bp_1 \wedge Bp_2 \wedge ... \wedge Bp_{100}$
   pela regra de aceitação probabilística.

3. $B(p_1 \wedge p_2 \wedge ... \wedge p_{100})$
   a partir de (2) e do princípio de que crenças racionais se aglomeram.

4. $B[(p_1 \wedge p_2 \wedge ... \wedge p_{100}) \wedge \neg(p_1 \wedge p_2 \wedge ... \wedge p_{100})]$
   a partir de (1) e (3) pelo princípio de que crenças racionais se aglomeram.

Mais informalmente, a regra de aceitação implica isto: cada crença de que um bilhete em particular perderá é provável o suficiente para justificar que se acredite nela. Por aplicações repetidas do princípio de aglomeração, juntar todas essas crenças justificadas resulta em uma crença justificada. Finalmente, juntar essa crença com a crença justificada de que um dos bilhetes é um vencedor resulta na crença contraditória de que todos perderão e um ganhará. No entanto, por aglomeração isso também é justificado.

Como a crença em uma contradição óbvia é um exemplo paradigmático de irracionalidade, Kyburg apresenta um dilema: rejeitar a aglomeração ou rejeitar regras que licenciam a crença para uma probabilidade menor que um. (Martin Smith 2016, pp. 186–196) adverte que mesmo a probabilidade de um leva à inconsistência conjunta para uma loteria que tem infinitamente muitos bilhetes. Kyburg rejeita a aglomeração. Ele promove a tolerância à inconsistência *conjunta* (ter crenças que não podem ser todas verdadeiras juntas) para evitar a crença em contradições. A razão nos proíbe de acreditar em uma proposição que é necessariamente falsa, mas nos permite ter um conjunto de crenças que necessariamente contém uma falsidade. A escolha de Henry Kyburg logo foi apoiada pela descoberta de um paradoxo relacionado.

1.5 Paradoxo do Prefácio

No prefácio de *Introduction to the Foundations of Mathematics*, Raymond Wilder (1952, p. iv) pede desculpas pelos erros em seu texto. A reimpressão de 1982 tem três páginas de errata que sustentam a humildade de Wilder. D. C. Makinson (1965, p. 205) cita o pedido de desculpas de Wilder de 1952 e extrai um paradoxo: Wilder acredita racionalmente em cada uma das afirmações apresentadas em seu livro. Contudo, como Wilder se considera falível, ele também acredita racionalmente que a conjunção de todas as suas afirmações é falsa. Se o princípio da aglomeração for válido, $(Bp \mathbin{\&} Bq) \rightarrow B(p \mathbin{\&} q)$, Wilder acredita racionalmente na conjunção de todas as afirmações em seu livro e também desacreditaria racionalmente na mesma coisa!

O paradoxo do prefácio não depende de uma regra de aceitação probabilística. A crença do prefácio é gerada organicamente de forma qualitativa. O autor está meramente refletindo, de maneira humilde, sobre sua semelhança com outros autores que são falíveis, sobre seus próprios erros do passado posteriormente descobertos, sobre sua imperfeição na verificação de fatos, e assim por diante.

Nesta conjuntura, muitos filósofos se juntam a Kyburg ao rejeitarem a aglomeração e concluem que pode ser racional ter crenças conjuntamente inconsistentes. A solução de Kyburg para o paradoxo do prefácio levanta uma questão metodológica sobre a natureza dos paradoxos. Como paradoxos podem mudar nossas maneiras de pensar se a inconsistência conjunta for permitida? Um paradoxo é comumente definido como um conjunto de proposições que são individualmente plausíveis, mas conjuntamente inconsistentes. A inconsistência funciona como o incômodo que nos impele a remover um dos membros do conjunto (ou como a dor que nos leva a nos afastarmos do estímulo). A título ilustrativo, muito trabalho epistemológico orbita em torno de um antigo enigma colocado pelo regresso da justificação: Qual das seguintes opções é falsa?

1. Uma crença só pode ser justificada por outra crença justificada.

2. Não há cadeias circulares de justificação.

3. Todas as cadeias de justificação são finitas.

4. Algumas crenças são justificadas.

Os fundacionalistas rejeitam (1). Eles tomam algumas proposições como sendo auto-evidentes ou permitem que crenças sejam justificadas por não-crenças (como percepções ou intuições). Os coerentistas rejeitam (2), tolerando algumas formas de raciocínio circular. Nelson Goodman (1965), por exemplo, caracterizou o método do equilíbrio reflexivo como virtuosamente circular. Charles Sanders Peirce (1933–35, 5.250) possivelmente rejeitava (3), mas o primeiro a claramente rejeitá-la foi Peter Klein (2007). Para um livro defendendo essa ideia, leia Scott F. Aikin (2011). Os infinitistas acreditam que cadeias infinitamente longas de justificação não são mais impossíveis do que cadeias infinitamente longas de causalidade. Finalmente, os anarquistas epistemológicos rejeitam (4) — como Paul Feyerabend, que entoa em *Contra o Método*, “vale tudo” (1988, vii, 5, 14, 19, 159).

Formular um paradoxo como um conjunto de crenças individualmente plausíveis, mas conjuntamente inconsistentes, é um feito de compressão de dados. Mas se a inconsistência conjunta é racionalmente tolerável, por que esses filósofos se preocupam em oferecer soluções para paradoxos como o do regresso da justificação? Por que é irracional acreditar nas proposições (1)–(4), apesar de sua inconsistência conjunta?

Kyburg pode responder que há um efeito de escala. Embora a sensação de inconsistência conjunta seja tolerável quando difusamente distribuída por um grande corpo de proposições, a sensação passa a ser incômoda quando a inconsistência torna-se localizada (Knight

2002). É por isso que os paradoxos são sempre representados como um pequeno conjunto de proposições. Um paradoxo é melhorado pela redução no número de suas proposições — como quando um membro do conjunto é exposto como supérfluo à inconsistência (ainda que, a rigor, um conjunto só possa mudar de tamanho no sentido metafórico em que um número pode crescer ou encolher).

Se você sabe que suas crenças são conjuntamente inconsistentes, mas nega que isso cria um grande paradoxo, então você deve rejeitar a definição de paradoxo oferecida por R. M. Sainsbury como "uma conclusão aparentemente inaceitável derivada de um raciocínio aparentemente aceitável a partir de premissas aparentemente aceitáveis" (1995, p. 1). Considere a negação de qualquer uma de suas crenças como uma conclusão e suas crenças restantes como as premissas. Você deve julgar esse argumento confuso como válido, e como tendo premissas que você aceita, e ainda assim como tendo uma conclusão que você rejeita (Sorensen 2003b, pp. 104–110). Se a conclusão desse argumento conta como um paradoxo, então a negação de qualquer uma de suas crenças conta como um paradoxo.

A semelhança entre o paradoxo do prefácio e o paradoxo da prova surpresa se torna mais visível por meio de um caso intermediário. O prefácio de *The Emperor of All Maladies: A Biography of Cancer*, de Siddhartha Mukherjee, alerta: "Em casos em que não havia conhecimento público prévio, ou quando os entrevistados solicitaram privacidade, usei um nome falso e deliberadamente confundi identidades para dificultar o rastreamento" (2010, p. xiv). Aqueles que se recusam a consentir em serem enganados são livres para fechar a crônica do Dr. Mukherjee, mas quase todos os leitores consideram aceitável a troca de mentiras por novas informações oferecida pelo médico. Eles antecipam racionalmente serem racionalmente enganados. No entanto, esses leitores aprendem muito sobre a história do câncer. Da mesma forma, os alunos que são avisados de que receberão uma prova surpresa esperam racionalmente ser racionalmente enganados sobre o dia do teste. A perspectiva de serem enganados não os leva a abandonar o curso.

O paradoxo do prefácio pressiona Kyburg a estender sua tolerância à inconsistência conjunta para a aceitação de contradições, pois o exemplo original de Makinson expressa o pesar de um lógico ao afirmar contradições, ao invés de falsas declarações contingentes. Considere um estudante de lógica que é obrigado a escolher cem verdades de uma lista mista de tautologias e contradições (Sorensen 2001, pp. 156–158). Embora o modesto estudante acredite em cada uma de suas respostas, $A_1, A_2, \dots, A_{100}$, ele também acredita que pelo menos uma dessas respostas é falsa — o que garantiria que ele acredita em uma contradição. Se alguma de suas respostas for falsa, então o aluno acredita em uma contradição (porque as únicas falsidades na lista são contradições). Se todas as suas respostas do teste forem verdadeiras, então o aluno acredita na seguinte contradição: $\neg(A_1 \wedge A_2 \wedge \ldots \wedge A_{100})$. Afinal, uma conjunção de tautologias é em si uma tautologia e a negação de qualquer tautologia é uma contradição.

Se os paradoxos fossem sempre conjuntos de proposições, argumentos ou conclusões, então eles sempre teriam significado. Mas alguns paradoxos são semanticamente falhos (Sorensen 2003b, p. 352) e alguns têm respostas que são apoiadas por um pseudo-argumento que emprega um "lema" defeituoso que não tem um valor de verdade. O paradoxo de Kurt Grelling, por exemplo, começa com uma distinção entre palavras autológicas e heterológicas. Uma palavra autológica descreve a si mesma; por exemplo, "polissilábico" é polissilábica, "português" é uma palavra em português,[3] "substantivo" é um substantivo, etc. Uma palavra heterológica não descreve a si mesma; por exemplo, "monossilábico" não é monossilábica, "chinês" não é uma palavra em chinês, "verbo" não é um verbo, etc. Agora, o enigma: "heterológico" é uma palavra heterológica ou autológica? Se "heterológico" é heterológica, então, uma vez que se descreve, é autológica. Mas se "heterológico" é autológica, então, como é uma palavra que não se descreve, é heterológica. A solução comum para esse enigma é que "heterológico", como definido por Grelling, não é um predicado bem definido (Thomson 1962). Em outras palavras, " 'heterológico' é heterológico?" não tem significado. Não pode haver um predicado que se aplique a todos e somente àqueles predicados aos quais ele não se aplica, pela mesma razão que não pode haver um barbeiro que faça a barba de todas e somente daquelas pessoas que não se barbeiam.

O eliminativista, que pensa que "saber" ou "justificado" não possuem significado, diagnosticará os paradoxos epistêmicos como questões que apenas *parecem* bem-formadas. Por exemplo, o eliminativista sobre justificação não aceitaria a proposição (4) no paradoxo do regresso: "Algumas crenças são justificadas". Seu ponto não é que nenhuma crença atende aos altos padrões de justificação, tal como um anarquista pode negar que quaisquer autoridades ostensivas atendem aos altos padrões de legitimidade. Em vez disso, o eliminativista diagnostica "justificado", sem romantismo, como um termo patológico. Assim como o astrônomo ignora "Existem um zilhão de estrelas?" com base no fato de que "zilhão" não é um numeral genuíno, o eliminativista ignora "Algumas crenças são justificadas?" com base em que "justificado" não é um adjetivo genuíno.

[3]N.T. No original, *"english" is english*.

No século XX, as suspeitas sobre patologia conceitual eram mais fortes para o paradoxo do mentiroso: “Esta frase é falsa” é verdadeira? Filósofos que pensavam que havia algo profundamente defeituoso com o paradoxo da prova surpresa o assimilaram ao paradoxo do mentiroso. Vamos dar uma olhada no processo de assimilação.

1.6 Anti-especialização

No paradoxo da prova surpresa, as premissas do aluno são auto-derrotantes. Qualquer razão que o aluno tenha para prever uma data para a prova ou uma data sem prova está disponível para o professor. Assim, o professor pode simular a previsão do aluno e saber o que o aluno espera.

A conclusão geral do aluno, de que a prova é impossível, também é auto-derrotante. Se o aluno acredita em sua conclusão, então ele não aguardará a prova. Logo, se ele receber uma prova, será uma surpresa. O evento será ainda mais inesperado porque o aluno se iludiu pensando que a prova era impossível.

Assim como a consciência de alguém sobre uma previsão pode afetar a probabilidade de ela ser verdadeira, a consciência dessa sensibilidade à sua consciência também pode afetar sua verdade. Se cada ciclo de consciência é autodestrutivo, então não há espaço estável para tirar uma conclusão.

Suponha que um psicólogo lhe ofereça uma caixa vermelha e uma caixa azul (Skyrms 1982). O psicólogo pode prever qual caixa você escolherá com 90% de precisão. Ele colocou um dólar na caixa que ele prevê que você escolherá e dez dólares na outra caixa. Você deve escolher a caixa vermelha ou a caixa azul? Não há como tomar uma decisão, pois qualquer escolha se torna uma razão para revertê-la.

Paradoxos epistêmicos afetam a teoria da decisão porque escolhas racionais são baseadas em *crenças* e desejos. Se o agente não consegue formar uma crença racional, é difícil interpretar seu comportamento como uma *escolha*. Na teoria da decisão, todo o ponto de atribuir crenças e desejos é estabelecer silogismos práticos que dão sentido às ações como meios para fins. Subtrair a racionalidade do agente torna esse esquema inútil. Dado esse compromisso com a interpretação caridosa, não há a possibilidade de você escolher racionalmente uma opção que você acredita ser inferior. Assim, ao realizar uma escolha, você não pode realmente acreditar que estava operando como um anti-especialista, isto é, alguém cujas opiniões sobre um tópico são confiavelmente erradas (Egan e Elga 2005).

O filósofo medieval John Buridan (*Sophismata*, Sophism 13) deu um exemplo notavelmente minimalista de tal instabilidade:

(B) Você não acredita nesta sentença.

Se você acredita em (B), então ela é falsa. Se você não acredita em (B), então é verdadeira. Você é um anti-especialista sobre (B); sua opinião está confiavelmente errada. Um estranho que monitora sua opinião pode avaliar se (B) é verdadeira, mas você não é capaz de fazer uso de sua anti-especialização.

O lado positivo disso é que você é capaz de explorar a anti-especialização dos outros. Quatro em cada cinco anti-especialistas recomendam parar agora a leitura!

1.6.1 O Paradoxo do Conhecedor

David Kaplan e Richard Montague (1960) consideram que o anúncio da professora em nosso exemplo da prova surpresa é equivalente ao autorreferencial

(K-3)

> Ou a prova é na segunda-feira, mas você não sabe disso antes de segunda-feira, ou a prova é na quarta-feira, e você não sabe disso antes de quarta-feira, ou a prova é na sexta-feira, mas você não sabe disso antes de sexta-feira, ou este anúncio é sabidamente falso.

Kaplan e Montague observam que o número de datas alternativas da prova pode ser aumentado indefinidamente. Surpreendentemente, eles afirmam que o número de alternativas pode ser reduzido a zero! O anúncio é então equivalente a

(K-0)

> Esta frase é sabidamente falsa.

Se (K-0) é verdadeiro, então é sabido que é falsa. Tudo o que é sabidamente falso é falso. Como nenhuma proposição pode ser verdadeira e falsa, provamos que (K-0) é falsa. Dado que a prova produz conhecimento, (K-0) é sabidamente falsa. Mas espere! É exatamente isso que (K-0) diz — então (K-0) deve ser verdadeira.

O argumento (K-0) tem uma semelhança suspeita com o paradoxo do mentiroso. Comentadores subsequentes trocam desleixadamente o sinal de negação nas apresentações formais do raciocínio de $K\neg p$ para $\neg Kp$ (isto é, de "é sabido que não-p", para "não é o caso que é sabido que p"). Ironicamente, essa transmissão distorcida resulta em uma variação mais limpa do conhecedor:

(K) Ninguém sabe esta sentença.

(K) é verdadeira? Por um lado, se (K) é verdadeira, então o que ela diz é verdadeiro e, portanto, ninguém a sabe. Por outro lado, esse mesmo raciocínio parece ser uma prova de (K). Acreditar em uma proposição ao vê-la ser provada é suficiente para o conhecimento dela, então alguém deve saber (K), mas então (K) é falsa! Já que ninguém pode saber uma proposição que é falsa, (K) não é sabida.

Um cético poderia esperar resolver (K-0) negando que qualquer coisa seja conhecida. Essa tática, contudo, não remedia (K). Se nada é conhecido, então (K) é verdadeira. Poderia o cético, em vez disso, desafiar a premissa de que uma prova sua é suficiente para seu conhecimento? Esta solução seria particularmente embaraçosa para o cético, que se apresenta como um defensor da prova. Se acontecer de nem mesmo a prova o influenciar, ele terá uma semelhança condenável com o dogmático por ele tão frequentemente repreendido.

Mas o cético não deve perder a cabeça. Provas nem sempre produzem conhecimento. Considere um aluno que adivinha corretamente que um passo em sua prova é válido. O aluno não sabe a conclusão, mas provou o teorema. Seu instrutor pode ter dificuldade em fazer o aluno entender por que sua resposta constitui uma prova válida. A intransigência pode advir da inteligência daquele que realiza a prova, e não de sua estupidez. L. E. J. Brouwer é mais conhecido em matemática por seu brilhante teorema do ponto fixo. Mas uma leitura duvidosa da filosofia da matemática de Immanuel Kant levou Brouwer a retratar sua prova. Brouwer também tinha dúvidas filosóficas sobre o Axioma da Escolha e a Lei do Terceiro Excluído. Brouwer persuadiu uma minoria de matemáticos e filósofos, conhecidos como intuicionistas, a imitar sua abstenção de provas não construtivas. Isso os levou a desenvolver provas construtivas de teoremas que foram provados anteriormente por meios menos informativos.

Todos concordam que aprende-se mais com uma prova de uma generalização existencial que procede a partir de uma instância provada do que com uma *reductio ad absurdum* inespecífica da generalização universal correspondente. Mas isso não justifica a recusa dos intuicionistas em serem persuadidos pela *reductio ad absurdum*. O intuicionista, mesmo aos olhos do cético, tem um padrão de prova muito alto, e um padrão de prova excessivamente alto pode impedir o conhecimento por prova.

O mito lógico de que “Você não pode provar uma negativa universal” é em si uma negativa universal, implicando sua própria improvabilidade. Essa implicação de improvabilidade está correta somente porque o princípio é falso. Por exemplo, uma inspeção exaustiva prova a negativa universal “Não há advérbios de lugar nesta frase”[4]. Uma *reductio ad absurdum* prova a negativa universal “Não há maior número primo”.

Trivialmente, proposições falsas não podem ser provadas como verdadeiras. Há alguma proposição verdadeira que não pode ser provada como verdadeira?

Sim, há uma infinidade. O teorema da incompletude de Kurt Gödel demonstrou que qualquer sistema que seja forte o suficiente para expressar a aritmética também é forte o suficiente para expressar uma contraparte formal da proposição autorreferencial presente no exemplo da prova surpresa “Esta afirmação não pode ser provada neste sistema”. Se o sistema não puder provar sua “sentença de Gödel”, então esta frase é verdadeira. Se o sistema puder provar sua sentença de Gödel, o sistema é inconsistente. Então, ou o sistema é incompleto ou inconsistente. (Veja a entrada sobre Kurt Gödel.)

Com efeito, esse resultado diz respeito à provabilidade relativa a um sistema. Um sistema pode provar a sentença de Gödel de outro sistema. Kurt Gödel (1983, p. 271) considerava que provas não eram necessárias para o conhecimento de que a aritmética é consistente.

J. R. Lucas (1964) afirma que isso revela que os seres humanos não são máquinas. Um computador é uma instância concreta de um sistema formal. Portanto, seu “conhecimento” é restrito ao que ele pode provar. Pelo teorema de Gödel, o computador será inconsistente ou incompleto. No entanto, qualquer ser humano *poderia* ter um conhecimento consistente e completo de aritmética. Portanto, necessariamente, nenhum ser humano é um computador.

Os críticos de Lucas defendem a paridade entre pessoas e computadores. Eles acham que temos nossas próprias sentenças de Gödel (Lewis 1999, pp. 166–173). Nesse espírito igualitário, G. C. Nerlich (1961) modela as crenças do aluno no exemplo da prova surpresa como um sistema lógico. O anúncio da professora é então uma sentença de Gödel sobre o aluno: Haverá uma prova na próxima semana, mas você não será capaz de provar em que dia ela ocorrerá com base nesse anúncio e na memória do que aconteceu nos dias de exame

[4]N.T. Em inglês “No adverbs appear in this sentence”. Contudo, “não” é considerado um advérbio da língua portuguesa, enquanto “no” tende a não ser classificado como advérbio por gramáticas do inglês.

anteriores. Quando o número de dias de exame é igual a zero, o anúncio é equivalente à sentença K.

Vários comentadores do paradoxo da prova surpresa objetam que interpretar surpresa como improvabilidade muda o tópico em questão. Ao invés de propor o paradoxo da prova surpresa, propõe-se uma variação do paradoxo do mentiroso. Outros conceitos podem ser misturados com o mentiroso como, por exemplo, noções aléticas, gerando o mentiroso possível: "Esta afirmação é possivelmente falsa" é verdadeira? (Posts 1970) (Se for falsa, então é falso que seja possivelmente falsa. O que não pode ser falso é necessariamente verdadeiro. Mas se for necessariamente verdadeiro, então não pode ser possivelmente falso). Dado que o conceito semântico de validade envolve a noção de possibilidade, também se pode derivar mentirosos da validade como o paradoxo de Pseudo-Scotus: "Quadrados são quadrados, portanto, este argumento é inválido" (Read 1979).

Suponha que o argumento de Pseudo-Scotus seja válido. Como a premissa é necessariamente verdadeira, a conclusão seria necessariamente verdadeira. Mas a conclusão contradiz a suposição de que o argumento é válido. Portanto, por *reductio*, o argumento é necessariamente inválido. Espere! O argumento pode ser inválido somente se for possível que a premissa seja verdadeira e a conclusão seja falsa. Mas já provamos que a conclusão de "Quadrados são quadrados, portanto, este argumento é inválido" é necessariamente verdadeira. Não há julgamento consistente da validade do argumento. Uma situação semelhante decorre de "A prova é na sexta-feira, mas esta previsão não pode ser deduzida corretamente deste anúncio".

Pode-se criar um complicado paradoxo do mentiroso que se assemelha ao paradoxo da prova surpresa, mas essa variante complexa do mentiroso não é um paradoxo *epistêmico*. Isso porque tais paradoxos giram em torno do conceito semântico de verdade, em vez de um conceito epistêmico.

1.6.2 O "Paradoxo da Cognoscibilidade"

Frederic Fitch (1963) relata que em 1945 ele soube pela primeira vez dessa prova de verdades não conhecíveis por meio do relatório de um parecerista sobre um manuscrito que ele nunca veio a publicar. Graças à pesquisa arquivística de Joe Salerno (2009), agora sabemos que o parecerista foi Alonzo Church.

Assuma que há uma sentença verdadeira da forma "*p* mas *p* não é conhecida". Embora esta frase seja consistente, princípios modestos de lógica epistêmica implicam que sentenças desta forma são incognoscíveis.

1. $K(p \wedge \neg Kp)$ *(Hipótese).*

2. $Kp \wedge K\neg Kp$ (de 1; conhecimento distribui sobre a conjunção).

3. $\neg Kp$ (de 2; conhecimento implica verdade, do segundo membro).

4. $Kp \wedge \neg Kp$ (de 2 e 3, por eliminação do segundo membro e depois introdução da conjunção).

5. $\neg K(p \wedge \neg Kp)$ (de 1 e 4, *reductio ad absurdum*).

Como todas as suposições foram descartadas, a conclusão é uma verdade necessária. Portanto, é uma verdade necessária que $p \wedge \neg Kp$ não é conhecida. Em outras palavras, $p \wedge \neg Kp$ é incognoscível.

Os cautelosos tiram uma conclusão condicional: se há verdades realmente desconhecidas, há verdades incognoscíveis. Não obstante, alguns filósofos rejeitam o antecedente porque acreditam que há um ser onisciente.

Mas idealistas seculares e positivistas lógicos admitem que há algumas verdades atuais desconhecidas. Como eles podem continuar a acreditar que todas as verdades são cognoscíveis? Surpreendentemente, esses eminentes filósofos parecem refutados pela pitada de lógica epistêmica que acabamos de ver. Também são prejudicados aqueles que limitam suas alegações de cognoscibilidade universal a um domínio restrito. Por exemplo, Immanuel Kant (A223/B272) afirma que todas as proposições empíricas são cognoscíveis. Esse pouco de otimismo seria suficiente para acender a contradição (Stephenson 2015).

Timothy Williamson duvida que essa lista de posições atingidas seja suficiente para que o resultado se qualifique como um paradoxo:

> A conclusão de que há verdades incognoscíveis é uma afronta a várias teorias filosóficas, mas não ao senso comum. Se os proponentes (e oponentes) dessas teorias ignoraram por muito tempo um simples contraexemplo, isso é um constrangimento, não um paradoxo. (2000, 271)

Aqueles que acreditam que o resultado de Church-Fitch é um paradoxo genuíno podem responder a Williamson com paradoxos que concordam com o senso comum (e com a ciência e a ortodoxia religiosa). Por exemplo, o senso comum concorda convictamente com a conclusão de que algo existe. Mas é surpreendente que isso possa ser provado sem premissas empíricas. Como os quantificadores da lógica tradicional (lógica de predicados de primeira ordem com identidade) têm uma conotação existencial, o lógico pode deduzir que algo existe a partir do princípio de que tudo é idêntico a si mesmo. A maioria dos filósofos se opõe a essa prova simples porque eles sentem que a existência de algo não pode ser provada por pura lógica. Eles não estão se opondo à afirmação que está de acordo com o senso comum (de que algo existe), mas à afirmação de que isso possa ser provado por pura lógica. Da mesma forma, muitos filósofos que concordam que há sentenças incognoscíveis resistem unicamente à ideia de que um resultado tão profundo possa ser alcançado por meios tão restritos.

1.6.3 O problema de Moore

O parecer de Church foi escrito em 1945. O momento e a estrutura de seu argumento das sentenças incognoscíveis sugerem que Church pode ter sido inspirado pela sentença de G. E. Moore (1942, 543):

(M)

Fui ao cinema na terça-feira passada, mas não acredito que eu tenha ido.

O problema de Moore é explicar a estranheza presente em proferimentos declarativos como (M). Esta explicação precisa abranger ambas as leituras de (M): "$p \wedge B\neg p$" e "$p \wedge \neg Bp$". (Essa ambiguidade de escopo é explorada por uma piada popular: René Descartes está sentado em um bar, tomando uma bebida. O barman pergunta "O senhor pensa em pedir mais alguma coisa?", Descartes responde "Não penso" e então desaparece. A piada é comumente criticada por ser falaciosa — não faz parte da crença de Descartes ele ser essencialmente pensante).

Uma explicação usual da absurdidade presente na sentença de Moore é que o falante conseguiu se contradizer sem proferir uma contradição. Portanto, a frase seria estranha por ser um contraexemplo à generalização de que qualquer um que se contradiz profere uma contradição.

Não há problema com contrapartes em terceira pessoa do proferimento (M). Qualquer outra pessoa pode dizer sobre Moore, sem paradoxo, "G. E. Moore foi ao cinema na terça-feira passada, mas não acredita nisso". (M) também pode ser incorporada sem paradoxo em condicionais: “Se eu fui ao cinema na terça-feira passada, mas não acredito, então estou sofrendo de um lapso de memória preocupante”. No pretérito também funciona: “Fui ao cinema na terça-feira passada, mas não acreditei ter ido”. No tempo futuro fica um pouco mais forçado: “Fui ao cinema na terça-feira passada, mas não vou acreditar nisso” (Bovens 1995). Temos a tendência de imaginar nossos eus futuros como mais bem informados. Eus do futuro são, por assim dizer, especialistas a quem eus anteriores devem se submeter. Quando um eu anterior prevê que seu eu do futuro acredita que p, essa predição passa a ser uma razão para crer que p. Bas van Fraassen (1984, 244) chama isso de “princípio da reflexão”: devo acreditar em uma proposição, dado que acreditarei nela em algum momento futuro.

Robert Binkley (1968) antecipa van Fraassen ao aplicar o princípio da reflexão ao paradoxo da prova surpresa. O aluno pode prever que não acreditará na declaração se nenhuma prova for dada até quinta-feira. A *conjunção* da história dos dias sem prova mais a declaração implicará na sentença mooreana:

(A’) O teste é na sexta-feira, mas você não acredita nisso.

Como o membro menos evidente da conjunção é a declaração, o aluno escolherá não acreditar nela. No começo da semana, o aluno prevê que seu eu futuro pode não acreditar no anúncio. Portanto, o aluno no domingo não acreditará na declaração quando ela for proferida pela primeira vez.

Binkley ilumina esse raciocínio com lógica doxástica (“doxa” é o termo grego para crença). As regras de inferência para essa lógica de crença podem ser entendidas como idealizando o aluno para ser um raciocinador ideal. Em termos gerais, um raciocinador ideal é alguém que infere o que deve inferir e se abstém de inferir mais do que deve. Como não há restrição em suas premissas, podemos discordar do raciocinador ideal. Mas se concordamos com as premissas do raciocinador ideal, parece que somos obrigados a concordar com sua conclusão. Binkley especifica alguns requisitos para dar força ao *status* do aluno como um raciocinador ideal: o aluno é perfeitamente consistente, acredita em todas as consequências lógicas de suas crenças e não se esquece. Binkley ainda assume que o raciocinador ideal está ciente de que ele é um raciocinador ideal. Segundo Binkley, isso garante que se o raciocinador ideal acredita que p, então ele acredita que acreditará que p posteriormente.

A descrição de Binkley sobre o estado epistêmico hipotético do aluno na quinta-feira é convincente. Mas seu argumento para espalhar a incredulidade do futuro para o passado está sujeito a três desafios.

A primeira objeção é que ela fornece o resultado errado. O aluno *é* informado pela declaração da professora, então Binkley não deveria usar um modelo em que o anúncio é tão absurdo quanto a conjunção “Fui ao cinema na terça-feira passada, mas não acredito nisso”.

Em segundo lugar, o futuro estado mental previsto por Binkley é apenas hipotético: se nenhum teste for dado até quinta-feira, o aluno achará a declaração inacreditável. No começo da semana, o aluno não sabe (ou acredita) que o professor vai esperar tanto tempo. O princípio da reflexão “Deferir às opiniões do meu eu futuro” não implica que eu deva me submeter às opiniões do meu eu futuro *hipotético*, pois meu eu futuro hipotético está reagindo a proposições que não precisam ser realmente verdadeiras.

Em terceiro lugar, o princípio da reflexão pode precisar de mais qualificações do que Binkley antecipa. Binkley percebe que um agente comum prevê que esquecerá detalhes. É por isso que escrevemos lembretes para nós mesmos. Um agente comum prevê que pode passar por períodos de juízo prejudicado. É por isso que limitamos quanto dinheiro levamos ao bar.

Binkley estipula que os alunos não se esquecem. Ele precisa acrescentar que os alunos sabem que não esquecerão, pois a mera ameaça de um lapso de memória às vezes é suficiente para minar o conhecimento. Considere o esquema do Professor de Anestesiologia para provas surpresa: “Uma prova surpresa será dada na quarta ou sexta-feira com a ajuda de um medicamento para amnésia. Se o teste ocorrer na quarta-feira, o medicamento será administrado cinco minutos após a aula de quarta-feira. O medicamento apagará instantaneamente a memória da prova e os alunos preencherão a lacuna por confabulação”. Você acabou de concluir a aula de quarta-feira e, portanto, sabe temporariamente que a prova será na sexta-feira. Dez minutos após a aula, você perde esse conhecimento. Nenhuma droga foi administrada e não há nada de errado com sua memória. Você está se lembrando corretamente de que nenhuma prova foi dada na quarta-feira. No entanto, você não sabe se sua memória é acurada porque também sabe que se o teste fosse dado na quarta-feira, você teria uma pseudo memória indistinguível de sua memória atual.

Apesar de não ganhar nenhuma nova evidência, você muda de ideia sobre a prova ocorrer na quarta-feira e perde seu conhecimento de que a prova será na sexta-feira. (A mudança de crença não é crucial; você ainda não teria *conhecimento* prévio da prova mesmo se persistisse dogmaticamente em acreditar que a prova será na sexta-feira).

Se os alunos sabem que não esquecerão e sabem que não haverá enfraquecimento por evidências externas, então podemos estar inclinados a concordar com a formulação de Binkley, segundo a qual seu aluno idealizado nunca perde o conhecimento que acumula. Como veremos, no entanto, isso ignora outras maneiras pelas quais agentes racionais podem perder conhecimento.

1.6.4 Pontos Cegos

"Eu sou um poeta, mas não sei disso" expressa uma proposição que não posso saber, mas que posso acessar por meio de outras atitudes, como esperar e desejar. Um ponto cego para uma atitude proposicional é uma proposição consistente que não pode ser acessada por essa atitude. Pontos cegos são relativos aos meios de acessar a proposição, à pessoa que faz a tentativa e ao momento em que ela tenta acessá-la. Embora *eu* não possa racionalmente acreditar que "ursos polares têm pele negra, mas eu acredito que eles não têm pele negra", *você* pode acreditar que eu erroneamente acredito que ursos polares não têm pele negra. A evidência que o convence de que estou cometendo esse erro não pode me convencer de que estou cometendo esse erro. Essa é uma assimetria imposta pela racionalidade ao invés da irracionalidade. Atribuições de erros específicos são pontos cegos pessoais para a pessoa que supostamente errou.

O antropólogo Gontran de Poncins começa seu capítulo sobre o missionário ártico, Padre Henry, com uma previsão:

> Vou dizer a você que um ser humano pode viver sem reclamar em uma casa de gelo construída para focas a uma temperatura de cinquenta e cinco graus abaixo de zero, e você vai duvidar da minha palavra. No entanto, o que eu digo é verdade, pois era assim que o Padre Henry vivia; .... (Poncins 1941 [1988], 240)

O testemunho subsequente de Gontran de Poncins pode levar o leitor a acreditar que alguém pode de fato se contentar em viver em uma casa de gelo. O mesmo testemunho pode levar outro leitor a acreditar que Poncins não está dizendo a verdade. Mas nenhum leitor deve acreditar que "Alguém pode se contentar em viver em uma casa de gelo e todo mundo acredita que isso não é o caso". Esse é um ponto cego universal.

Se Gontran acredita em uma proposição que é um ponto cego para seu leitor, então ele não pode fornecer bons motivos para que seu leitor compartilhe sua crença. Isso vale mesmo se eles forem raciocinadores ideais. Por conseguinte, uma implicação dos pontos cegos pessoais é que pode haver desacordo entre raciocinadores ideais na medida em que eles diferem no que concerne aos seus pontos cegos.

Isso é relevante para o paradoxo da prova surpresa. Os alunos são os surpreendidos pela prova. Como a declaração implica que a data da prova surpresa é um ponto cego para eles, os não surpreendidos não podem persuadi-los.

O mesmo ponto vale para desacordos intrapessoais ao longo do tempo. Evidências que me persuadiram no domingo de que "Este código de segurança é 390524, mas na sexta-feira não vou acreditar nisso" não devem mais me persuadir na sexta-feira (dada minha crença de que o dia é sexta-feira), pois essa proposição é um ponto cego para o meu eu de sexta-feira.

Embora cada ponto cego seja inacessível, uma disjunção de pontos cegos normalmente não é um ponto cego. Posso racionalmente acreditar que "Ou o número de estrelas é par e eu não acredito nisso, ou o número de estrelas é ímpar e eu não acredito nisso". A declaração do autor no prefácio de que há algum erro em seu livro é equivalente a uma disjunção muito longa de pontos cegos. O autor está dizendo que ele ou acredita falsamente em sua primeira declaração, ou acredita falsamente em sua segunda declaração, ou ..., ou acredita falsamente em sua última declaração.

O anúncio da professora de que haverá uma prova surpresa equivale a uma disjunção de erros futuros: "Ou haverá uma prova na segunda-feira e o aluno não acreditará de antemão, ou haverá uma prova na quarta-feira e o aluno não acreditará de antemão, ou a prova será na sexta-feira e o aluno não acreditará de antemão."

Os pontos levantados até agora sugerem uma solução para o paradoxo da prova surpresa (Sorensen 1988, 328–343). Como Binkley (1968) afirma, a prova seria uma surpresa mesmo se a professora esperasse até o último dia para aplicá-la, e ainda poderia ser verdade que o anúncio da professora é informativo. No início da semana, os alunos estão justificados em acreditar no anúncio da professora de que haverá uma prova surpresa. Esse anúncio é equivalente a:

(A) Ou

i. o teste é na segunda-feira e o aluno não sabe disso antes de segunda-feira, ou

ii. o teste é na quarta-feira e o aluno não sabe disso antes de quarta-feira, ou

iii. o teste é na sexta-feira e o aluno não sabe disso antes de sexta-feira.

Considere a situação do aluno na quinta-feira (dado que o teste não foi na segunda ou quarta-feira). Se ele sabe que nenhum teste foi dado, ele não pode saber também que (A) é verdadeira. Porque isso implicaria:

iii. O teste é na sexta-feira e o aluno não sabe disso antes de sexta-feira.

Embora (iii) seja consistente e possa ser sabida por outras pessoas, (iii) não pode ser sabida pelo aluno antes de sexta-feira. (iii) é um ponto cego para os alunos, mas não para, digamos, os colegas da professora. Portanto, a professora pode dar uma prova surpresa na sexta-feira na medida em que isso forçaria os alunos a perderem o conhecimento do anúncio original (A). O conhecimento pode ser perdido sem que nada seja esquecido.

Esta solução torna quem você é um fator relevante para o que você pode saber. Além de comprometer a impessoalidade do conhecimento, compromete-se também sua neutralidade temporal.

Como o paradoxo da prova surpresa também pode ser formulado em termos de crença racional, haverá ajustes paralelos para o que devemos acreditar. Somos criticados por falhar em crer nas consequências lógicas do que acreditamos e criticados por acreditar em proposições que entram em conflito entre si. Qualquer um que atenda a esses ideais de completude e consistência será incapaz de acreditar em uma gama de proposições consistentes que são acessíveis a outros pensadores completos e consistentes. Em particular, essas pessoas não serão capazes de acreditar em proposições que atribuem erros específicos a elas próprias, e proposições que implicam essas proposições inacessíveis.

Algumas pessoas usam camisetas com *Question Authority!* escrito nelas. Questionar autoridade é geralmente considerado uma questão de critério individual. O paradoxo da prova surpresa mostra que às vezes é obrigatório. O aluno é racionalmente obrigado a duvidar da declaração da professora, mesmo que a professora não tenha dado nenhuma nova evidência de não ser confiável. Isso porque quando resta apenas um dia, o anúncio implica (iii), uma declaração que é impossível para o aluno saber. O aluno pode prever que essa perda forçada de conhecimento abre uma oportunidade para a professora dar a prova surpresa. Esse conhecimento prévio está disponível no momento da declaração.

Essa solução implica que pode haver desacordo entre raciocinadores ideais que concordam com os mesmos dados impessoais. Considere os colegas dos professores. Eles não estão entre aqueles que a professora almeja para a surpresa. Como “surpresa” aqui significa “surpresa para os alunos”, os colegas da professora podem consistentemente inferir que a prova será no último dia a partir da premissa de que ela não foi dada em nenhum dia anterior. Mas esses colegas não são úteis *para os alunos* como informantes.

1.7 Paradoxos Epistêmicos Dinâmicos

As anomalias acima (perder conhecimento sem esquecê-lo, discordância entre raciocinadores ideais igualmente bem informados, racionalmente mudar de ideia sem a aquisição de contraevidências) seriam mais toleráveis se reforçadas por linhas separadas de raciocínio. A fonte mais fértil desse tipo de suporte paralelo está em enigmas sobre atualização de crenças.

A estratégia natural é focar no conhecedor estático. No entanto, assim como é mais fácil para um esquimó observar uma raposa do ártico quando ela se move, frequentemente obtemos uma melhor compreensão do conhecedor dinamicamente, quando ele está no processo de ganhar ou perder conhecimento.

1.7.1 Paradoxo da investigação de Mênon: Um enigma sobre a obtenção de conhecimento

Quando em julgamento por impiedade, Sócrates levou sua curiosidade até o Oráculo de Delfos (Apologia 21d em Cooper 1997). Antes de começar sua missão de investigação, Querefonte perguntou ao Oráculo: "Quem é o mais sábio dos homens?", e o Oráculo respondeu: "Ninguém é mais sábio do que Sócrates". Isso surpreendeu Sócrates, que acreditava não saber de nada. Enquanto um filósofo menos piedoso poderia ter questionado a confiabilidade do Oráculo de Delfos, Sócrates seguiu a prática costumeira de tratar o Oráculo como infalível. A única cogitação apropriada para uma resposta infalível é a interpretação. Consequentemente, Sócrates resolveu sua perplexidade inferindo que sua sabedoria estava em reconhecer sua própria ignorância. Enquanto outros podem não saber nada, Sócrates sabe que ele não sabe nada.

Ainda hoje Sócrates continua a ser elogiado por sua perspicácia, mas, na realidade, sua "descoberta" é uma contradição. Se Sócrates sabe que não sabe nada, então ele sabe algo (a proposição de que ele não sabe nada) e ainda assim não sabe nada (porque conhecimento implica verdade).

Sócrates poderia recuperar a consistência ao rebaixar seu metaconhecimento ao *status* de uma crença. Se ele acredita que não sabe nada, então ele naturalmente deseja remediar sua ignorância perguntando sobre tudo. Com efeito, esse raciocínio está presente ao longo dos primeiros diálogos. Mas quando chegamos ao *Mênon*, um de seus interlocutores tem uma epifania. Depois de receber o tratamento padrão oferecido por Sócrates no contexto de uma discussão acerca da natureza da virtude, Mênon observa um conflito entre a ignorância e a investigação socrática (Mênon 80d, em Cooper 1997). Como Sócrates reconheceria a resposta correta mesmo se Mênon a fornecesse?

A estrutura geral do paradoxo de Mênon é um dilema: Se você sabe a resposta para a pergunta que está fazendo, então nada pode ser aprendido perguntando; se você não sabe a resposta, então você não pode reconhecer uma resposta correta mesmo que ela seja dada a você. Portanto, não se pode aprender nada fazendo perguntas.

A solução natural para o paradoxo de Mênon é caracterizar o inquiridor como apenas parcialmente ignorante. Ele sabe o suficiente para reconhecer uma resposta correta, mas não o suficiente para responder sozinho. Dicionários ortográficos são inúteis para crianças de seis anos porque elas raramente sabem mais do que a primeira letra da palavra buscada. Crianças de dez anos têm conhecimento parcial suficiente da grafia da palavra para reduzir o número de possíveis grafias. Dicionários de ortografia também são inúteis para aqueles com conhecimento total de ortografia e aqueles com total ignorância de ortografia. Mas a maioria de nós tem uma quantidade intermediária de conhecimento.

É natural analisar conhecimento parcial como conhecimento de condicionais. A criança de dez anos conhece a versão falada de "Se o dicionário ortográfico soletra o mês após janeiro como *f-e-v-e-r-e-i-r-o*, então essa grafia está correta". Consultar o dicionário ortográfico lhe dá conhecimento do antecedente do condicional.

Grande parte do nosso aprendizado com condicionais ocorre tão suavemente quanto esse exemplo sugere. Como conhecemos o condicional, estamos preparados para aprender o consequente meramente aprendendo o antecedente (e aplicando a regra de inferência *modus ponens*: Se P então Q, P, portanto Q). Mas a próxima seção é dedicada a alguns condicionais conhecidos que são repudiados quando aprendemos seus antecedentes.

1.7.2 Paradoxo do dogmatismo: Um enigma sobre a perda de conhecimento

As ruminações de Saul Kripke sobre o paradoxo da prova surpresa o levaram a um paradoxo sobre o dogmatismo. Ele deu uma palestra sobre ambos os paradoxos na Universidade de Cambridge para o *Moral Sciences Club* em 1972. (Uma versão da palestra agora aparece como Kripke 2011.) Gilbert Harman apresentou o novo paradoxo de Kripke da seguinte forma:

> Se eu sei que *h* é verdadeira, eu sei que qualquer evidência contra *h* é uma evidência contra algo que é verdade; eu sei que tal evidência é enganosa. Mas eu deveria desconsiderar evidências que eu sei que são enganosas. Então, uma vez que eu saiba que *h* é verdadeira, estou em posição de desconsiderar qualquer evidência futura que pareça ir contra *h*. (1973, 148)

Os dogmáticos aceitam esse raciocínio. Para eles, o conhecimento encerra a investigação. Qualquer "evidência" que entre em conflito com o que já é conhecido pode ser descartada como evidência enganosa. Quem é prevenido vale por dois.

Esse conservadorismo cruza a linha que vai da confiança para a intransigência. Para ilustrar a inflexibilidade excessiva, aqui está um argumento em cadeia para a conclusão dogmática de que meu confiável colega Doug me deu um relato enganoso (corrigido de Sorensen 1988b):

(C_1) Meu carro está no estacionamento.

(C_2) Se meu carro estiver no estacionamento e Doug fornecer evidências de que meu carro não está no estacionamento, então as evidências de Doug são enganosas.

(C_3) Se Doug relata que viu um carro igual ao meu sendo rebocado do estacionamento, então seu relato é uma evidência enganosa.

(C_4) Doug relata que um carro igual ao meu foi rebocado do estacionamento.

(C_5) O relato de Doug é uma evidência enganosa.

Por hipótese, estou justificado em acreditar em (C_1). A premissa (C_2) é uma certeza porque é analiticamente verdadeira. O argumento que leva de (C_1) e (C_2) para (C_3) é válido. Portanto, meu grau de confiança em (C_3) deve ser igual ao meu grau de confiança em (C_1). Uma vez que também estamos assumindo que obtenho justificação suficiente para (C_4), parece seguir-se que estou justificado a acreditar em (C_5) por *modus ponens*. Argumentos semelhantes me levarão a rejeitar novas evidências, como um telefonema do serviço de reboque e minha falha em ver meu carro quando caminho confiantemente até o estacionamento.

Gilbert Harman diagnostica o paradoxo da seguinte forma:

> O argumento para o paradoxo ignora a maneira como realmente ter evidências pode fazer a diferença. Como agora sei [que meu carro está no estacionamento], agora sei que qualquer evidência que pareça indicar outra coisa é enganosa. Isso não me autoriza a simplesmente desconsiderar qualquer evidência adicional, pois obter essa evidência adicional pode mudar o que eu sei. Em particular, depois de obter essa evidência adicional, posso não saber mais que ela é enganosa. Pois ter a nova evidência pode tornar verdade que não sei mais que a nova evidência é enganosa. (1973, 149)

Na verdade, Harman nega a robustez do conhecimento. O princípio da robustez afirma que alguém só detém conhecimento sobre algo se não houver nenhuma evidência tal que, caso a pessoa soubesse dessa evidência, ela não estaria justificada em acreditar em sua conclusão.

A conclusão de Harman de que novos conhecimentos podem minar conhecimentos antigos pode ser aplicada ao paradoxo da prova surpresa: os alunos perdem o conhecimento do anúncio da prova, embora não se esqueçam do anúncio ou façam qualquer outra coisa incompatível com suas credenciais como raciocinadores ideais. Um aluno na quinta-feira está mais bem informado sobre os resultados dos dias da prova do que estava no domingo. Ele sabe que a prova não foi na segunda-feira e nem na quarta-feira. Mas ele só pode prever que o teste será na sexta-feira se continuar a saber o anúncio. O conhecimento extra dos dias sem prova mina o conhecimento do anúncio.

A maioria dos epistemólogos aceitou o apelo de Harman aos derrotadores. Alguns tentaram torná-lo mais preciso com detalhes sobre a atualização de condicionais indicativos (Sorensen 1988b). Isso pode justificar e generalizar a previsão de Harman de que as evidências futuras mudarão sua opinião sobre o que são evidências enganosas. O conhecimento de tais condicionais é inútil para um *modus ponens* futuro. O dogmático diz corretamente que conhecemos o condicional "Se eu sei que *p*, então qualquer evidência conflitando com *p* é uma evidência enganosa". De fato, é uma tautologia! Mas o dogmático falha em reconhecer que essa tautologia conhecida é um conhecimento inútil. Adquirir a evidência enganosa me fará deixar de conhecer *p*. Se um auditor prevê ser apresentado a uma lista tendenciosa de fatos, ele pode proferir a tautologia ao seu assistente para transmitir outra proposição para a qual ele tem suporte empírico. Essa proposição empírica não precisa ser conhecimento inútil. Quando a lista prevista é apresentada, o auditor prevenido ignora os fatos. Mas a base não é seu conhecimento *a priori* da tautologia do dogmático.

Kripke observa que essa solução não impedirá o dogmático de raciocínio rápido que toma medidas para impedir a aquisição de evidências que ele agora considera enganosas (Kripke 2011, 43–44). Uma segunda preocupação é que o dogmático ainda pode ignorar evidências fracas. Se eu sei que uma moeda é justa, então sei que, se os primeiros vinte lançamentos derem cara, essa é uma evidência enganosa de que a moeda não é justa. Tal sequência não derrota minha alegação de conhecimento. (Substitua por uma sequência mais curta se você acha que ela derrota a alegação de conhecimento). Sendo assim, a solução de Harman não se aplica. No entanto, é dogmatismo ignorar essa evidência.

Além desse problema do dogmatismo fraco, Rachel Fraser (2022) acrescenta um terceiro problema: o da ancoragem[5] dogmática. Quando Robert Millikan e Harvey Fletcher mediram a carga elétrica elementar com gotículas eletricamente carregadas, eles descartaram algumas gotas por estarem enganadoramente fora do intervalo plausível para o valor verdadeiro. Gotas localizadas mais centralmente dentro do intervalo eram consideradas "belas". A exclusão dos valores discrepantes proporcionou a Millikan uma medição mais precisa — e um Prêmio Nobel em 1923. Em 1978, o físico Gerald Holton examinou os cadernos de anotações e ficou chocado com a quantidade de dados contrários que haviam sido omitidos por Millikan. Fraser considera que há uma circularidade viciosa na purificação dos dados.

Mas a ancoragem dogmática considerará a circularidade como virtuosa. Quando as evidências são uma mescla de evidências fortes e contra evidências fracas, o corpo de evidências mais forte expõe o corpo mais fraco como sendo de evidências enganosas. Pense em um quebra-cabeça que foi misturado com peças perdidas de outro quebra-cabeça. Quando você consegue obter uma imagem completa com um subconjunto das peças, as peças restantes são jogadas fora. Escurecer as evidências enganosas permite com que as evidências principais brilhem mais visivelmente. Portanto, podemos de fato estar mais confiantes do que estávamos antes de descartar as evidências fracas. Millikan estava sendo um guardião responsável em vez de um esperançoso ingênuo. Assim como os dados devem controlar a teoria, a teoria deve controlar os dados. O experimentador deve atingir um equilíbrio delicado entre descartar dados contrários em excesso e descartá-los menos que o devido. As soluções propostas para o paradoxo do dogmatismo têm dificuldade em sustentar esse equilíbrio.

I. J. Good (1967) demonstrou que, quando o custo das evidências é insignificante, a obtenção de novas evidências maximiza o valor epistêmico esperado. Sob essa suposi-

[5]N.T. "Bootstrapping" no original.

ção simplificadora, Good mostra que desvios do princípio da evidência total são, no mínimo, imprudentes. Dado o utilitarismo epistêmico, essa irracionalidade prática se torna irracionalidade teórica. Bob Beddor (2019) agora adiciona a premissa de que é irracional pretender fazer o que se prevê ser irracional. Por exemplo, se fosse oferecido um milhão de dólares para beber uma toxina amanhã que o deixaria doente por um dia, você poderia lucrar com a oferta (Kavka 1983). Contudo, se o milhão fosse ganho imediatamente após você ter a intenção de beber a toxina, então você não poderia lucrar porque você sabe que não haveria razão para seguir adiante. Por analogia, Beddor conclui que seria irracional pretender evitar contra evidências (Beddor 2019, 738). Ninguém nunca tem o direito de descartar evidências, mesmo depois de terem sido previstas como evidências enganosas.

Mas se o custo da evidência for significativo, a conexão entre racionalidade prática e racionalidade teórica *favorece* ignorar evidências contrárias. Julgar que *p* incorpora uma resolução de não investigar mais a questão de se *p* é verdadeiro. Ou ao menos assim responde o volicionista (Fraser 2022).

1.7.3 O futuro dos paradoxos epistêmicos

Não podemos prever que algum novo paradoxo epistêmico específico aguarda para ser descoberto. Para ver o porquê, considere a previsão que Jon Wynne-Tyson atribui a Leonardo Da Vinci:

> Aprendi desde cedo a abjurar o uso da carne, e chegará o tempo em que homens como eu olharão para o assassinato de animais como agora olham para o assassinato de homens. (1985, 65)

Ao prever esse progresso, Leonardo inadvertidamente revela que ele *já acredita* que o assassinato de animais é o mesmo que o assassinato de homens. Se você acredita que uma proposição é verdadeira, mas também que só se acreditará nela num momento futuro, então você já a acredita agora — e, portanto, há uma inconsistência (o que é realmente verdade é irrelevante).

Regressos específicos podem ser antecipados. Quando tento prever minha crença futura em uma verdade específica, eu me antecipo. Quando tento prever minha crença futura em uma falsidade específica, não há preempção.

Não haveria problema em prever o progresso se Leonardo pensasse que o progresso moral reside na preferibilidade moral da crença vegetariana em vez da verdade das crenças vegetarianas. Alguém pode admirar o vegetarianismo sem aceitar a correção do vegetarianismo. Mas Leonardo está endossando a correção da crença. Essa frase incorpora um absurdo mooreano. É como dizer: "Leonardo levou vinte e cinco anos para completar *A Virgem das Rochas*, mas só acreditarei nisso a partir de amanhã". (Este absurdo levará alguns a objetar que interpretei Leonardo de forma pouco caridosa; ele deve ter pretendido abrir uma exceção para si mesmo e estar se referindo apenas a homens de seu tipo.)

Não posso antecipar especificamente a aquisição inicial da crença verdadeira de que *p*, pois essa previsão mostraria que eu já tenho a crença verdadeira de que *p*. A verdade não espera. A impaciência da verdade impõe um limite à previsão de descobertas.

Referências Bibliográficas

K. Scott Aikin. *Epistemology and the Regress Problem*. Routledge, London, 2011.

C. Anthony Anderson. The paradox of the knower. *The Journal of Philosophy*, 80:338–355, 1983.

Robert Binkley. The surprise examination in modal logic. *Journal of Philosophy*, 65(2):127–136, 1968.

Nicolas Bommarito. Rationally self-ascribed anti-expertise. *Philosophical Studies*, 151:413–419, 2010.

Luc Bovens. 'P and I will believe that not-P': Diachronic Constraints on Rational Belief. *Mind*, 104(416):737–760, 1995.

Tyler Burge. Buridan and epistemic paradox. *Philosophical Studies*, 34:21–35, 1978.

Tyler Burge. Epistemic paradox. *Journal of Philosophy*, 81(1):5–29, 1984.

John Buridan. *John Buridan on Self-Reference: Chapter Eight of Buridan's 'Sophismata'*. Cambridge University Press, Cambridge, 1982.

Rudolf Carnap. *The Logical Foundations of Probability*. University of Chicago Press, Chicago, 1950.

David Christensen. Higher order evidence. *Philosophy and Phenomenological Research*, 81:185–215, 2010.

Cicero. *On the Nature of the Gods, Academica*. Loeb Classical Library, Cambridge, MA, 1933.

Arthur Collins. Could our beliefs be representations in our brains? *Journal of Philosophy*, 74 (5):225–243, 1979.

Earl Conee. Heeding misleading evidence. *Philosophical Studies*, 103:99–120, 2004.

John Cooper, editor. *Plato: The Complete Works*. Hackett, Indianapolis, 1997.

Keith DeRose. *The Appearance of Ignorance: Knowledge, Skepticism, and Context (Volume 2)*. Oxford University Press, Oxford, 2017.

Andy Egan and Adam Elga. I can't believe i'm stupid. *Philosophical Perspectives*, 19(1): 77–93, 2005.

Sextus Empiricus. *Outlines of Pyrrhonism*. Harvard University Press, Cambridge, MA, 1933.

Paul Feyerabend. *Against Method*. Verso, London, 1988.

Frederic Fitch. A logical analysis of some value concepts. *Journal of Symbolic Logic*, 28(2): 135–142, 1963.

Rachel Fraser. The will in belief. *Oxford Studies in Epistemology*, 2022.

I. J. Good. On the principle of total evidence. *British Journal for the Philosophy of Science*, 17(4):319–321, 1967.

Kurt Gödel. What is cantor's continuum problem? In Paul Benacerraf and Hilary Putnam, editors, *Philosophy of Mathematics*, pages 258–273. Cambridge University Press, Cambridge, 1983.

Ian Hacking. *The Emergence of Probability*. Cambridge University Press, Cambridge, 1975.

Alan Hajek. The cable guy paradox. *Analysis*, 65(2):112–119, 2005.

Gilbert Harman. Knowledge, inference, and explanation. *American Philosophical Quarterly*, 5(3):164–173, 1968.

Gilbert Harman. *Thought*. Princeton University Press, Princeton, 1973.

John Hawthorne. *Knowledge and Lotteries*. Clarendon Press, Oxford, 2004.

Piet Hein. *Grooks*. MIT Press, Cambridge, MA, 1966.

Jaakko Hintikka. *Knowledge and Belief*. Cornell University Press, Ithaca, NY, 1962.

Wesley Holliday. On being in an undiscoverable position. *Thought*, 5(1):33–40, 2016.

Wesley Holliday. Epistemic logic and epistemology. In Sven Ove Hansson and Vincent F. Hendricks, editors, *The Handbook of Formal Philosophy*. Springer, Dordercht, 2017.

G. E. Hughes. *John Buridan on Self-Reference*. Cambridge University Press, Cambridge, 1982.

Daniel Immerman. Question closure to solve the surprise test paradox. *Synthese*, 194(11): 4583–4596, 2017.

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

Peter Klein. How to be an infinitist about doxastic justification. *Philosophical Studies*, 134: 77–25–29, 2007.

Kevin Knight. Measuring inconsistency. *Journal of Philosophical Logic*, 31(1):77–98, 2002.

Saul Kripke. Two paradoxes of knowledge. In *Philosophical Troubles: Collected Papers (Volume 1)*, pages 27–51. Oxford University Press, New York, 2011.

Jonathan L. Kvanvig. The epistemic paradoxes. In *Routledge Encyclopedia of Philosophy*. Routledge, London, 1998.

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

David Lewis. Lucas against mechanism. In *Papers in Philosophical Logic*, pages 166–169. Cambridge University Press, Cambridge, 1998.

David Lewis and Jane Richardson. Scriven on human unpredictability. *Philosophical Studies*, 17(5):69–74, 1966.

J. R. Lucas. Minds, machines and gödel. In Alan Ross Anderson, editor, *Minds and Machines*, pages 112–117. Prentice Hall, Englewood Cliffs, N.J., 1964.

D. C. Makinson. The paradox of the preface. *Analysis*, 25:205–207, 1965.

Norman Malcolm. *Knowledge and Certainty*. Prentice Hall, Englewood Cliffs, NJ, 1963.

G. E. Moore. A reply to my critics. In P. A. Schilpp, editor, *The Philosophy of G. E. Moore*. Northwestern University, Evanston, IL, 1942.

G. C. Nerlich. Unexpected examinations and unprovable statements. *Mind*, 70(280):503–514, 1961.

Charles Sanders Peirce. *The Collected Works of Charles Sanders Peirce*. Harvard University Press, Cambridge, MA, 1931–1935.

Plato. *Plato: The Complete Works*. Hackett, Indianapolis, 1997.

Gontran de Poncins. *Kabloona in collaboration with Lewis Galantiere*. Carroll & Graff Publishers, New York, 1941. 1988 reprint.

John F. Post. The possible liar. *Noûs*, 4:405–409, 1970.

W. V. O Quine. On a so-called paradox. *Mind*, 62(245):65–67, 1953.

W. V. O Quine. Epistemology naturalized. In *Ontological Relativity and Other Essays*, pages 69–90. Columbia University Press, New York, 1969.

W. V. O Quine. *Quiddities*. Harvard University Press, Cambridge, MA, 1987.

Stephen Read. Self-reference and validity. *Synthese*, 42(2):265–274, 1979.

R. M. Sainsbury. *Paradoxes*. Cambridge University Press, Cambridge, 1995.

Joseph Salerno. *New Essays on the Knowability Paradox*. Oxford University Press, New York, 2009.

Michael Scriven. An essential unpredictability in human behavior. In Benjamin B. Wolman and Ernest Nagel, editors, *Scientific Psychology: Principles and Approaches*, pages 411–425. Basic Books, New York, 1964.

Brian Skyrms. Causal decision theory. *Journal of Philosophy*, 79(11):695–711, 1982.

Martin Smith. *Between Probability and Certainty*. Clarendon Press, Oxford, 2016.

Roy Sorensen. *Blindspots*. Clarendon Press, Oxford, 1988a.

Roy Sorensen. Dogmatism, junk knowledge, and conditionals. *Philosophical Quarterly*, 38: 433–454, 1988b.

Roy Sorensen. *Vagueness and Contradiction*. Clarendon Press, Oxford, 2001.

Roy Sorensen. Paradoxes of rationality. In Al Mele, editor, *The Handbook of Rationality*, pages 257–275. Oxford University Press, Oxford, 2003a.

Roy Sorensen. *A Brief History of the Paradox*. Oxford University Press, New York, 2003b.

Andrew Stephenson. Kant, the paradox of knowability, and the meaning of experience. *Philosophers Imprint*, 15(17):1–19, 2015.

J. F. Thomson. On some paradoxes. In R. J. Butler, editor, *Analytical Philosophy*, pages 104–119. Barnes & Noble, New York, 1962.

Thomas Tymoczko. An unsolved puzzle about knowledge. *The Philosophical Quarterly*, 34: 437–458, 1984.

Bas van Fraassen. Belief and the will. *Journal of Philosophy*, 81:235–256, 1984.

Bas van Fraassen. Belief and the problem of ulysses and the sirens. *Philosophical Studies*, 77:7–37, 1995.

Paul Weiss. The prediction paradox. *Mind*, 61(242):265–269, 1952.

Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.

Jon Wynne-Tyson. *The Extended Circle*. Centaur Press, Fontwell, Sussex, 1985.

(IV) Paradoxo de Zenão[1]

Título Original: Zeno's Paradox
Autor: Nick Huggett
Tradução: Rafael Cavalcanti de Souza
Revisão: Maria Amélia Reis de Castro Rodrigues

Quase tudo o que sabemos sobre Zenão de Eleia está nas páginas iniciais do *Parmênides* de Platão. No diálogo, aprendemos que Zenão tinha quase 40 anos quando Sócrates era um jovem, por volta de 20 anos. Como Sócrates nasceu em 469 a.C., podemos estimar que Zenão tenha nascido por volta de 490 a.C. Além disso, o que realmente sabemos é que ele estava próximo de Parmênides (Platão relata o boato de que eles tiveram um relacionamento sexual quando Zenão era jovem) e que escreveu um livro de paradoxos defendendo a filosofia de Parmênides. Infelizmente, esse livro não sobreviveu ao passar do tempo, e o que sabemos de seus argumentos é por terceiros, principalmente por meio de Aristóteles e seus comentadores (aqui nos baseamos especialmente em Simplício, que, embora tenha escrito mil anos após Zenão, aparentemente possuía ao menos alguns de seus livros). Aparentemente, havia 40 'paradoxos da pluralidade', que tentavam mostrar que o pluralismo ontológico – a crença na existência de várias coisas em vez de apenas uma – leva a conclusões absurdas; desses paradoxos, apenas dois definitivamente sobreviveram, embora um terceiro argumento possa ser atribuído a Zenão. Aristóteles fala de mais quatro

---

[1]HUGGETT, Nick, "Zeno's Paradoxes", 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/paradox-zeno/.
A seguir está a tradução da entrada sobre os Paradoxos de Zenão de Nick Huggett na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/fall2024/entries/paradox-zeno/. 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/paradox-zeno/. Agradecemos aos editores Edward N. Zalta e Uri Nodelman pela permissão para traduzir e publicar esta entrada.

argumentos contra o movimento (e por extensão a mudança em geral), todos os quais ele apresenta e tenta refutar. Aristóteles atribui dois outros paradoxos a Zenão. Infelizmente, mais uma vez, quase nenhum desses paradoxos é citado nas palavras originais de Zenão, sendo reproduzido em paráfrases por vários comentadores.

1. Contexto

1.1 Contexto Antigo

Antes de nós olharmos para os paradoxos mesmo será útil esboçar algumas das suas significâncias histórica e lógica. Primeiro, Zenão procurou defender Parmênides atacando os seus críticos. Parmênides rejeitou o pluralismo e a realidade de qualquer tipo de mudança: para ele, tudo era um indivisível, realidade imutável, e qualquer aparências do contrário eram ilusões, para serem dissolvida pela razão e relelação. Não surpreendentemente, essa filosofia encontrou muitos críticos, que ridicularizaram a sugestão; afinal, ela contradiz algumas de nossas crenças mais básicas sobre o mundo. Em resposta a essas críticas, Zenão fez algo que pode parecer óbvio, mas que teve um impacto profundo na filosofia grega sentido até os dias de hoje: ele tentou demonstrar que absurdos igualmente graves decorriam logicamente da negação das ideias de Parmênides. Você acredita que existem muitas coisas? Você precisa concluir, então, que tudo é ao mesmo tempo infinitamente pequeno e infinitamente grande! Você acredita que o movimento é infinitamente divisível? Então, disso decorre que nada se move! (É isso que se entende por 'paradoxo': uma demonstração de que uma contradição ou consequência absurda decorre de premissas aparentemente razoáveis.)

Ao analisarmos os argumentos, é crucial manter esse método em mente. Eles são sempre direcionados a um alvo mais ou menos específico: as visões de alguma pessoa ou escola. Devemos lembrar que tais argumentos são 'ad hominem' no sentido literal do latim – dirigidos às posições ou crenças das pessoas – e não no sentido de atacar as pessoas em si que apresentam essas visões, em vez de atacar as próprias visões. Esses argumentos operam assumindo temporariamente, pro argumento, que as afirmações são verdadeiras, e depois argumentam que, se forem verdadeiras, consequências absurdas se seguem— como, por exemplo, que nada se move. Esses são argumentos de '*reductio ad absurdum*' (ou 'dialética' no sentido da época). Assim, se o argumento for logicamente válido e a conclusão for genuinamente inaceitável, então as afirmações devem ser falsas. Ao analisarmos os argumentos de Zenão, portanto, devemos fazer duas perguntas relacionadas: quem ou qual posição Zenão está atacando, e o que exatamente é assumido para que o argumento possa ser feito? Se descobrirmos que Zenão faz suposições ocultas além do que a posição atacada implica, então a conclusão absurda pode ser evitada ao negar uma dessas suposições ocultas, enquanto se mantém a posição. De fato, comentaristas, pelo menos desde Aristóteles, têm respondido a Zenão dessa maneira.

Então, de quem são as visões que os argumentos de Zenão atacam? Há uma vasta literatura debatendo o alvo histórico exato de Zenão. Como discutiremos brevemente adiante, alguns argumentam que o alvo era uma doutrina técnica dos pitagóricos, mas a maioria dos estudiosos hoje interpreta Zenão como se opondo às noções sobre pluralidade e movimento do senso comum. Abordaremos os paradoxos nesse espírito e direcionaremos o leitor à literatura que trata do debate interpretativo.

1.2 Contexto Moderno

Seguindo uma indicação dada por Russell (1929, pp. 182–198), diversos filósofos – notavelmente Grünbaum (1967), mas também Salmon (2001) – empreenderam a tarefa de demonstrar como a matemática moderna poderia resolver todos os paradoxos de Zenão (ver §5 para uma importante ampliação desse projeto). Esse trabalho permanece como a visão "recebida"e influencia profundamente a discussão a seguir, tanto na abordagem quanto nas análises específicas.

Essa abordagem sustenta que – com certas qualificações – os paradoxos de Zenão revelam problemas que não podem ser resolvidos sem os recursos completos da matemática desenvolvida no século XIX (e possivelmente além disso). Isso não significa (necessariamente), contudo, que a matemática moderna seja indispensável para responder aos problemas que Zenão *pretendia* levantar; é possível argumentar que Aristóteles e outros antigos já ofereceram respostas que poderiam – ou deveriam – ter satisfeito Zenão. (Além disso, não faremos nenhuma alegação específica sobre a influência de Zenão na história da matemática.) No entanto, à medida que a matemática se desenvolveu e os paradoxos foram mais analisados, surgiram novas dificuldades a partir deles, dificuldades essas que exigem a matemática moderna para serem resolvidas. Essas novas dificuldades derivam, em parte, da evolução da nossa compreensão sobre o que o rigor matemático exige: soluções que atenderiam aos padrões de rigor de Zenão não satisfariam os nossos. Assim, abordaremos vários dos paradoxos, desde suas formulações de senso comum até suas resoluções na matemática moderna. (Uma qualificação: apresentaremos resoluções em termos da matemática “padrão”, mas outras formulações modernas também são capazes de lidar com Zenão, e, possivelmente, de modos que melhor representam seus conceitos matemáticos.)

Embora muitos continuem a considerar frutífera a abordagem geral de Russell e Grünbaum, mais recentemente, diversos comentaristas têm questionado as interpretações específicas da visão predominante. Caso estejam corretas, essas críticas enriquecem nossa compreensão tanto da matemática antiga quanto da matemática do século XIX, como veremos a seguir.

2. Os Paradoxos da Pluralidade

2.1 O Argumento da Densidade

> Se há muitos, devem ser tantos quantos são, nem mais nem menos do que isso. Mas, se são tantos quantos são, seriam limitados. Se há muitos, as coisas que existem são ilimitadas. Pois há sempre outros entre as coisas que existem, e ainda outros entre estes, e assim as coisas que existem são ilimitadas. (Simplício, *Sobre a Física de Aristóteles*, 140.29)

Esse primeiro argumento, apresentado nas palavras de Zenão segundo Simplício, tenta demonstrar que não poderia haver mais de uma coisa, sob pena de contradição: se há muitas coisas, então elas são ao mesmo tempo “limitadas”e “ilimitadas”, o que é uma contradição. Por um lado, ele afirma que qualquer conjunto deve conter um número *definido* de coisas, ou, em suas palavras, “nem mais nem menos”. Mas, se há um número definido de coisas, conclui ele, então esse número deve ser finito – ou seja, ‘limitado’. Ao inferir isso, Zenão pressupõe que ter infinitas coisas seria o mesmo que ter um número indefinido delas. Por outro lado, imagine qualquer coleção de “muitas”coisas dispostas no espaço – para maior clareza, visualize-as alinhadas em uma dimensão. Entre quaisquer duas delas, ele afirma, há uma terceira; e entre essas três, outras duas; e entre essas cinco, outras quatro; e assim por diante, sem fim. Portanto, a coleção também é ‘ilimitada’. Assim, nossa suposição inicial de uma pluralidade leva a uma contradição e, consequentemente, é falsa: no fim das contas, não há muitas coisas. Pelo menos, é assim que o raciocínio de Zenão se desenvolve.

Consideremos os dois argumentos subordinados, em ordem inversa. Primeiro: haveria “sempre outros entre as coisas que existem”? (Em terminologia moderna: por que os objetos deveriam estar sempre ordenados de forma ”densa”?) Suponha que imaginássemos uma fileira de dez maçãs; de fato, há uma maçã entre a sexta e a oitava, mas não há uma entre a sétima e a oitava! Se assumirmos que Zenão não está simplesmente confuso, o que ele tinha em mente? Os textos não dizem, mas há duas possibilidades: primeiro, pode-se defender que, para qualquer par de objetos físicos (digamos, duas maçãs) serem dois objetos distintos e não apenas um (uma 'maçã dupla'), deve haver um terceiro entre eles, separando-os fisicamente – mesmo que seja apenas ar. E pode-se pensar que, para esses três serem distintos, deve haver mais dois objetos separando-os, e assim por diante (essa visão pressupõe que o fato de serem feitos de substâncias diferentes não basta para torná-los distintos). Assim, talvez Zenão esteja argumentando contra a pluralidade com base em uma certa concepção de distinção física. Mas, em segundo lugar, também se poderia sustentar que qualquer corpo tem *partes* que podem ser ordenadas de forma densa. Claro, metades, quartos, oitavos etc. de maçãs não são densos – tais partes podem ser adjacentes —, mas pode haver partes suficientemente pequenas – digamos, 'partes-ponto' – que o sejam. De fato, se entre quaisquer duas partes-ponto houver uma distância finita, e se essas partes-ponto puderem estar arbitrariamente próximas, então elas são densas; uma terceira estaria no ponto médio entre quaisquer duas. Em particular, os pontos geométricos familiares são assim e, portanto, são densos. Logo, talvez Zenão esteja apresentando um argumento sobre a divisibilidade dos corpos. De qualquer forma, a suposição de densidade feita por Zenão exige alguma suposição adicional sobre a pluralidade em questão e, consequentemente, direciona o alvo de seu paradoxo.

Mas suponha que alguém sustente que certo conjunto (como os pontos em uma reta) é densa e, portanto, 'ilimitada' ou infinita. O primeiro aspecto do ataque de Zenão pretende mostrar que, por conter um número definido de elementos, ela também é 'limitada', ou finita. Seria possível escapar dessa contradição? A suposição de que qualquer número definido é finito parece intuitiva, mas hoje sabemos, graças ao trabalho de Cantor no século XIX, como entender números infinitos de modo que eles sejam tão definidos quanto os números finitos. O elemento central dessa teoria dos 'números transfinitos' é uma definição precisa de quando conjuntos infinitos têm o mesmo tamanho e quando uma é maior que a outra. Com tal definição, é possível ordenar os números infinitos assim como os finitos: por exemplo, há diferentes números infinitos definidos para frações e para pontos geométricos em uma reta, mesmo que ambos sejam densos. (Veja as Leituras Adicionais abaixo para referências sobre essas ideias matemáticas e sua história.) Portanto, ao contrário da suposição de Zenão, é significativo comparar conjuntos infinitos quanto ao número de seus elementos, afirmando se uma tem mais, menos ou 'tantos quanto' outra: há, por exemplo, mais números decimais do que números inteiros, mas há tantos números pares quanto números inteiros. Assim, matematicamente, o raciocínio de Zenão é incorreto quando ele afirma que, por uma coleção ter um número definido, ela deve ser finita, tornando o primeiro subargumento falacioso. (Embora, é claro, isso apenas mostre que coleções infinitas são matematicamente consistentes, não que existam fisicamente.)

2.2 O Argumento da Extensão Finita

> ... se fosse adicionado a algo já existente, não o tornaria maior. Pois, se não tivesse tamanho e fosse adicionado, não poderia aumentar de tamanho. E, portanto, segue-se imediatamente que o que é adicionado é nada. Mas se, quando subtraído, a outra coisa não é menor, nem aumenta quando é adicionada, claramente a coisa adicionada ou subtraída é nada. (Simplício *Sobre a Física de Aristóteles*, 139.9) Mas, se existe, cada coisa deve ter algum tamanho e espessura, e parte dela deve estar separada do resto. E o mesmo raciocínio se aplica à parte que está na frente. Pois esta também terá tamanho e parte dela estará na frente. Ora, é a mesma coisa dizer isso uma vez e continuar dizendo para sempre. Pois nenhuma parte dela será a última, nem haverá uma parte que não esteja relacionada a outra. Portanto, se há muitas coisas, elas devem ser pequenas e grandes; tão pequenas que não tenham tamanho, mas tão grandes que sejam ilimitadas. (Simplício *Sobre a Física de Aristóteles*, 141.2)

Mais uma vez temos as próprias palavras de Zenão. Segundo sua conclusão, há três partes neste argumento, mas apenas duas sobrevivem. O primeiro argumento – ausente – pretende mostrar que, se muitas coisas existem, elas não devem ter tamanho algum. Em segundo lugar, a partir disso Zenão argumenta que segue-se que elas não existem de forma alguma; como o resultado de juntar (ou remover) um objeto sem tamanho a qualquer coisa não produz mudança alguma, ele conclui que a coisa adicionada (ou removida) é literalmente nada. O argumento até este ponto é uma refutação autossuficiente do pluralismo, mas Zenão continua gerando um problema adicional para quem insiste na existência de uma pluralidade. Essa terceira parte do argumento é formulada de maneira bastante inadequada, mas parece ser algo assim: suponha que há uma pluralidade, então algum objeto espacialmente extenso existe (afinal, ele acabou de argumentar que coisas não extenso não existem). Como é estendido, tem duas partes espacialmente distintas (uma 'à frente' da outra). E como as partes existem, elas têm extensão, e portanto cada uma também tem duas partes espacialmente distintas; e assim por diante sem fim. E portanto, a conclusão final do argumento parece ser que o objeto, se é extenso de fato, é infinito em extensão.

Mas o que poderia justificar esse último passo? Não parece que pelo fato de um objeto ter duas partes ele deva ser infinitamente grande! E tampouco segue-se de qualquer outra das divisões que Zenão descreve aqui; quatro, oito, dezesseis, ou qualquer número finito de partes formam um todo finito. Novamente, certamente Zenão está ciente desses fatos, e portanto deve ter algo mais em mente, presumivelmente o seguinte: ele assume que se a série infinita de divisões que descreve fosse repetida infinitas vezes, então resultaria um conjunto definido de partes. E note que ele não precisa assumir que alguém poderia realmente realizar as divisões – não há tempo suficiente e as facas não são afiadas o suficiente - apenas que um objeto pode ser decomposto geometricamente em tais partes (ele também não assume que essas partes sejam o que naturalmente categorizaríamos como objetos físicos distintos como maçãs, células, moléculas, elétrons ou algo assim, mas apenas que são partes geométricas desses objetos). Agora, se – como um pluralista poderia muito bem aceitar – tais partes existem, segue-se da segunda parte de seu argumento que elas são extensas, e, ele aparentemente assume, uma soma infinita de partes finitas é infinita.

Aqui devemos notar que há duas maneiras pelas quais ele pode estar imaginando o resultado da divisão infinita.

Primeiro, poderíamos lê-lo como primeiro dividindo o objeto em metades, então uma das metades – digamos a segunda – em dois quartos, então um dos quartos – digamos o segundo novamente – em dois oitavos e assim por diante. Neste caso, o resultado da divisão infinita resulta em uma sequência interminável de pedaços de tamanho 1/2 do comprimento total, 1/4 do comprimento, 1/8 do comprimento... E então o comprimento total é $(1/2 + 1/4 + 1/8 + ...)$ do comprimento, o qual Zenão conclui ser uma distância infinita, de modo que o pluralista está comprometido com o absurdo de que corpos finitos são 'tão grandes quanto ser ilimitado'.

O que frequentemente é apontado em resposta é que Zenão não nos dá nenhuma razão para pensar que a soma é infinita em vez de finita. Ele pode ter tido a intuição de que qualquer soma infinita de quantidades finitas, já que cresce incessantemente com cada novo termo, deve ser infinita, mas também se poderia tomar esse tipo de exemplo como mostrando que algumas somas infinitas são, afinal, finitas. Assim, ao contrário do que ele pensava, Zenão não provou que a conclusão absurda se segue. No entanto, o que nem sempre é apreciado é que o pluralista não escapa tão facilmente, pois não basta apenas dizer que a soma *pode* ser finita, ela também deve mostrar que *é* finita – caso contrário permanecemos incertos sobre a sustentabilidade de sua posição. Como ilustração da dificuldade enfrentada aqui, considere o seguinte: muitos comentaristas falam como se fosse simplesmente óbvio que a soma infinita das frações é 1, que não há nada a somar infinitamente. Mas e a seguinte soma:$1 - 1 + 1 - 1 + 1 - ...$ Obviamente, parece, a soma pode ser reescrita como $(1-1)+(1-1)+... = 0+0+... = 0$. Certamente esta resposta parece tão intuitiva quanto a soma das frações. Mas esta soma também pode ser reescrita como $1 - (1 - 1 + 1 - 1 + ...) = 1 - 0$ – já que acabamos de mostrar que o termo entre parênteses desaparece – =1. Confiar em intuições sobre como realizar somas infinitas leva à conclusão de que $1 = 0$. Até que se possa dar uma teoria de somas infinitas que possa dar uma resposta satisfatória a qualquer problema, não se pode dizer que a soma infinita de Zenão é obviamente finita. Tal teoria só foi totalmente desenvolvida no século XIX por Cauchy. (No sistema de Cauchy,$1/2 + 1/4 + ... = 1$, mas $1 - 1 + 1 - ...$ é indefinido.)

Em segundo lugar, pode ser que Zenão queira dizer que o objeto é dividido ao meio, então ambas as metades são divididas ao meio, então todos os quartos são divididos ao meio e assim por diante. Nesse caso, os pedaços em qualquer estágio particular têm todos o mesmo tamanho finito, e assim poderia-se concluir que o resultado de continuar o procedimento infinitamente seria pedaços do mesmo tamanho, que, se existem - de acordo com Zenão - são maiores que zero; mas um infinito de partes extensas *iguais* é de fato infinitamente grande.

Mas essa linha de pensamento pode ser resistida. Primeiro, suponha que o procedimento descrito divida completamente o objeto em partes não sobrepostas. (Há um problema com essa suposição que veremos logo abaixo.) Envolve dobrar o número de pedaços após cada divisão e assim após N divisões há 2^N pedaços. Mas acontece que para qualquer número natural ou infinito, N, $2^N > N$, e assim o número de (supostas) partes obtidas pela infinidade de divisões descritas é um infinito ainda maior. Este resultado não apresenta dificuldade imediata, pois, como mencionamos acima, infinitos vêm em tamanhos diferentes. O número de vezes que tudo é dividido em dois é dito ser 'infinito contável': há um infinito contável de coisas em uma coleção se elas puderem ser rotuladas pelos números $1, 2, 3, ...$ sem sobra em nenhum dos lados. Mas o número de pedaços que a divisão infinita produz é

’infinito incontável’, o que significa que não há como rotulá-los $1, 2, 3, \ldots$ sem perder alguns deles - de fato, infinitos deles. No entanto, a definição de Cauchy de uma soma infinita só se aplica a séries infinitas contáveis de números, e portanto não se aplica aos pedaços que estamos considerando. No entanto, poderíamos considerar apenas um número contável deles, cujos comprimentos, de acordo com Zenão - já que ele afirma que são todos iguais e não nulos - somariam um comprimento infinito; o comprimento de todos os pedaços não poderia ser menor que isso.

Nesse ponto, o pluralista que acredita que a divisão de Zenão divide completamente os objetos em partes não sobrepostas (veja o parágrafo seguinte) poderia responder que as partes de fato não têm extensão, mesmo que existam. Isso bloquearia a conclusão de que objetos finitos são infinitos, mas parece empurrá-la de volta para o outro chifre do argumento de Zenão, pois como todas essas partes de comprimento zero podem formar um todo de tamanho não nulo? (Note que, de acordo com Cauchy, 0+0+0+...=0, mas este resultado não mostra nada aqui, pois, como vimos, há incontáveis pedaços para adicionar - mais do que são adicionados nesta soma.) Adiaremos essa questão para a discussão do próximo paradoxo, na qual ela surge explicitamente.

O segundo problema com interpretar a divisão infinita como uma divisão repetida de todas as partes é que ela não divide um objeto em partes distintas, se os objetos são compostos da maneira natural. Para ver isso, vamos fazer a pergunta sobre quais partes são obtidas por essa divisão em metades, quartos, oitavos, ... Como a divisão é repetida indeterminavelmente, não há um último pedaço que possamos dar como resposta, e assim precisamos pensar na questão de uma maneira diferente. Se supomos que um objeto pode ser representado por um segmento de linha de comprimento unitário, então a divisão produz conjuntos de segmentos, dos quais o primeiro é a primeira ou a segunda metade do segmento inteiro, o segundo é o primeiro ou segundo quarto, ou terceiro ou quarto quarto, e em geral o segmento produzido por *N* divisões é a primeira ou segunda metade do segmento anterior. Por exemplo, escrevendo o segmento com extremidades *a* e *b* como [a,b], alguns desses conjuntos (tecnicamente conhecidas como ’cadeias’, já que os elementos da coleção são ordenados por tamanho) começariam [0,1], [0,1/2], [1/4,1/2], [1/4,3/8], ... (Quando argumentamos antes que a divisão de Zenão produzia incontáveis pedaços do objeto, o que deveríamos ter dito com mais cuidado é que produz incontáveis cadeias como essa.)

A questão de quais partes a divisão seleciona é então a questão de qual parte qualquer cadeia dada seleciona; é natural dizer que uma cadeia seleciona a parte da linha que está contida *em cada um* de seus elementos. Considere, por exemplo, a cadeia [0,1/2], [1/4,1/2], [3/8,1/2], ..., em outras palavras, a cadeia que começa com a metade esquerda da linha e para a qual cada outro elemento é a metade direita do anterior. O ponto médio está em cada um dos segmentos dessa cadeia; é o ponto final direito de cada um. Mas nenhum outro ponto está em todos os seus elementos: claramente nenhum ponto além do meio está; e escolha qualquer ponto p antes do meio, se você tomar metades direitas de [0,1/2] vezes suficientes, o final esquerdo do segmento estará à direita de *p*. Por consequinte, a única parte da linha que está em todos os elementos desta cadeia é o ponto médio, e, portanto, esse é a parte da linha selecionada pela cadeia. (De fato, segue-se de um postulado da teoria dos números que há exatamente um ponto que todos os membros de qualquer cadeia desse tipo têm em comum.) O problema é que, por raciocínio paralelo, o ponto médio também é selecionado pela cadeia distinta [1/2,1], [1/2,3/4], [1/2,5/8], ..., onde cada segmento após o primeiro é a metade esquerda do precedente. E desse modo ambas as cadeias selecionam o mesmo pedaço da linha: o ponto médio. E assim por diante para muitos outros pares de cadeias. Assim, o argumento de Zenão, interpretado em termos de uma divisão repetida de *todas* as partes ao meio, não divide a linha em partes distintas. Portanto, se pensarmos que os objetos são compostos da mesma maneira que a linha, segue-se que, apesar das aparências, essa versão do argumento não corta os objetos em partes cujo tamanho total pode adequadamente ser discutido.

(Você pode pensar que esse problema poderia ser corrigido tomando os elementos das cadeias como segmentos sem ponto final à direita. Então, a primeira das duas cadeias que consideramos não tem mais o ponto médio em nenhum de seus segmentos, e, consequentemente, não seleciona aquele ponto. O problema agora é que ela falha em selecionar qualquer parte da linha: o raciocínio anterior mostrou que ela não seleciona nenhum ponto maior ou menor que o ponto médio, e agora nem mesmo seleciona aquele ponto!)

2.3 O Argumento da Divisibilidade Completa

> ... Sempre que um corpo for, por natureza, divisível completamente, seja por bisseção, seja, de modo geral, por qualquer outro método, nada impossível terá ocorrido se ele de fato tiver sido dividido – embora, talvez, ninguém na prática pudesse dividi-lo assim. O que, então, restará? Uma magnitude? Não: isso é impossível, pois então haveria algo não dividido, enquanto, ex hypothesi, o corpo era divisível completamente. Mas, se se admitir que não restará nem

> um corpo nem uma magnitude... o corpo consistirá ou de pontos (e seus constituintes serão destituídos de magnitude) ou será absolutamente nada. Se for este último caso, então poderia tanto vir a ser a partir do nada quanto existir como um composto de nada; e, assim, presumivelmente, o corpo todo não será nada além de uma aparência. Mas, se consistir de pontos, não possuirá nenhuma magnitude. (Aristóteles, Da Geração e da Corrupção, 316a19)

Estas palavras são de Aristóteles, não de Zenão, e, de fato, o argumento nem sequer é atribuído a Zenão por Aristóteles. No entanto, temos a opinião de Simplício (Sobre a Física de Aristóteles, 139.24) de que ele se origina com Zenão, razão pela qual está incluído aqui. Aristóteles começa supondo que algum corpo é completamente divisível, "através e através"; o segundo passo do argumento deixa claro que, com isso, ele quer dizer que o corpo é divisível em partes que, elas mesmas, não possuem tamanho – partes que possuem qualquer magnitude permanecem incompletamente divididas. (Mais uma vez, o que importa é que o corpo seja genuinamente composto de tais partes, não que alguém tenha tempo e ferramentas para realizar a divisão; e lembrando, da seção anterior, que não se obtêm tais partes por meio de divisões sucessivas de todas as partes pela metade.) Suponha-se, então, que o corpo seja dividido em suas partes sem dimensão. Essas partes poderiam ser ou nada – como Zenão argumentou acima – ou "partes-ponto". Se as partes forem nada, então o corpo também o será: trata-se apenas de uma ilusão.

E quanto ao passo final? Por que algo composto inteiramente de pontos não teria tamanho? Aristóteles não diz. Uma compreensão natural – e padrão – é que Aristóteles entende que o problema reside em somar os comprimentos de partes de comprimento zero. Certamente qualquer soma – mesmo uma infinita – de zeros é zero; isso certamente é o caso (como observado acima) na definição de somas infinitas de Cauchy. No entanto, Grünbaum (1967) apontou que essa definição se aplica apenas a somas enumeráveis, e Cantor forneceu uma bela, surpreendente e extremamente influente prova "diagonal" de que o número de pontos em um segmento é não enumeravelmente infinito. Não há como rotular *todos* os pontos de uma linha com a infinidade dos números 1, 2, 3, ..., de modo que há mais pontos em um segmento de linha do que somandos em uma soma de Cauchy. Em suma, a análise empregada para divisões infinitas enumeráveis não se aplica aqui.

Suponha, então, que sejam dados apenas o número de pontos em uma linha e que seus comprimentos sejam todos zero; como se determinaria o comprimento? Precisaríamos de uma nova definição, uma que estendesse a de Cauchy para somas não enumeráveis?

Acontece que isso não ajudaria, porque Cantor mostrou que qualquer segmento, de qualquer comprimento que seja (e, de fato, uma linha infinita inteira), *possui exatamente o mesmo número de pontos que nosso segmento unitário*. Assim, saber o número de pontos não determinará o comprimento da linha, e, portanto, nada semelhante à adição familiar – na qual o todo é determinado pelas partes – é possível. Em vez disso, devemos considerar as propriedades de distância de uma linha como logicamente posteriores à sua composição pontual: *primeiro* temos um conjunto de pontos (ordenados de certa maneira, de modo que haja algum fato, por exemplo, sobre qual dos três está entre os outros) e, *então*, definimos uma função dos pares de pontos que especifica quão distantes eles estão (satisfazendo condições como a de que a distância entre *A* e *B* somada à distância entre *B* e *C* é igual à distância entre *A* e *C* – se *B* está entre *A* e *C*). Assim, respondemos a Zenão da seguinte forma: o argumento pressupunha que o tamanho do corpo era a soma dos tamanhos das partes-ponto, mas esse não é o caso; segundo a matemática moderna, um segmento de linha geométrica é uma infinidade não enumerável de pontos mais uma função de distância. (Note-se que Grünbaum utilizou o fato de que a composição pontual não determina um comprimento para sustentar sua visão 'convencionalista' de que uma linha não possui comprimento determinado algum, independente de um padrão de medida.)

Como enfatiza Ehrlich (2014), poderíamos até mesmo estipular que uma "soma não enumerável" de zeros é zero, porque o comprimento de uma linha não é igual à soma dos comprimentos dos pontos que ela contém (respondendo à preocupação de Sherry (1988) de que recusar-se a estender a definição seria ad hoc). Assim, se se estipula que o comprimento de uma linha é a soma de qualquer conjunto completo de partes próprias, então decorre que os pontos não são, propriamente falando, partes de uma linha (ao contrário de metades, quartos e assim por diante de uma linha). Em um sentido estrito, na teoria da medida moderna (que generaliza a estrutura de Grünbaum), os pontos de uma linha são incomensuráveis com ela, e a própria formulação dada por Aristóteles, na qual o comprimento do todo é analisado em termos de seus pontos, é ilegítima.

A resposta de Grünbaum, embora aborde um paradoxo, foi recentemente significativamente contestada quanto à sua precisão, tanto no que se refere à matemática antiga quanto à matemática do século XIX. Reese et al. (2022) interpretam o paradoxo – especificamente o raciocínio de Aristóteles na frase final – de uma forma que não depende de somar zeros, mas pressupõe apenas que o tamanho de um ponto é limitado pelo tamanho das linhas das quais ele faz parte. O paradoxo resultante, eles demonstram, teria causado problemas para os princípios da matemática antiga; além disso, causou problemas na matemática da década de 1880, sendo, por fim, resolvido por Borel. Embora essa solução dependa do fato de que o número de pontos em uma linha é não enumerável, ela não se baseou na prova de Cantor nem em ideias sobre somas infinitas. Pesquisas futuras sérias sobre esse paradoxo precisarão levar esse trabalho em consideração.

3. Os Paradoxos do Movimento

3.1 A Dicotomia

> O primeiro [argumento] afirma a não existência do movimento com base na ideia de que aquilo que está em locomoção deve chegar à metade do caminho antes de chegar ao destino. (Aristóteles, *Física*, 239b11)

Esse paradoxo é conhecido como ‘dicotomia’ porque envolve uma divisão repetida em dois (assim como o segundo paradoxo da pluralidade). Como os outros paradoxos do movimento, ele nos foi transmitido por Aristóteles, que buscou refutá-lo.

Suponha que uma corredora muito rápida – como a mítica Atalanta – precise correr para pegar o ônibus. Claramente, antes de chegar ao ponto de ônibus, ela deve percorrer a metade do caminho, como diz Aristóteles. Não há problema nisso; supondo um movimento constante, ela levará metade do tempo para percorrer a metade do caminho, e a outra metade do tempo para percorrer o restante. Agora, ela também deve percorrer a metade da metade do caminho – ou seja, um quarto da distância total – antes de chegar à metade, mas novamente, ela tem um número finito de distâncias finitas a percorrer, e bastante tempo para isso. E antes de percorrer 1/4 do caminho, ela deve percorrer 1/2 de 1/4, que é 1/8 do caminho; e antes disso, 1/16; e assim por diante. Não há problema em nenhum ponto finito dessa série, mas e se essa divisão pela metade for feita infinitas vezes? A série resultante não contém uma primeira distância a ser percorrida, pois qualquer possível primeira distância poderia ser dividida ao meio e, portanto, não seria realmente a primeira. No entanto, ela contém uma última distância, que é 1/2 do caminho; uma penúltima, que é 1/4 do caminho; uma antepenúltima, que é 1/8 do caminho; e assim por diante. Assim, a sequência de distâncias que Atalanta precisa percorrer é: …, depois 1/16 do caminho, depois 1/8 do caminho, depois 1/4 do caminho e, finalmente, 1/2 do caminho (e aqui não estamos sugerindo que ela *pare* no fim de cada segmento e comece novamente no próximo – estamos pensando na sua corrida contínua, composta por essas partes). E agora surge um problema, pois essa descrição de sua corrida a faz percorrer um número *infinito* de distâncias *finitas*, o que, segundo Zenão, nos levaria à conclusão de que isso deveria levar um tempo infinito, ou seja, nunca seria concluído. E, como o argumento não depende da distância nem de quem ou do que seja o movimento, conclui-se que nenhuma distância finita pode jamais ser percorrida – ou seja, todo movimento é impossível. (Note que o paradoxo pode ser formulado facilmente na direção oposta, de modo que Atalanta deve primeiro percorrer metade do caminho, depois metade do restante, depois metade disso, e assim por diante, levando-a a uma sequência infinita de frações da distância total: 1/2, depois 1/4, depois 1/8, e assim sucessivamente.)

Algumas respostas comuns não são adequadas. Alguém poderia – como Simplicio (*Sobre a Física de Aristóteles*, 1012.22) relata que Diógenes, o Cínico, fez – simplesmente se levantar e começar a andar em silêncio, apontando que é uma questão de experiência comum que as coisas de fato se movem, e que sabemos muito bem que Atalanta não teria dificuldade em chegar ao ponto de ônibus. Mas isso não impressionaria Zenão, que, como um seguidor fiel de Parmênides, sustentava que muitas coisas não são como parecem: pode parecer que Diógenes está andando ou que Atalanta está correndo, mas as aparências podem enganar – e, afinal, temos uma prova lógica de que eles, na verdade, não estão se movendo. Alternativamente, se alguém não aceita que Zenão tenha realmente provado que o movimento é ilusório – como, esperançosamente, não aceitamos – então essa pessoa tem a obrigação de explicar o que há de errado em seu argumento: ele forneceu razões pelas quais o movimento é impossível, então uma resposta adequada precisa mostrar por que essas razões não são suficientes. E não basta simplesmente apontar que existem maneiras de dividir a corrida de Atalanta – por exemplo, apenas em duas metades – em que não há problema. Pois, se você aceita todos os passos do argumento de Zenão, então deve aceitar sua conclusão (assumindo que ele tenha raciocinado de forma logicamente dedutiva): não basta apresentar uma divisão não problemática, é necessário também mostrar por que *essa* divisão específica não apresenta problemas.

Outra resposta – dada pelo próprio Aristóteles – é apontar que, à medida que dividimos as distâncias percorridas, devemos também dividir o tempo total gasto: há 1/2 do tempo para percorrer a última metade, 1/4 do tempo para o quarto anterior, 1/8 do tempo para o oitavo de corrida, e assim por diante. Assim, cada fração da distância tem exatamente a fração correspondente do tempo total finito para que Atalanta a complete, e, portanto, a distância pode ser percorrida em um tempo finito. Essa resposta pode parecer trivial aos olhos modernos, mas, em um estudo importante, Sattler (2020) argumenta que entender o movimento por meio dessa correspondência matemática entre distância e duração foi um avanço crucial feito por Aristóteles, em resposta aos seus predecessores. Certamente, Aristóteles considerava que sua resposta deveria satisfazer Zenão. No entanto, ele também percebeu (*Física*, 263a15) que isso não encerrava completamente a questão. Afinal, agora estamos dizendo que o *tempo* que Atalanta leva para chegar ao ponto de ônibus é composto por um número infinito de partes finitas – ..., 1/8, 1/4 e 1/2 do tempo total – e isso não seria um tempo infinito?

É claro que alguém poderia novamente argumentar que algumas somas infinitas têm totais finitos e, em particular, que a soma dessas partes é exatamente 1× o tempo total, que, evidentemente, é finito (e, novamente, uma solução completa exigiria uma explicação rigorosa sobre somas infinitas, como a desenvolvida por Cauchy). No entanto, Aristóteles não fez esse passo. Em vez disso, ele traçou uma distinção rigorosa entre o que ele chamou de linha 'contínua' e uma linha dividida em partes. Considere uma divisão simples de uma linha em duas: de um lado, há a linha não dividida; de outro, a linha com um ponto médio selecionado como fronteira entre as duas metades. Aristóteles afirma que essas são duas coisas distintas e que a segunda é apenas 'potencialmente' derivada da primeira. Em seguida, Aristóteles adota a visão do senso comum de que o tempo é como uma linha geométrica e considera o tempo necessário para completar a corrida. Podemos novamente distinguir os dois casos: há o intervalo contínuo do início ao fim e há o intervalo dividido na infinidade de meias-corridas de Zenão. O primeiro é 'potencialmente infinito' no sentido de que poderia ser dividido no segundo, que é uma 'infinidade atual'. Aqui está o ponto crucial: Aristóteles acredita que, uma vez que esses intervalos são *geometricamente* distintos, eles também devem ser *fisicamente* distintos. Mas como isso seria possível? Ele afirma que a corredora deve fazer algo no fim de cada meia-corrida para torná-la distinta da próxima: ela deve parar, tornando a própria corrida descontínua. (Não está claro por que alguma outra ação não seria suficiente para dividir o intervalo.) Assim, a resposta completa de Aristóteles ao paradoxo é que a questão de saber se a série infinita de corridas é possível ou não é ambígua: a série potencialmente infinita de metades em uma corrida contínua é possível, enquanto uma infinidade atual de meias-corridas descontínuas não é – Zenão, de fato, identifica uma impossibilidade, mas ela não descreve a maneira usual de correr por uma pista!

É difícil – talvez do nosso ponto de vista moderno— ver como essa resposta poderia ser completamente satisfatória. Em primeiro lugar, ela assume que é possível traçar uma distinção clara entre infinitos potenciais e infinitos atuais, algo que nunca foi plenamente alcançado. Em segundo lugar, suponha que o problema de Zenão dependa da afirmação de que somas infinitas de quantidades finitas são invariavelmente infinitas. Então, a distinção de Aristóteles só ajudaria se ele pudesse explicar por que somas potencialmente infinitas são, de fato, finitas (não poderíamos, potencialmente, somar 1 + 1 + 1 + ..., o que não tem um total finito); ou se ele pudesse apresentar uma razão pela qual somas potencialmente infinitas simplesmente não existem. Ou talvez Aristóteles não visse as somas infinitas como o problema, mas sim saber se completar uma infinidade de ações finitas é metafisicamente, conceitualmente e fisicamente possível. Discutiremos brevemente essa questão – das 'supertarefas' – mais adiante, mas vale notar que existe uma corrida bem definida na qual os estágios da corrida de Atalanta são pontuados por descansos finitos, o que, por argumentos, mostra a possibilidade de completar uma série infinita de tarefas finitas em um tempo finito (Huggett 2010, pp. 21–22). Por fim, a distinção entre infinito potencial e infinito atual não desempenha nenhum papel na matemática desde que Cantor domou os números transfinitos – certamente o infinito potencial não tem desempenhado papel algum nas soluções matemáticas modernas discutidas aqui.

3.2 Aquiles e a Tartaruga

> O [segundo] argumento foi chamado "Aquiles", consequentemente, a partir do fato de que Aquiles foi tomado [como personagem] nele, e o argumento diz que é impossível para ele alcançar a tartaruga quando a persegue. Pois, de fato, é necessário que aquilo que está para alcançar [algo], antes de alcançar [isso], primeiro alcance o limite de onde aquilo que está fugindo partiu. No [tempo no] qual aquilo que persegue chega a isso, aquilo que está fugindo avançará um certo intervalo, mesmo se for menor do que aquele que aquilo que persegue avançou... E no tempo, de novo, no qual aquilo que persegue irá percorrer esse [intervalo] que aquilo que está fugindo avançou, nesse tempo, de novo, aquilo que está fugindo irá percorrer alguma quantidade... E assim, em todo tempo no qual aquilo que persegue irá percorrer o [intervalo] que aquilo que está fugindo, sendo mais lento, já avançou, aquilo que está fugindo também irá avançar alguma distância. (Simplício, *Sobre a Física de Aristóteles* 6, 1014.10)

Esse paradoxo gira em torno de considerações muito semelhantes ao anterior. Imagine

Aquiles perseguindo uma tartaruga e suponha que Aquiles corre a 1 *m/s*, que a tartaruga rasteja a 0,1 *m/s* e que a tartaruga começa 0,9m à frente de Aquiles. Aparentemente, Aquiles deveria alcançar a tartaruga após 1s, a uma distância de 1m de onde ele começa (e, portanto, 0,1m de onde a tartaruga começa). Poderíamos dividir o movimento de Aquiles em metades, como fizemos com Atalanta, ou da seguinte forma: antes que Aquiles possa alcançar a tartaruga, ele deve chegar ao ponto onde a tartaruga começou. Mas no tempo que ele leva para fazer isso, a tartaruga rasteja um pouco mais para frente. Então, Aquiles deve alcançar esse novo ponto. Porém, no tempo que Aquiles leva para isso, a tartaruga avança um pouquinho mais. E assim até o infinito: toda vez que Aquiles chega ao lugar onde a tartaruga estava, ela teve tempo suficiente para avançar mais um pouco, e então Aquiles tem outra corrida a fazer. Assim, Aquiles tem um número infinito de alcançadas finitas antes de pegar a tartaruga e, portanto, Zenão conclui que ele nunca a alcança.

Um aspecto do paradoxo, então, é que Aquiles deve percorrer a seguinte série infinita de distâncias antes de alcançar a tartaruga: primeiro 0,9m, depois mais 0,09m, depois 0,009m, … Essas são as distâncias que a tartaruga percorre no início de cada alcançada de Aquiles. Visto dessa forma, o quebra-cabeça é idêntico à Dicotomia, pois é o mesmo que dizer que "aquilo que está em movimento deve chegar [a nove décimos do caminho] antes de chegar ao objetivo". E, assim, tudo o que dissemos antes se aplica aqui também.

Mas o que esse paradoxo evidencia de forma mais vívida é o problema de completar uma série de ações que não tem um membro final – nesse caso, a série infinita de alcançadas antes que Aquiles atinja a tartaruga. Mas qual é exatamente o problema? Talvez o seguinte: a corrida de Aquiles até o ponto em que ele deveria alcançar a tartaruga pode, aparentemente, ser completamente decomposta na série de alcançadas, nenhuma das quais o leva à tartaruga. Portanto, em nenhum momento de sua corrida ele de fato a alcança. Mas, se isso era o que Zenão tinha em mente, não funciona. É claro que Aquiles não alcança a tartaruga em nenhum ponto da sequência, pois cada corrida na sequência ocorre *antes* que esperemos que ele a alcance! Pensando em termos dos pontos que Aquiles deve atingir em sua corrida, 1m não aparece na sequência 0,9m, 0,99m, 0,999m, …, então é óbvio que ele nunca pega a tartaruga durante essa sequência de corridas! (E a mesma situação surge na Dicotomia: não há uma primeira distância na série, então ela não inclui a partida de Atalanta!) Assim, a série de alcançadas não decompõe completamente a corrida: o ponto final – no qual Aquiles de fato alcança a tartaruga – deve ser adicionado a ela. Então, há algum paradoxo? Argumentavelmente, sim.

A corrida de Aquiles passa pela sequência de pontos 0,9m, 0,99m, 0,999m, …, 1m. Mas uma sequência tão estranha—composta por uma infinidade de membros seguida por mais um—faz sentido matematicamente? Se não, então nossa descrição matemática da corrida não pode estar correta, mas então qual é? Felizmente, a teoria dos transfinitos, com relação à qual Cantor é pioneiro, nos assegura que tal série é perfeitamente respeitável. Percebeu-se que as propriedades de ordem das séries infinitas são muito mais elaboradas que as das séries finitas. Qualquer forma de organizar os números 1, 2 e 3 resulta em uma série com o mesmo padrão, por exemplo, mas há muitas maneiras distintas de ordenar os números naturais: 1, 2, 3, … por exemplo. Ou …, 3, 2, 1. Ou …, 4, 2, 1, 3, 5, …. Ou 2, 3, 4, …, 1, que é exatamente o mesmo tipo de série que as posições que Aquiles deve percorrer. Assim, a teoria dos transfinitos trata não apenas dos números 'cardinais'—que dependem apenas de quantas coisas existem—mas também dos números 'ordinais', que dependem ainda de como as coisas estão ordenadas. Como os ordinais são normalmente considerados números matematicamente legítimos, e como a série de pontos que Aquiles deve passar tem um número ordinal, assumiremos que a série é matematicamente legítima. (Novamente, veja 'Supertarefas' abaixo para outro tipo de problema que pode surgir para Aquiles.)

3.3 A flecha

> O terceiro [paradoxo] é... que a flecha em voo está em repouso, conclusão que decorre da suposição de que o tempo é composto de momentos... Ele afirma que, se tudo que ocupa um espaço igual está em repouso, e se aquilo que está em movimento está sempre em um 'agora', então a flecha em voo está, portanto, imóvel. (Aristóteles, *Física*, 239b30).

> Zenão suprime o movimento, dizendo: "O que está em movimento não se move nem no lugar onde está, nem no lugar onde não está". (Diógenes Laércio, *Vidas e Doutrinas dos Filósofos Ilustres*, IX.72)

Esse argumento contra o movimento baseia-se explicitamente em um tipo específico de suposição de pluralidade: a de que o tempo é composto de momentos (ou 'agoras') *e nada mais*. (De acordo com Pemberton, 2022, Aristóteles rejeita por completo a afirmação de que o tempo é *composto* de instantes.) Considere uma flecha, aparentemente em movimento, em qualquer instante. Primeiro, Zenão assume que ela não percorre nenhuma distância durante esse momento – 'ela ocupa um espaço igual' durante todo o instante. Mas todo o período de seu movimento contém apenas instantes, todos os quais contêm a flecha em repouso e, portanto, Zenão conclui que a flecha não pode estar se movendo.

Uma preocupação imediata é entender por que Zenão está justificado ao assumir que a flecha está em repouso em qualquer instante. Isso decorre de modo imediato se assumirmos que um instante dura 0s: qualquer que seja a velocidade da flecha, ela não chegará a lugar algum se não tiver tempo algum. Mas e se considerarmos que as menores partes do tempo são finitas—ainda que minúsculas—, de modo que uma flecha em movimento poderia, de fato, percorrer alguma distância durante um instante? Uma maneira de sustentar essa suposição – que exige uma interpretação mais elaborada do texto – começa assumindo que os instantes são indivisíveis. Em seguida, suponha que a flecha realmente se moveu durante um instante. Ela estaria em lugares diferentes no início e no fim do instante, o que implica que o instante tem um 'início' e um 'fim', o que, por sua vez, sugere que ele tem pelo menos duas partes e, portanto, é divisível—contrariando nossa suposição inicial. (Observe que esse argumento apenas estabelece que nada pode se mover durante um instante, não que os instantes não podem ser finitos.)

Assim, nada se move durante qualquer instante, mas o tempo é inteiramente composto de instantes; logo, nada jamais se move. Uma primeira resposta é observar que determinar a velocidade da flecha implica dividir a distância percorrida em certo tempo pela duração desse tempo. Mas – assumindo, a partir de agora, que os instantes têm duração zero—essa fórmula não faz sentido no caso de um instante: a flecha percorre 0m em 0s (a duração do instante), mas 0/0 m/s não é um número. Portanto, é falacioso concluir, a partir do fato de que a flecha não percorre distância alguma em um instante, que ela está em repouso; se ela está ou não em movimento em um instante depende se ela percorre alguma distância em um intervalo *finito* que inclua o instante em questão.

A resposta está correta, mas traz uma implicação contraintuitiva: o movimento não é algo que acontece em qualquer instante, mas apenas ao longo de períodos finitos de tempo. Pense da seguinte maneira: o tempo, como dissemos, é composto apenas de instantes. Nenhuma distância é percorrida durante qualquer instante. Então, quando a flecha realmente se move? Como ela vai de um lugar a outro em um momento posterior? Há apenas uma resposta: a flecha vai do ponto *X* no tempo 1 ao ponto *Y* no tempo 2 simplesmente em virtude de estar em pontos intermediárias sucessivas em tempos intermediários sucessivos—a flecha nunca muda sua posição durante um instante, mas apenas ao longo de intervalos compostos de instantes, pela ocupação de diferentes posições em diferentes momentos. Nas palavras memoráveis de Bergson – que ele considerava expressar um absurdo –, “o movimento é composto de imobilidades”(1911, 308): ir de *X* a *Y* consiste em ocupar exatamente um lugar intermediário em cada instante (na ordem correta, é claro). Para uma discussão mais aprofundada dessa concepção ‘ponto a ponto’ do tempo, consulte Arntzenius (2000) e Salmon (2001, 23-4).

3.4 O Estádio

> O quarto argumento é o que diz respeito a corpos iguais que se movem ao lado de corpos iguais no estádio a partir de direções opostas – uns do final do estádio, outros do meio – em velocidades iguais, no qual ele pensa que se segue que metade do tempo é igual ao seu dobro.... (*Física* de Aristóteles, 239b33)

Aristóteles prossegue elaborando e refutando um argumento para o paradoxo final do movimento de Zenão. O texto é bastante enigmático, mas geralmente é interpretado da seguinte maneira: imagine três conjuntos de cubos idênticos e justapostos em movimento relativo. Um conjunto—os As—permanece em repouso, enquanto os outros—os *B*s e *C*s—movem-se para a direita e para a esquerda, respectivamente, em velocidades iguais e constantes. Suponha ainda que, em certo momento, o *B* mais à direita e o *C* mais à esquerda estejam alinhados com o *A* central, conforme ilustrado (três de cada são representados por simplicidade):

```
  A A A
B B B
    C C C
```

Como os *B*s e *C*s se movem na mesma velocidade, eles se alinharão simultaneamente com os *A*s em um momento posterior:

```
A A A
B B B
C C C
```

Nesse instante, o *B* mais à direita terá passado por todos os *C*s, mas apenas pela metade dos *A*s. Como os cubos são do mesmo tamanho, ele terá percorrido uma certa distância *e também* metade dessa distância. No entanto, a contradição não é explicitada aqui, presumivelmente porque é claro que essas distâncias contraditórias são relativas aos *C*s e *A*s, respectivamente—e não há contradição geral em estar em relações diferentes com coisas diferentes. Em vez disso, as distâncias são convertidas em tempos ao serem divididas pela velocidade dos *B*s: metade da distância, a uma dada velocidade, leva metade do tempo. Surge então uma contradição, pois o tempo entre os estados é de fato inequívoco, não relativo—o processo leva um tempo (não nulo) e metade desse tempo.

O veredito geral é que Zenão estava profundamente confuso sobre velocidades relativas nesse paradoxo. Se os *B*s se movem a uma velocidade *V m/s* para a direita em relação aos *A*s, e os *C*s a *V m/s* para a esquerda em relação aos *A*s, então os *C*s se movem a *2V m/s* para a esquerda em relação aos *B*s. Assim, embora os Bs percorram o dobro da distância em relação aos *C*s do que em relação aos *A*s, fazem-no com o dobro da velocidade relativa, de modo que os tempos são iguais em ambos os casos. Mas será que Zenão poderia estar tão confuso? (Sattler, 2015, argumenta contra essa e outras interpretações comuns do estádio.)

Talvez (Davey, 2007) ele tivesse em mente o seguinte (embora essa leitura, que atribui mais perspicácia a Zenão, não se encaixe tão bem nas palavras de Aristóteles): suponha que os *A*s, *B*s e *C*s tenham a menor extensão espacial possível—sejam "pontuais"—, onde "pontos"têm tamanho zero se o espaço for contínuo, ou tamanho finito se o espaço for "atômico". Suponha ainda que não haja espaços entre os *A*s, *B*s ou *C*s. Durante o movimento descrito, o *B* mais à frente passa por todos os *C*s e pela metade dos *A*s—ou seja, por metade do número de *A*s em relação aos *C*s. Agora, como um ponto se move continuamente ao longo de uma linha sem lacunas, há uma correspondência biunívoca entre os instantes de tempo e os pontos na linha—para cada instante, um ponto; para cada ponto, um instante. Portanto, o número de "instantes-*A*"que o *B* está na frete leva para passar pelos *A*s é metade do número de "instantes-*C*"necessários para passar pelos *C*s—mesmo que esses processos levem o mesmo tempo. Se então, crucialmente, assumirmos que metade dos instantes significa metade do tempo, concluímos que metade do tempo é igual ao tempo total, uma contradição.

Como vimos na discussão sobre divisibilidade completa, há um problema nesse raciocínio aplicado a linhas contínuas: qualquer segmento de linha tem o mesmo número de pontos, logo nada pode ser inferido a partir do número de pontos dessa maneira— certamente não pode ser inferido que metade dos pontos (aqui, instantes) signifique metade do comprimento (ou tempo). O paradoxo, como formulado, falha. Mas a própria afirmação de que os intervalos contêm o mesmo número de instantes não entra em conflito com a conclusão de que há metade de instantes-*A* em relação aos instantes-*C*? A questão é sutil para conjuntos infinitos: por exemplo, 1, 2, 3, … está em correspondência biunívoca com 2, 4, 6, …, e portanto há o mesmo ”número”de cada. É nesse sentido de correspondência—o sentido matemático preciso de ”mesmo número”—que qualquer segmento finito tem a mesma quantidade de pontos que outro. No entanto, informalmente, há ”metade”dos números pares em relação aos inteiros: os pares (1,2), (3,4), (5,6), … também podem ser postos em correspondência biunívoca com 2, 4, 6, …. Analogamente, há—informalmente—metade dos instantes-A em relação aos instantes-*C*: os instantes-*A* correspondem biunivocamente a pares de instantes-*C*. Assim, não há contradição na contagem de pontos: a metade informal equivale ao todo estrito (uma solução diferente é necessária para uma teoria atômica, conforme sugerido no parágrafo final dessa seção).

(Permita-me mencionar um paradoxo similar do movimento—a “roda de moinho”—atribuído a Maimônides. Imagine duas rodas, uma com o dobro do raio e circunferência da outra, fixas a um mesmo eixo. Deixe-as rolar sobre um trilho, com um dos lados elevado para manter o eixo horizontal, para que ambas as rodas deem uma volta [elas giram na mesma velocidade por causa do eixo, cada ponto de cada roda entra em contato com exatamente um ponto do trilho correspondente, e cada ponto de cada trilho com exatamente um ponto de sua roda correspondente. Após uma volta completa, a montagem percorreu uma distância igual à circunferência da roda maior? Da menor? De Ambas? De alguma outra? Como? Esse problema também exige compreensão do contínuo, mas não é um paradoxo de Zenão—deixemos sua resolução à engenhosidade do leitor.)

Uma última reconstrução possível do Estádio de Zenão o interpreta como um argumento contra uma teoria atômica do espaço e tempo—o que é interessante porque a física contemporânea explora essa ideia ao tentar “quantizar”o espaço-tempo. Suponha, então, que os lados de cada cubo equivalem ao “quantum”de comprimento e que os dois momentos considerados estão separados por um único quantum de tempo. Algo estranho deve ocorrer: o B mais à direita e o C central passam um pelo outro durante o movimento, mas não há momento em que estejam alinhados—já que os dois momentos estão separados pelo menor tempo possível, não pode haver um instante intermediário (seria um tempo menor que o menor intervalo considerado). Por outro lado, se insistíssemos que, ao passarem, deve haver um momento de alinhamento, isso provaria que não pode haver um intervalo finito maiscurto—qualquer que fosse, esse argumento o refutaria. Mas por que insistir nessa suposição? O problema é que se imagina naturalmente o espaço quantizado como um tabuleiro de xadrez, no qual as peças permanecem imóveis durante cada quantum de tempo. Então, pergunta-se quando a rainha vermelha, por exemplo, vai de uma casa à outra, ou como ultrapassa a rainha branca sem se alinhar a ela. Mas a analogia é enganosa. É melhor pensar no espaço quantizado como uma matriz gigante de lâmpadas, onde cada padrão de luzes acesas representa um quantum de tempo. Nessa analogia, uma lâmpada acesa representa a presença de um objeto—por exemplo, uma sequência de luzes acendendo-se em linha representa um corpo em movimento retilíneo. Aqui, não há tentação de perguntar quando a luz "chega"de uma lâmpada à próxima—ou, em analogia, como o corpo se move de um local a outro. (Aqui tangenciamos questões sobre partes temporais e se objetos "perduram"ou "subsistir".)

4. Dois Paradoxos Adicionais

Dois outros paradoxos são atribuídos a Zenão por Aristóteles, mas eles são apresentados no contexto de outros argumentos que ele desenvolve, de modo que a intenção original de Zenão não pode ser determinada com certeza—nem mesmo se eles foram concebidos para argumentar contra a pluralidade e o movimento. Discutiremos brevemente esses paradoxos para que a exposição seja completa.

4.1 O Paradoxo do Lugar

> A dificuldade levantada por Zenão exige uma explicação; pois, se tudo o que existe ocupa um lugar, então o próprio lugar também deverá ter um lugar, e assim sucessivamente, *ad infinitum*. (Aristóteles, *Física*, 209a23)

Quando Aristóteles estabelece sua teoria do lugar - noção espacial crucial em sua teoria do movimento - ele enumera várias teorias e problemas que seus predecessores, incluindo Zenão, formularam sobre o assunto. O argumento novamente levanta questões sobre o infinito, já que a segunda etapa do argumento defende uma regressão infinita de lugares. No entanto, Aristóteles o apresenta como um argumento contra a própria ideia de lugar, em vez de pluralidade (provavelmente tirando-o de contexto). ÉÉ difícil entender a necessidade da conclusão, pois por que não deveria haver uma série infinita de lugares de lugares de lugares...? Presumivelmente, a preocupação seria maior para alguém que (como Aristóteles) acreditasse que não poderia haver um infinito real de coisas, pois o argumento parece mostrar que há. Mas como discutimos acima, hoje não precisamos ter tais preocupações; não parece haver nada problemático com um infinito real de lugares.

A única outra maneira pela qual alguém poderia achar a regressão problemática é se sustentasse que os corpos têm lugares ‘absolutos’, no sentido de que há sempre uma resposta privilegiada única para a pergunta ‘onde está’? O problema então não é que existam infinitos lugares, mas apenas que existem muitos. E Aristóteles pode ter tido essa preocupação, pois em sua teoria do movimento, o movimento natural de um corpo é determinado pela relação de seu lugar com o centro do universo: uma explicação que exige que o lugar seja determinado, porque o movimento natural é. (Veja Sorabji 1988 e Morrison 2002 para considerações gerais e concorrentes sobre as visões de Aristóteles sobre o lugar; capítulo 3 deste último especialmente para uma discussão sobre o tratamento do paradoxo por Aristóteles.) Mas supondo que alguém sustente que o lugar é absoluto por qualquer motivo, então, por exemplo, onde estou enquanto escrevo? Se o paradoxo estiver correto, então estou no meu lugar, e também no lugar do meu lugar, e no lugar do lugar do meu lugar... Como estou em todos esses lugares, qualquer um pode parecer uma resposta apropriada à pergunta. Várias respostas são concebíveis: negar lugares absolutos (especialmente porque nossa física não os exige), definir uma noção de lugar que seja única em todos os casos (provavelmente a solução de Aristóteles), ou talvez afirmar que lugares são seus próprios lugares, interrompendo assim a regressão!

4.2 O grão de Trigo

> O raciocínio de Zenão é falso ao afirmar que não existe parte alguma do trigo que não produza som, pois não há razão pela qual qualquer parte deva - em qualquer intervalo de tempo - mover o ar que a medida inteira move ao cair. (Aristóteles, *Física*, 250a19)

No contexto, Aristóteles explica que uma fração de força pode não produzir proporcionalmente a mesma fração de movimento. Por exemplo, enquanto cem estivadores podem rebocar um barco, um único estivador pode não conseguir movê-lo de forma alguma, muito menos a 1/100 da velocidade; portanto, mesmo com tempo ilimitado, ele não o moveria na mesma proporção que os cem (hoje atribuímos isso ao efeito do atrito). Da mesma forma, apenas porque uma medida de trigo em queda produz um ruído, não se segue que cada grão individual faça: mesmo com tempo ilimitado, um grão não moveria o mesmo volume de ar que a medida completa. Entretanto, ao refutar essa premissa, Aristóteles não esclarece qual função ela cumpria no argumento de Zenão, restando-nos apenas conjecturar. Nem mesmo está claro se fazia parte de um paradoxo ou de outra discussão: teria Zenão também alegado que um único grão de trigo não produz som? Uma hipótese é que nossos sentidos confirmariam que *não*, já que não ouvimos um grão isolado cair. Nesse caso, a resposta de Aristóteles seria apropriada - assim como a observação correlata de que a própria audição requer movimento do ar acima de determinado limiar.

5. A Influência de Zenão na Filosofia

Nesta seção final, devemos considerar brevemente o impacto que Zenão teve em vários filósofos; uma pesquisa da literatura revelará que esses debates continuam.

Os Pitagóricos: Na primeira metade do século XX, a interpretação predominante - seguindo Tannery (1885) - sustentava que os argumentos de Zenão eram dirigidos contra uma doutrina técnica dos pitagóricos. Segundo essa leitura, eles acreditavam que todas as coisas eram compostas de elementos que tinham as propriedades de um número unitário, um ponto geométrico e um átomo físico: essa posição se alinharia com sua doutrina de que a realidade é fundamentalmente matemática. No entanto, em meados do século, uma série de comentaristas (Vlastos, 1967, resume o argumento e contém referências) argumentou vigorosamente que o alvo de Zenão era, em vez disso, uma compreensão comum da pluralidade e do movimento - baseada em noções geométricas familiares - e que essa doutrina não era uma parte importante do pensamento pitagórico. Assumimos implicitamente que esses argumentos estão corretos em nossas leituras dos paradoxos. Dito isso, a interpretação de Tannery ainda tem seus defensores (ver, por exemplo, Matson 2001).

Os Atomistas: Aristóteles (*Sobre a Geração e Corrupção* 316b34) afirma que nosso terceiro argumento - aquele sobre a divisibilidade completa - foi o que convenceu os atomistas de que deveria haver partes mínimas e indivisíveis da matéria. Ver Abraham (1972) para uma discussão mais aprofundada sobre a conexão de Zenão com os atomistas.

O Devir Temporal: No início do século XX, vários filósofos influentes tentaram empregar os argumentos de Zenão a serviço de uma metafísica do ”devir temporal”, o (suposto) processo pelo qual o presente vem a ser. Pensadores como Bergson (1911), James (1911, Cap. 10-11) e Whitehead (1929) argumentaram que os paradoxos de Zenão mostram que o espaço e o tempo não estão estruturados como um *continuum* matemático: eles argumentaram que a maneira de preservar a realidade do movimento era negar que o espaço e o tempo são compostos de pontos e instantes. No entanto, vimos claramente que as ferramentas da matemática moderna padrão são capazes de resolver os paradoxos, portanto nenhuma conclusão desse tipo parece justificada: se o presente de fato "vem a ser", não há razão para pensar que o processo não seja capturado pelo *continuum*.

Aplicando o *Continuum* Matemático ao Espaço e Tempo Físicos: Como observado na §1.2, a "visão recebida"de Zenão (desenvolvida na segunda metade do século XX por filósofos que elaboraram as ideias de Grünbaum, 1967) visava mostrar como a matemática moderna resolve os paradoxos. No entanto, central para esse projeto era o reconhecimento de que uma solução puramente matemática não é suficiente: os paradoxos questionam não apenas a matemática abstrata, mas também a natureza da realidade física. Portanto, o que eles buscavam era um argumento não apenas de que Zenão não representava uma ameaça à matemática do infinito, mas também de que essa matemática descreve corretamente objetos, tempo e espaço. Não seria uma resposta aos paradoxos de Zenão se o arcabouço matemático que invocamos não fosse uma boa descrição do espaço, tempo e movimento reais! A ideia de que uma lei matemática - digamos, a lei da gravitação universal de Newton - pode ou não descrever corretamente as coisas é familiar, mas alguns aspectos da matemática do infinito - a natureza do *continuum*, a definição de somas infinitas etc. - parecem tão básicos que pode ser difícil perceber a princípio que eles também se aplicam contingentemente. Mas certamente o fazem: nada garante *a priori* que o espaço tenha a estrutura do *continuum*, ou mesmo que partes do espaço se somem de acordo com a definição de Cauchy. (Salmon oferece um bom exemplo para ajudar a esclarecer o ponto: como o álcool se dissolve na água, se você misturar os dois, acabará com menos do que a soma de seus volumes, mostrando que mesmo a adição comum não é aplicável a todo tipo de sistema.) Nossa crença de que a teoria matemática do infinito descreve o espaço e o tempo é justificada na medida em que as leis da física assumem que sim, e na medida em que essas leis são confirmadas pela experiência. Embora seja verdade que quase todas as teorias físicas assumem que o espaço e o tempo de fato têm a estrutura do *continuum*, também é o caso que teorias quânticas da gravidade provavelmente implicam que não. Embora ninguém saiba ao certo onde essa pesquisa levará, é bem possível que o espaço e o tempo acabem sendo, no nível mais fundamental, muito diferentes do *continuum* matemático que assumimos aqui.

Deve-se notar também que Grünbaum considerou que a tarefa de mostrar que a matemática moderna descreve o espaço e o tempo envolvia algo bastante diferente de argumentar que ela é confirmada pela experiência. A visão dominante na época (embora não atualmente) era que os termos científicos tinham significado na medida em que se referiam diretamente a objetos de experiência - como "uma régua de 1m- ou, se se referissem a entidades "teóricas"em vez de "observáveis- como "um ponto do espaço"ou "1/2 de 1/2 de ... 1/2 de uma pista- então eles obtinham significado por suas relações lógicas - por meio de definições e leis teóricas - com tais termos de observação. Assim, Grünbaum empreendeu um programa impressionante para dar significado a todos os termos envolvidos na teoria moderna do infinito, interpretada como uma descrição do espaço e do tempo.

Supertarefas: Uma linha adicional de pensamento diz respeito ao que Black (1950-51) chamou de "máquinas infinitas". Black e seus seguidores queriam mostrar que, embora os paradoxos de Zenão não representassem um problema para a matemática, eles mostravam que, afinal, a matemática não era aplicável ao espaço, tempo e movimento. Mais drasticamente, nossa resolução para a Dicotomia e Aquiles assumiu que a corrida completa poderia ser dividida em uma série infinita de meias corridas, que poderiam ser somadas. Mas é realmente possível completar qualquer série infinita de ações: completar o que é conhecido como uma "supertarefa"? Se não, e assumindo que Atalanta e Aquiles podem completar suas tarefas, suas corridas completas não podem ser corretamente descritas como uma série infinita de meias corridas, embora a matemática moderna as descreva assim. O que as máquinas infinitas supostamente estabelecem é que uma série infinita de tarefas não pode ser completada - portanto, qualquer tarefa completável não pode ser dividida em uma infinidade de tarefas menores, independentemente do que a matemática sugira.

Infinitesimais: Finalmente, vimos como abordar os paradoxos usando os recursos da matemática desenvolvida no século XIX. Por muito tempo, considerou-se uma das grandes virtudes desse sistema que ele finalmente mostrou que quantidades infinitesimais, menores que qualquer número finito, mas maiores que zero, são desnecessárias. (O cálculo de Newton, por exemplo, efetivamente fazia uso de tais números, tratando-os às vezes como zero e às vezes como finitos; o problema com essa abordagem é que a forma de tratar os números é uma questão de intuição, não de rigor.) No entanto, no século XX, Robinson mostrou como introduzir números infinitesimais na matemática: esse é o sistema da "análise não-padrão"(o sistema familiar de números reais, fundamentado rigorosamente por Dedekind, é, em contraste, apenas “análise”). Analogamente, Bell (1988) explica como segmentos de linha infinitesimais podem ser introduzidos na geometria e comenta sua relação com Zenão. Além disso, McLaughlin e Miller (1992) e McLaughlin (1994) mostram como os paradoxos de Zenão podem ser resolvidos na análise não-padrão; eles não são mais um argumento contra a análise não-padrão do que contra a matemática padrão que assumimos aqui. Deve-se enfatizar, no entanto, que - ao contrário das sugestões de McLaughlin e Miller - não há necessidade de análise não-padrão para resolver os paradoxos: qualquer um dos sistemas é igualmente bem-sucedido. (Reeder, 2015, argumenta que a análise não-padrão é insatisfatória em relação à flecha e oferece uma explicação alternativa usando uma concepção diferente de infinitesimais.) A construção da análise não-padrão levanta, no entanto, uma questão adicional sobre a aplicabilidade da análise ao espaço e tempo físicos: parece plausível que todas as teorias físicas possam ser formuladas em ambos os termos e, portanto, até onde nossa experiência se estende, ambas parecem igualmente confirmadas. Mas ambas não podem ser verdadeiras para o espaço e o tempo: ou o espaço tem partes infinitesimais ou não.

6. Leituras Complementares

Após as entradas relevantes nessa enciclopédia, o ponto de partida para qualquer investigação adicional é Salmon (2001), que contém alguns dos artigos mais importantes sobre Zenão até 1970, além de uma bibliografia impressionantemente abrangente de obras em inglês no século XX. Um tratamento mais recente e aprofundado, que contextualiza o debate no desenvolvimento do conceito de movimento na Antiguidade, é Sattler (2020).

Também se pode consultar Huggett (1999, cap. 3) e Huggett (2010, caps. 2–3) para mais trechos de fontes e discussões. Para introduções às ideias matemáticas por trás das resoluções modernas, o Apêndice de Salmon (2001) ou Stewart (2017) são bons pontos de partida; Russell (1919) e Courant et al. (1996, caps. 2 e 9) também são excelentes fontes. Por fim, três coletâneas de fontes originais sobre os paradoxos de Zenão: Lee (1936 [2015]) contém tudo o que se conhece, Kirk *et al*. (1983, cap. 9) reúne uma grande quantidade de material (em inglês e grego) com comentários úteis, e Cohen *et al*. (1995) também apresenta os principais trechos.

Referências Bibliográficas

W. E. Abraham. The nature of zeno's argument against plurality in dk 29 b i. *Phronesis*, 17: 40–52, 1972.

Aristotle. On generation and corruption. In Jonathan Barnes, editor, *The Complete Works of Aristotle*. Princeton University Press, Princeton, 1984a.

Aristotle. Physics. In Jonathan Barnes, editor, *The Complete Works of Aristotle*. Princeton University Press, Princeton, 1984b.

F. Arntzenius. Are there really instantaneous velocities? *The Monist*, 83:187–208, 2000.

J. L. Bell. Infinitesimals. *Synthese*, 75(3):285–315, 1988.

G. Belot and J. Earman. Pre-socratic quantum gravity. In C. Callender and N. Huggett, editors, *Physics Meets Philosophy at the Planck Scale: Contemporary Theories in Quantum Gravity*. Cambridge University Press, Cambridge, 2001.

H. Bergson. *Creative Evolution*. Holt, Reinhart and Winston, New York, 1911.

M. Black. Achilles and the tortoise. *Analysis*, 11:91–101, 1950.

S. M. Cohen, P. Curd, and C. D. C. Reeve, editors. *Readings in Ancient Greek Philosophy From Thales to Aristotle*. Hackett Publishing Co. Inc., Indianapolis/Cambridge, 1995.

R. Courant, H. Robbins, and I. Stewart. *What is Mathematics? An Elementary Approach to Ideas and Methods*. Oxford University Press, New York, Oxford, 2 edition, 1996.

K. Davey. Aristotle, zeno, and the stadium paradox. *History of Philosophy Quarterly*, 24: 127–146, 2007.

P. Ehrlich. An essay in honor of adolf grünbaum's ninetieth birthday: A reexamination of zeno's paradox of extension. *Philosophy of Science*, 81(4):654–675, 2014.

A. Grünbaum. *Modern Science and Zeno's Paradoxes*. Connecticut Wesleyan University Press, Middletown, 1967.

N. Huggett, editor. *Space from Zeno to Einstein: Classic Readings with a Contemporary Commentary*. MIT Press, Cambridge, MA, 1999.

N. Huggett. *Everywhere and Everywhen: Adventures in Physics and Philosophy*. Oxford University Press, Oxford, 2010.

W. James. *Some Problems of Philosophy*. Longmans, Green & Co., New York, 1911.

G. S. Kirk, J. E. Raven, and M. Schofield, editors. *The Presocratic Philosophers: A critical History with a Selection of Texts*. Cambridge University Press, Cambridge, 2 edition, 1983.

Diogenes Laertius. Lives of famous philosophers. In G. S. Kirk, J. E. Raven, and M. Schofield, editors, *The Presocratic Philosophers: A Critical History with a Selection of Texts*, page 273. Cambridge University Press, Cambridge, 2 edition, 1983.

H. D. P. Lee, editor. *Zeno of Elea: A Text, with Translation and Notes*. Cambridge University Press, Cambridge, 1936. Reprinted 2015.

W. I. Matson. Zeno moves! In A. Preus, editor, *Essays in Ancient Greek Philosophy VI: Before Plato*. State University of New York Press, Albany, 2001.

W. I. McLaughlin. Resolving zeno's paradoxes. *Scientific American*, 271(5):84–89, 1994.

W. I. McLaughlin and S. L. Miller. An epistemological use of nonstandard analysis to answer zeno's objections against motion. *Synthese*, 92:371–384, 1992.

B. Morison. *On Location: Aristotle's Concept of Place*. Oxford University Press, Oxford, 2002.

I. Newton. *The Principia: Mathematical Principles of Natural Philosophy*. University of California Press, Berkeley, 1999.

J. M. Pemberton. Aristotle's solution to zeno's arrow paradox and its implications. *Ancient Philosophy Today*, 4:73–95, 2022.

Plato. Parmenides. In John M. Cooper, editor, *Plato: Complete Works*. Hackett Publishing Co. Inc., Indianapolis, 1997.

P. Reeder. Zeno's arrow and the infinitesimal calculus. *Synthese*, 192:1315–1335, 2015.

B. Reese, M. Vazquez, and S. Weinstein. How can a line segment with extension be composed of extensionless points? from aristotle to borel, and beyond. *Synthese*, 200(2):1–28, 2022.

B. Russell. *Introduction to Mathematical Philosophy*. George Allen and Unwin Ltd., London, 1919.

B. Russell. *Our Knowledge of the External World*. W. W. Norton & Co. Inc., New York, 1929.

W. C. Salmon. *Zeno's Paradoxes*. Hackett Publishing Co. Inc., Indianapolis, 2 edition, 2001.

B. M. Sattler. Time is double the trouble: Zeno's moving rows. *Ancient Philosophy*, 35:1–22, 2015.

B. M. Sattler. *The Concept of Motion in Ancient Greek Thought: Foundations in Logic, Method, and Mathematics*. Cambridge University Press, Cambridge, 2020.

D. M. Sherry. Zeno's metrical paradox revisited. *Philosophy of Science*, 55:58–73, 1988.

Simplicius. *On Aristotle's Physics 6*. Gerald Duckworth & Co. Ltd, London, 1989.

Simplicius. On aristotle's physics. In S. M. Cohen, P. Curd, and C. D. C. Reeve, editors, *Readings in Ancient Greek Philosophy From Thales to Aristotle*, pages 58–59. Hackett

Publishing Co. Inc., Indianapolis, 1995.

R. Sorabji. *Matter, Space, and Motion Theories in Antiquity and Their Sequel*. Cornell University Press, Ithaca, 1988.

I. Stewart. *Infinity a Very Short Introduction*. Oxford University Press, Oxford, 2017.

P. Tannery. Le concept scientifique du continu: Zenon d'elee et georg cantor. *Revue Philosophique de la France et de l'Etranger*, 20:385–410, 1885.

G. Vlastos. Zeno of elea. In P. Edwards, editor, *The Encyclopedia of Philosophy*. The Macmillan Co. and The Free Press, New York, 1967.

A. N. Whitehead. *Process and Reality*. Macmillan Co., New York, 1929.

(V) Paradoxo de Skolem[1]

Título Original: Skolem's Paradox
Autor: Thimothy Bays
Tradução: Mahan Vaz[2]
Revisão: Diego Fernandes

O paradoxo de Skolem envolve um conflito aparente entre dois teoremas da lógica clássica. O teorema de Löwenheim-Skolem diz que se uma teoria de primeira ordem possui modelos infinitos, então ela tem modelos cujos domínios são enumeráveis . O teorema de Cantor diz que alguns conjuntos são não-enumeráveis. O paradoxo de Skolem ocorre quando percebemos que os princípios básicos da teoria de conjuntos de Cantor – i.e., aqueles princípios utilizados para provar o teorema de Cantor sobre a existência de conjuntos não-enumeráveis – podem eles próprios serem formulados como uma coleção de sentenças de primeira ordem. Como podem os mesmos princípios que provam a existência de conjuntos não-enumeráveis serem satisfeitos por um modelo que é ele próprio apenas enumerável? Como pode um modelo enumerável satisfazer uma sentença de primeira ordem

[1] "Skolem's Paradox", In: ZALTA, E. N. (ed.). Stanford Encyclopedia of Philosophy. Stanford, CA: The Metaphysics Research Lab. Disponível em: https://plato.stanford.edu/entries/paradox-skolem/.
A seguir está a tradução da entrada sobre o Paradoxo de Skolem na Stanford Encyclopedia of Philosophy. A tradução segue a versão da entrada nos arquivos da SEP em https://plato.stanford.edu/archives/spr2025/entries/paradox-skolem/. 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/paradox-skolem/. Gostaríamos de agradecer aos editores da Stanford Encyclopedia of Philosophy, em especial o Prof. Dr. Edward N. Zalta, por concederem permissão para traduzir e publicar esta entrada.
[2] Faz-se aqui indispensável a menção à colaboração das agências de fomento que financiam o desenvolvimento deste trabalho. O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processo nº 2022/16816-9 e pela DAAD, sob o programa Binationallysupervised/Cotutelle Doctoral Degree. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade do(s) autor(es) e não necessariamente refletem a visão da FAPESP ou da DAAD.

que diz que há uma quantidade não-enumerável de objetos – por exemplo, uma quantidade não-enumerável de números reais?

A discussão filosófica deste paradoxo teve uma tendência de se focar em três questões principais. Primeiro, há uma questão puramente matemática: por que o paradoxo da Skolem não introduz uma contradição imediata à teoria de conjuntos? Segundo, há uma questão histórica: o próprio Skolem deu uma explicação muito boa do porquê esse paradoxo não constitui uma contradição matemática imediata; por que, então, Skolem e suas contemporâneas continuaram a considerar o paradoxo tão problemático filosoficamente? Finalmente há a questão filosófica: o que nos diz o Paradoxo de Skolem sobre nosso entendimento sobre teoria de conjuntos e/ou sobre a semântica da linguagem da teoria de conjuntos?

1. Introdução

Para entendermos o Paradoxo de Skolem, precisamos lembrar de dois teoremas da lógica clássica[3]. O primeiro data do final do século XIX. Em 1873, Georg Cantor formulou uma nova técnica para mediar o tamanho – ou cardinalidade – de um conjunto de objetos. A ideia de Cantor era a de que dois conjuntos deveriam ter o mesmo tamanho apenas no caso em que seus elementos pudessem ser colocados numa correspondência um-para-um entre

---

[3]Aqui é provavelmente o local para destacar algumas convenções notacionais. Nessa entrada, assumiremos que $=, \neg, \rightarrow,$ e $\exists$ constituem o vocabulário oficial da lógica de primeira-ordem, e trataremos $\wedge, \vee, \leftrightarrow,$ e $\forall$ como abreviações. (Faremos uso livre dessas abreviações quando elas pareçam melhorar a leitura). A menos quando dito o contrário, trabalharemos na linguagem da teoria de conjuntos de primeira-ordem – isto é, a linguagem que possui "$\in$" como seu único primitivo não-lógico. Dizemos que uma fórmula está *nesta* linguagem quando ela é construída a partir de $\in$ juntamente com os conectivos padrão da lógica.
Nessa entrada utilizaremos letras em negrito para denotar modelos e suas letras correspondentes sem negrito para denotar os domínios desses modelos. Assim, **M** é um modelo e M é seu domínio, **N** é um modelo e M é seu domínio, etc. Tendo dito isto, frequentemente abusamos a notação e escrevemos coisas como "**M** é enumerável" ou "$m \in$ **M**" quando na verdade queremos dizer "M é enumerável" e "$m \in M$"; no contexto, isso nunca deve causar confusão. Finalmente, a menos que especificado o contrário, assume-se que todos os modelos são para a linguagem da teoria de conjuntos – isto é, como o caso acima, a linguagem com $\in$ como seu único primitivo não-lógico.
A menos quando observado o contrário, utilizaremos $\vDash$ para denotar a satisfação de primeira-ordem e $\vdash$ para provabilidade de primeira ordem. Se m é um elemento de algum modelo **M** e $\phi(x)$ é uma fórmula com apenas x livre, então escreveremos **M** $\vDash \varphi[m]$ para dizer que m habita o subconjnto de **M** em que $\varphi(x)$ é verdadeira. Para mais sobre notação de teoria de modelos, veja as entradas em teoria de modelos e teoria de modelos de primeira ordem. Para o básico de notação em teoria de conjuntos, veja a entrada em teoria de conjuntos.

eles. Por exemplo, o conjunto $\{1, 2, \dots, 26\}$ pode ser colocado em uma correspondência um-para-um com o conjunto $\{A, B, \dots, Z\}$, através do mapemamento natural que relaciona 1 a A, 2 a B, 3 a C, etc., etc; de modo similar, o conjunto dos números naturais pode ser colocado em uma correspondência um-para-um com o conjunto dos números pares através do mapeamento: $x \mapsto 2x$.

Ao aplicar este conceito de cardinalidade a conjuntos infinitos, Cantor chegou à conclusão inicialmente surpreendente de que há diferentes tipos de infinitude. Há conjuntos infinitos relativamente pequenos, como o conjunto dos números pares, o conjunto dos inteiros, ou o conjunto dos números racionais. Estes são conjuntos que podem todos ser colocados em uma correspondência um-para-um com o conjuntos dos números naturais; eles são chamados *infinitos enumeráveis*. Em contraste, há conjuntos infinitos muitos "maiores", como o conjuntos dos números reais, o conjunto dos números complexos, ou o conjuntos de todos os subconjuntos dos números naturais. Estes conjuntos são muito grandes para serem colocados em correspondência um-para-um com os números naturais; eles são chamados *infinitos não-enumeráveis*. O Teorema de Cantor, então, é apenas a asserção de que há conjuntos infinitos não-enumeráveis – conjuntos que são, como vistos na época, muito grandes para serem enumeráveis[4].

Nosso segundo teorema data do início do século XX. Em 1915, Leopold Löwenheim demonstrou que se uma sentença de primeira-ordem tem um modelo, então ela tem um modelo cujo domínio é enumerável[5]. Em 1922, Thoralf Skolem generalizou este resultado para

[4]Dissemos acima que a conclusão de Cantor de que há diferentes tipos de infinito pode ser inicialmente surpreendente. Isso necessita qualificação. Num primeiro momento, afinal, parece óbvio que alguns conjuntos infinitos são maiores que outros. Se, por exemplo, medirmos o "tamanho" de um conjunto por meio de uma relação de subconjunto – assim, A é menor que B se, e somente se, A é um subuconjunto próprio de B – então pode ser trivial mostrar que há diferentes tipos de infinito. *Nessa* definição, afinal, o conjunto dos números pares é menor que o conjunto dos números naturais, o conjunto dos naturais é menor que o conjunto dos inteiros e o conjunto dos inteiros é menor que o conjuntos dos racionais. É só quando filtramos toda a nossa análise através da nova definição de Cantor de cardinalidade – e então descobrimos que toda a sequência dos números pares, naturais, inteiros, racionais, etc, consiste em conjuntos com o mesmo tamanho – que começamos a suspeitar que *todos* os conjuntos infinitos possuem o mesmo tamanho, o que, de um jeito ou de outro, por bem ou por mal, deveríamos ser capazes de encontrar uma bijeção entre quaisquer dois conjuntos infinitos. Tendo *esse* pano de fundo é que o teorema de Cantor começa a parecer surpreendente.

[5]Essa não é a forma como Löwenheim teria formulado o teorema, mas é a formulação mais clara para nossos fins. Para uma discussão detalhada de formulação e demonstração deste teorema pelo próprio Löwenheim, veja (Badesa, 2004)

conjuntos de sentenças . Ele demonstrou que se uma coleção enumerável de sentenças de primeira-ordem tem um modelo infinito, então ela tem um modelo que é apenas enumerável. Este é o resultado que tipicamente recebe o nome de Teorema de Löwenheim-Skolem. Antes de prosseguirmos, é importante mencionarmos três, de algum modo mais refinadas, versões deste teorema[6].

Seja T uma coleção enumerável de sentenças de primeira-ordem e seja A um conjunto infinito. O Teorema Ascendente de Löwenheim-Skolem diz que se T possui algum modelo, então T possui um modelo cujo domínio tem o mesmo tamanho de A (de fato, podemos assumir sem perda de generalidade que o domínio deste segundo modelo é justamente A)[7].O Teorema Descendente de Löwenheim-Skolem diz que se **N** é um modelo de cardinalidade (infinita) κ e se λ é um cardinal infinito menor que κ, então **N** tem um submodelo de cardinalidade λ que satisfaz exatamente as mesmas sentenças que **N**[8].

Finalmente, o Teorema do Submodelo Transitivo diz que se nosso **N** inicial vem a ser um modelo *transitivo* de **ZF**, então ele contém um submodelo transitivo enumerável, que

[6]Para uma investigação completa da matemática que circunda os Teoremas de Löwenheim Skolem, veja (Ebbinghaus, 2007).

[7]Este teorema é originalmente atribuído a Tarski. Ele possui este nome por conta do fato que nos permite iniciarmos com um modelo enumerável de T e gerarmos modelos de cardinalidade arbitrariamente grande.

[8]Este teorema é, mais uma vez, atribuído a Tarski. Chama-se teorema *descendente* porque permite-nos começar com um modelo grande e depois gerarmos um submodelo (menor). Duas observações adicionais são aqui necessárias. Primeira, ao dizermos que nosso modelo novo, **N**, é um *submodelo* do nosso modelo original, **M**, queremos dizer que o domínio de **N** é um subconjunto do domínio de **M** e que os dois modelos concordam na interpretação das constantes, predicados, relações e funções na nossa linguagem – por exemplo, para todos $n_1, \ldots, n_m$ no domínio de **N** e qualquer R na nossa linguagem, **N** $\vDash R[n_1, \ldots, n_m] \Leftrightarrow$ **M** $\vDash R[n_1, \ldots, n_m]$. Segunda, como mencionado acima, o teorema descendente depende da suposição que nossa linguagem é enumerável. (Se nossa linguagem não é enumerável, então precisaríamos da suposição adicional de que λ é pelo menos tão grande quando o tamanho da nossa linguagem). Por conveniência de exposição, o restante desta entrada se limitará ao caso de linguagens enumeráveis – portanto, deste ponto em diante "linguagem" = "linguagem enumerável".

também satisfaz **ZF**[9].

Voltando, agora, para a versão original do teorema de Löwenheim-Skolem – aquele que diz que qualquer teoria que possua um modelo infinito também possui um modelo infinito enumerável. O Paradoxo de Skolem surge quando percebemos que os axiomas clássicos da teoria de conjuntos podem, eles próprios, serem formulados como uma coleção (enumerável) de sentenças de primeira-ordem. Se é o caso que esses axiomas possuem um modelo, então o teorema de Löwenheim-Skolem garante que eles possuem um modelo com domínio enumerável[10] . Isto, porém, parece deveras intrigante. De que modo podem os axiomas que demonstram o teorema de Cantor sobre a existência de conjuntos não-enumeráveis serem satisfeitos por um modelo que é, ele próprio, enumerável? Como é possível um modelo enumerável satisfazer a sentença de primeira-ordem que "diz que" há uma quantidade não-enumerável de coisas?

Essas questões podem ser tornadas mais concretas ao considerarmos um caso específico. Seja T uma axiomatização padrão, em primeira-ordem da teoria de conjuntos. (Por conveniência, esta entrada focará no caso em que T é **ZFC**, mas qualquer outra axiomatização padrão da teoria de conjuntos teria um resultado semelhante.) Assumindo que T possui um modelo, os teoremas de Löwenheim-Skolem garantem que T possui um modelo enumerável. Chame este modelo de **M**. Agora, como $T \vdash \exists x$ " x é enumerável", então

[9]De fato, há um único modelo transitivo e enumerável – o modelo chamado de modelo minimal, que satisfaz ZFC + $V = L$ e que é um submodelo de todo modelo transitivo de ZF. Se quisermos que nosso submodelo enumerável satisfaça exatamente as mesmas sentenças que o original N, então as coisas ficam um pouco mais complicadas. Em muitos casos podemos fazê-lo – por exemplo, quando N tem a forma V_γ para algum ordinal γ. Hamkins, Woodin e Button, porém, recentemente mostraram que há casos em que um modelo transitivo de ZF não possui submodelos transitivos enumeráveis que satisfaçam exatamente a mesma teoria. Para uma definição de transitividade – e uma discussão de seu significado filosófico no contexto do Paradoxo de Skolem – veja a seção 2.3.

[10]Devemos observar aqui que a suposição inicial de que nossos axiomas possuem qualquer modelo não é trivial – afinal, o segundo teorema de Incompletude garante que não podemos provar a existência de tais modelos de dentro de nossa axiomatização. Além disso, alguns dos resultados que analisaremos mais adiante nesta entrada envolvem suposições de existência de modelos ainda mais fortes – por exemplo, a existência de um modelo transitivo para nossos axiomas é estritamente mais forte que a de mera existência de um modelo para nossos axiomas. Agora, tendo destacado esses tipos de preocupação aqui, nós vamos suprimi-los do restante dessa entrada (então, daqui em diante, simplesmente assumiremos que nossos axiomas possuem modelos, modelos transitivos, etc.). Para uma discussão mais detalhada dos problemas filosóficos levantados pela possível não-existência de tais modelos, veja (Bays, 2007b). Para informações mais técnicas sobre a força de várias suposições de existência de modelos, veja (Ebbinghaus, 2007) e a seção 1 de (Bays, 2007a).

deve existir algum $\hat{m} \in$ **M** tal que **M** $\vDash$ "$\hat{m}$ é não-enumerável". Mas, como **M** é somente enumerável, há apenas uma quantidade enumerável de $m \in$ **M** tais que **M** $\vDash m \in \hat{m}$. Superficialmente, então, pode parecer que há uma contradição imediata: de uma perspectiva, $\hat{m}$ parece não-enumerável, enquanto de outra perspectiva, $\hat{m}$ é claramente enumerável.

Isto, portanto, nos dá uma formulação bastante simples do Paradoxo de Skolem. Antes de nos voltarmos à solução deste paradoxo, um ponto sobre a sua motivação deve ser levantado. De uma perspectiva, não há nada surpreendente no fato de que um modelo em particular falha ao capturar acuradamente toda característica da realidade da qual ele é modelo. Um modelo matemático de uma teoria física, por exemplo, pode conter apenas números reais e conjuntos de números reais, ainda que a teoria, ela própria seja sobre, digamos, partículas subatômicas e regiões do espaço-tempo. Semelhantemente, uma maquete do sistema solar reproduzirá corretamente certos aspectos do sistema solar, enquanto reproduzirá outros erroneamente. Assim, por exemplo, ela pode reproduzir os tamanhos relativos do planetas corretamente, ao passo que reproduz seus tamanhos absolutos (ou até seus tamanhos proporcionais) erroneamente; ou pode estar correta sobre o fato de que os planetas se movem ao redor do Sol, estando incorreta sobre os mecanismos deste movimento (por exemplo, os planetas não se movem ao redor do Sol porque um demonstrador gira uma alavanca!). Dados todos esses fatos, pode não ser claro por quê deveríamos esperar modelos de primeira-ordem de teoria de conjuntos capturem acuradamente a distinção entre conjuntos enumeráveis e não-enumeráveis. Logo, de início pode não ser claro por que deveríamos pensar que o Paradoxo de Skolem aparenta ser paradoxal.

Ainda que falaremos mais sobre esse tipo de problema (veja em especial as seções 2.1 e 3.1), algumas notas preliminares são apropriadas neste momento. Primeiro, é importante notar que há alguns conceitos de teoria de conjuntos que modelos de primeira-ordem capturam com bastante precisão. Como veremos na seção 3.1, modelos de primeira-ordem capturam noções de cardinalidade finita – por exemplo, "x é vazio", "x tem dois elementos", "x tem dezessete elementos", etc – muito bem[11] . Se nos permitirmos utilizar uma quantidade

[11] É preciso ter algum cuidado aqui. Há muitas maneiras de se entender o que significa para um modelo "capturar" uma noção da teoria de conjuntos; em alguns deles, mesmo noções finitas de cardinalidade não podem ser adequadamente capturadas. Por enquanto, a consideração importante é apenas de que há um modo de entender o que significa para um modelo "capturar" uma noção da teoria de conjuntos sobre a qual noções de cardinalidade finita podem ser capturadas, mas não a distinção enumerável/não-enumerável. Veja a seção 3.1 para um discussão detalhada das distinções relevantes.

infinita de fórmulas, então podemos também capturar a noção mais geral "x é infinito"[12].

Finalmente, se fixarmos nosso entendimento de pertencimento – isto é, se restringirmos nossa atenção aos modelos que utilizam a relação real de pertencimento para interpretarmos o símbolo "$\in$" – então podemos também capturar a noção geral de "x é finito"[13].

Dados todos estes pontos, o Paradoxo de Skolem mostra que a linha entre conjuntos enumeráveis e não-enumeráveis é, em um sentido bastante profundo, o primeiro lugar em que a nossa teoria de modelos perde a habilidade de capturar noções de cardinalidade. Esse fato ajuda a explicar por quê o Paradoxo de Skolem pode continuar aparentando ser paradoxal, mesmo após termos absorvido os pontos gerais sobre modelos e teoria de modelos que foram apresentados no penúltimo parágrafo. Em suma: é exatamente pelo fato de que conseguimos capturar tantas noções de cardinalidade que habitam, por assim dizer, abaixo da distinção enumerável/não-enumerável o que faz a nossa repentina falta de habilidade de capturar a distinção enumerável/não-enumerável ela própria a princípio tão surpreendente.

Segundo, o Paradoxo de Skolem não depende da axiomatização específica da teoria de conjuntos que porventura estejamos usando. *Qualquer* axiomatização de primeira-ordem da teoria de conjuntos pode ter os teoremas de Löwenheim-Skolem aplicadas a ela, logo toda axiomatização é sujeita ao Paradoxo de Skolem. Isso significa, em particular, que não podemos resolver o paradoxo apenas escolhendo uma nova axiomatização da teoria de conjuntos (ou adicionado alguns novos axiomas à axiomatização que já estamos usando). O fato de que o paradoxo de Skolem é, desse modo, *intrínseco* ao contexto de primeira-ordem – que ele é um fato *inescapável* sobre axiomatizações de primeira-ordem da teoria de conjuntos – é outra razão pela qual o Paradoxo de Skolem aparenta inicialmente ser tão intrigante.

Isto, portanto, nos dá um primeiro passo para a formulação do Paradoxo de Skolem. Na

[12]Em particular, seja $\phi_n(x)$ a formula que diz que x tem ao menos n elementos (para n um número natural). Então, para qualquer modelo **M** e qualquer elemento $m \in$ **M**:

$$\forall x(\mathbf{M} \vDash \phi_n[m]) \Leftrightarrow \{m' \in \mathbf{M} : \mathbf{M} \vDash m' \in \mathbf{M}\} \text{ é infinito.}$$

Então podemos utilizar todo o conjunto de ϕ_n's para capturar a noção de "x tem uma quantidade infinita de elementos".

[13]Qualquer modelo que obtenha corretamente pertencimento necessita possuir uma relação de pertencimento bem-fundada. Logo, ele também obterá a estrutura dos números naturais corretamente. Portanto, a formulação corriqueira de "x é finito" selecionará todos e apenas aqueles conjuntos que possuem um número finito de elementos em **M** – isto é, aqueles $m \in$ **M** tal que $\{m' \in \mathbf{M} : \mathbf{M} \vdash m' \in m\}$.

próxima seção, explicaremos por que essa versão simples do paradoxo não constitui uma contradição genuína, e veremos a formulações mais refinadas do paradoxo. Na seção 3, voltamo-nos para questões históricas e filosóficas. A seção 3.1 olha para o entendimento do próprio Skolem de seu paradoxo. Seções 3.2–3.4 olham para tentativas mais recentes de argumentar que, ainda que o paradoxo não constitua uma contradição matemática genuína, ele ainda nos diz algo filosoficamente importante sobre a natureza do entendimento da teoria de conjuntos.

2. Questões Matemáticas

Em uma introdução ao artigo de 1922, em que Thoralf Skolem apresentou o Paradoxo de Skolem pela primeira vez, Jean van Heijenoort escreve que o paradoxo "não é um paradoxo no sentido de uma antinomia … é uma propriedade nova e inesperada de sistemas formais."[14]. Esse comentário reflete o consenso geral sobre o Paradoxo de Skolem dentro da comunidade matemática. Quaisquer que sejam os problemas filosóficos que o paradoxo supostamente engendre, ele não constitui um problema *para a matemática*.

Para entender porque o paradoxo não constitui um problema para a matemática, precisamos levantar duas questões. Na formulação simples do paradoxo dada acima, notamos que há um $\hat{m} \in$ **M** específico tal que **M** $\vDash$ "m é não-enumerável". Literalmente, é claro, isso não está de todo correto. O que realmente queremos dizer aqui é que há uma fórmula de certo modo complicada na linguagem de teoria formal de conjuntos – uma fórmula que matemáticas às vezes acham conveniente abreviar pela expressão em português "x é não-enumerável" – e que **M** satisfaz essa fórmula particular em $\hat{m}$. Por conveniência, denotemos a fórmula relevante por "$\Omega(x)$". Então podemos reformular o fato mencionado acima dizendo que **M** $\vDash \Omega[\hat{m}]$ [15]. Nossas duas questões, portanto, são essas:

1. Por que é tão natural abreviar $\Omega(x)$ por "x é não-enumerável"? Por que, em particular, alguém iria pensar que o fato de **M** $\vDash \Omega[\hat{m}]$ implicaria $\hat{m}$ ser não-enumerável?

2. Por que o fato de **M** $\vDash \Omega[\hat{m}])$ não implica, de fato, $\hat{m}$ é não-enumerável?

[14] A introdução foi escrita para uma reimpressão do artigo em (Van Heijenoort, 1967). Veja pp. 290–291.

[15] Logo, $\Omega(x)$ é uma fórmula aberta tendo x como sua única variável livre e $\Omega[\hat{m}]$ é o resultado de avaliar essa fórmula em **M** sob a suposição de que x designa $\hat{m}$ (veja a nota de rodapé 1).

Com efeito, a primeira dessas questões indaga se o Paradoxo de Skolem é *simplesmente* um efeito colateral das nossas abreviações, efeito este que poderia desaparecer se o Paradoxo de Skolem fosse formulado de maneira mais cuidadosa e clara. Assumindo que ele não desapareceria, a segunda questão pede uma explicação mais detalhada de como o paradoxo pode ser de fato dissolvido.

2.1 A Aparência de Paradoxo

Há duas maneiras de se abordar a primeira questão. Por um lado, poderíamos começar com a fórmula $\Omega(x)$ e dar a essa fórmula o que podemos chamar de sua interpretação em "português comum". Essa é a interpretação em que "$\in$" se refere à relação real de pertencimento da teoria de conjuntos, em que "$\forall$" e "$\exists$" abrangem todo o universo (real) da teoria de conjuntos, e em que "$=$" e os conectivos proposicionais são interpretados de maneira usual[16]. Então para qualquer conjunto m, $\Omega(m)$ será verdadeira se, e somente se, m é não-enumerável[17]. Isso mostra que há pelo menos uma interpretação de $\Omega(x)$ em que esta fórmula realmente captura – ao menos de uma perspectiva extensional – o noção matemática comum de não-enumerabilidade. Portanto, há ao menos uma interpretação em que $\Omega(\hat{m})$ realmente diz que $\hat{m}$ é não-enumerável.

Por outro lado, poderíamos começar, não com a fórmula $\Omega(x)$, mas com a frase em português "x é não-enumerável". Se perguntada o que essa frase significa, uma teórica de conjuntos diria algo sobre a falta de bijeção entre x e os números naturais[18]. Se perguntada sobre a frase "é uma bijeção", ela passará a falar sobre coleções de pares ordenados satisfazendo determinadas propriedades desejáveis, e se perguntada sobre o termo "pares ordenados", ela dirá algo sobre os modos que alguém pode identificar pares ordenados e conjuntos particulares. Se ela leva o processo adiante por uma distância suficiente – e se ela economiza tempo usando símbolos por exemplo $\neg$ e $\exists y$ como abreviações para "não"

[16]Então "$\neg$" significa *não*, "$\vee$" significa *ou*, "$\wedge$" significa *e*, etc.

[17]Como anteriormente, $\Omega(x)$ é uma fórmula aberta com x sendo sua única variável livre; $\Omega(m)$ é o resultado de substituir as instâncias livres de x com um nome nome para m. Em particular, portanto, $\Omega(\hat{m})$ é obtida ao se trocar as instâncias relevantes de x com um nome para $\hat{m}$.

[18]Relembre aqui que uma bijeção entre os conjuntos x e yé apenas uma correspondência um-para-um entre os elementos de x e os elementos de y. Como vimos na seção 1, dizer que x é enumerável é apenas dizer que existe uma tal bijeção entre x e os números naturais; Dizer que x é não-enumerável é dizer que não existem tais bijeções.

e "existe um conjunto y tal que" – então ela obterá eventualmente uma explicação detalhada de "x é não-enumerável" que se parece exatamente com a fórmula $\Omega(x)$. Isto é, se compararmos apenas a sintaxe de sua explicação de "x é não-enumerável" com a sintaxe de $\Omega(x)$, então descobriremos que ambas as expressões contém exatamente os mesmos símbolos exatamente na mesma ordem[19]. Mais uma vez, por conseguinte, descobrimos que há uma semelhança real, ainda que de certo modo superficial, entre $\Omega(x)$ e "x é não-enumerável" – uma semelhança que se mantém mesmo depois de pararmos de utilizar "x é não-enumerável" como uma abreviação imediata de $\Omega(x)$ e uma semelhança que explica o porquê de até uma versão claramente formulada do Paradoxo de Skolem poder continuar a parecer, de algum modo, enigmática.

Estes são, assim, os dois modos de se pensar sobre a relação entre $\Omega(x)$ e "x é não-enumerável". Juntas, elas podem explicar por que é tão natural para matemáticas utilizar "x é contável" como uma abreviação de $\Omega(x)$ e (portanto) por que alguém pode estar inclinada a pensar que o fato de **M** $\vDash \Omega[\hat{m}]$ deveria implicar $\hat{m}$ é não-enumerável. Elas também nos remontam à segunda questão: Por que o fato de **M** $\vDash \Omega[\hat{m}]$ não implica, de fato, $\hat{m}$ é não-enumerável?

2.2 Uma Solução Genérica

Para responder a esta segunda questão, é útil começarmos por comparar a interpretação em português cotidiano de $\Omega(x)$ – aquela introduzida há três parágrafos e que realmente implica x ser não-enumerável – com a interpretação de $\Omega(x)$ na teoria de modelos dada por **M** e $\vDash$. Claramente, é esta última interpretação em teoria de modelos a mais relevanta para se entender o fato de **M** $\vDash \Omega[\hat{m}]$. Além disso, é apenas no caso em que essa interpretação em teoria de modelos seja intimamente ligada à interpretação em português comum – e então, derivadamente, à interpretação em português comum de "x é não-enumerável" – que teremos algum fundamento real para acreditarmos no fato de que **M** $\vDash \Omega[\hat{m}]$ deveria implicar $\hat{m}$ é não-enumerável.

Felizmente, mesmo uma descrição grosseira da interpretação em teoria de modelos é

[19]Devemos observar aqui que há muitas maneiras distintas de se explicar "x é não-enumerável", a depender de como decidimos "codificar" noções básicas tais como pares ordenados ou números naturais. A discussão acima assume que nossa teóric de conjuntos tomou as mesmas decisões que nós tomamos ao fomularmos originalmente $\Omega(x)$. Como qualquer formulação particular de $\Omega(x)$ corresponde a uma dessas explanações, essa suposição não envolve perda de generalidade.

suficiente para mostrar que não há tais "ligações íntimas". A interpretação em teoria de modelos é obtida ao se assumir que o significado de "$\in$" seja fixado pela interpretação da função interpretação de **M**, assumindo que os quantificadores em $\Omega(X)$ abranjam o domínio de **M**, e assumindo que o significado de "$=$" e dos conectivos proposicionais sejam fixados pelas cláusulas recursivas na definição de satisfação em primeira-ordem. Essa descrição salienta duas diferenças fundamentais entre a interpretação em teoria de modelos e a interpretação em português comum.

Em primeiro lugar, a interpretação em teoria de modelos entende que "$\in$" se refere a qualquer relação binária em **M** que por acaso seja capturada pela função interpretação de **M**; em contraste, a interpretação em português de $\Omega(x)$ entende que "$\in$" se refere à verdadeira relação de pertencimento em teoria de conjuntos. Não há razão, porém, para pensar que esses dois entendimentos concordam entre si. Podemos encontrar caso em que **M** $\models m_1 \in m_2$, apesar do fato de que nem m_1, nem m_2 são conjuntos (de fato, em se tratando de teoria de modelos, m_1 e m_2 poderiam ser ambos gatos, ou coelhos, ou ouriços, ou ...)[20]. Ademais, até quando todos os elementos de **M** são conjuntos, isto não nos dá garantia de que o entendimento de "$\in$" em teoria de modelos concordará com o entendimento de "$\in$" em português usual. Podemos encontrar um caso em que m_1 e m_2 são conjuntos genuínos e que **M** $\models m_1 \in m_2$, apesar do fato de que m_1 não seja elemento de m_2; de modo semelhante, podemos encontrar um caso em que **M** $\models m_1 \notin m_2$, apesar do fato de que m_1 seja realmente um elemento de m_2 (e num caso em que, novamente, m_1 e m_2 são conjuntos genuínos)[21].

Depois, a interpretação em teoria de modelos entende "$\exists y$" e "$\forall y$" abrangendo apenas o domínio de **M**, enquanto a interpretação em português usual entende que esses quantificadores abrangem todo o universo da teoria de conjuntos. Claramente, esses dois entendimentos são deveras distintos. Além disso, as diferenças em questão são fortemente relacionadas aos tipos de conjuntos envolvidos no Paradoxo de Skolem. Suponha, por exemplo, que **M** $\models$ "$\hat{m}$ é o conjunto dos números reais". Então um simples argumento por cardinalidade mostra que há $2^{\aleph_0}$ números reais, que não estão no domínio de **M** (e então, em particular, não estão em $\{m :$ **M** $\models m \in \hat{m}\}$). Logo, há uma diferença real entre o conjunto genuinamente não-

[20]Em geral, para quaisquer dois objetos, a e b, podemos encontrar um modelo enumerável **M** tal que 1.) **M** $\models$ ZFC e 2.) a e b são elementos de tal que **M** $\models a \in b$. Para mais detalhes sobre essa construção, veja a seção 2 de (Bays, 2007a)

[21]Novamente, veja a seção 2 de (Bays, 2007a) para mais detalhes sobre essas construções.

enumerável $\mathfrak{R}$ e o conjunto meramente enumerável $\{m : \mathbf{M} \vDash m \in \hat{m}\}$ – isto é, entre o verdadeiro conjunto de números reais e o conjunto de coisas que **M** meramente pensa ser números reais. Na interpretação em teoria de modelos de $\Omega(x)$, os quantificadores abrangem apenas o último conjunto, que é menor, enquanto que a interpretação em português, abrange a conjunto maior por completo. Semelhantemente, suponha que **M** $\vDash$ "x é infinito". Então podemos mostrar que há exatamente $2^{\aleph_0}$ bijeções $f : \omega \to \{m' \in \mathbf{M} : \mathbf{M} \vDash m' \in m\}$[22]. Não obstante, há no máximo uma quantidade enumerável destas bijeções no domínio de **M**. Portanto, apenas uma quantidade enumerável delas são "vistas" por $\exists x$ e $\forall x$ na interpretação em teoria de modelos de $\Omega(x)$, ainda que todas as $2^{\aleph_0}$ delas sejam "vistas" sob a interpretação em português usual.

Quando tomados em conjunto, esses resultados sugerem que o Paradoxo de Skolem pode simplesmente resultar de um colapso sub-reptício de duas interpretações distintas de $\Omega(x)$. Dado um modelo enumerável de ZFC, apenas a interpretação em teoria de modelos de $\Omega(x)$ permite-nos encontrar um elemento $\hat{m} \in$ **M** tal que **M** $\vDash \Omega[\hat{m}]$. No entanto, apenas a interpretação em português usual nos dá os verdadeiros fundamentos para pensarmos que $\Omega(\hat{m})$ implica $\hat{m}$ ser enumerável. Além disso, e como acabamos de ver, há bastantes diferenças entre a interpretação em teoria de modelos e aquela em português usual para que suspeitemos de qualquer colapso imediato entre ambas (mesmo que não soubéssemos que esse colapso eventualmente nos levaria longinquamente ao Paradoxo de Skolem). Particularmente, portanto, deveríamos resistir a qualquer tentativa de passar do fato de que **M** $\vDash \Omega[\hat{m}]$ para a afirmação de que $\hat{m}$ é não-enumerável.

Com efeito, essa análise trata o Paradoxo de Skolem como um caso imediato de equivocação. Há uma interpretação de $\Omega(\hat{m})$ na qual esta fórmula realmente implica $\hat{m}$ ser um conjunto não-enumerável; há uma outra – assaz distinta – garantindo que **M** $\vDash \Omega[\hat{m}]$; o Paradoxo de Skolem depende da confusão entre essas duas interpretações. À princípio, notar que essa confusão nos induz ao erro não deveria ser mais surpreendente que notar que o banco da praça não é o lugar mais adequado para depositar nosso dinheiro. De fato, o caso de teoria de modelos pode ser até ligeiramente pior do que o caso do banco; alguém poderia ter sorte e encontrar um banco de praça confiável para depositar seu dinheiro, en-

[22]Relembre que essas são precisamente as bijeções que demonstram que o conjunto $\{m' : \mathbf{M} \vDash m' \in m\}$ é enumerável; logo, essas são bijeções que, por uma razão ou outra, **M** falha em propriamente compreender (veja a seção 2.3 para um modo de como isso pode acontecer; veja as seções 3–5 de (Bays, 2007a) para mais detalhes e outras possibilidades).

quanto é um teorema imediato que se **M** é enumerável, então $\{m : \mathbf{M} \vDash m \in \hat{m}\}$ também é enumerável.

Isso nos dá, por conseguinte, uma solução bastante simples para o Paradoxo de Skolem. É uma solução que explica por que a maioria das matemáticas não considera o paradoxo muito preocupante além de ser deveras popular na literatura filosófica. Ela é, por exemplo, essencialmente a solução que o próprio Skolem deu em 1922 (Skolem 1922), e variações desta solução apareceram em discussões mais recentes do paradoxo (Resnik 1966; Myhill 1967; Hart 1970; McIntosh 1979; Benacerraf 1985; Shapiro 1991; Giaquinto 2002). Também aparece em vários livros-texto introdutórios recentes (Shoenfield 1967; Kleene 1967; Fraenkel et al. 1984; Ebbinghaus et al. 1994; van Dalen 1997).

2.3 Submodelos Transitivos

Antes de examinarmos alguns dos problemas mais puramente filosóficos relacionados ao Paradoxo de Skolem, há algumas considerações adicionais sobre a matemática do paradoxo. Em primeiro lugar, para se ter uma intuição de como as diferenças entre as interpretações em teoria de modelos e português usual de fato dão origem ao Paradoxo de Skolem, é importante olhar para essas diferenças por meio de uma versão um tanto mais refinada do paradoxo. Dizemos que X é transitivo se todo elemento de X é um conjunto e todo elemento de um elemento de X também pertence a X (assim, $y \in x \in X \Rightarrow y \in X$).Dizemos que um modelo para a linguagem de teoria de conjuntos é transitivo se o domínio do modelo é um conjunto transitivo e sua relação de "pertencimento" é justamente a relação de pertencimento verdadeira, restrita ao domínio do modelo (assim, para todo $m_1, m_2 \in \mathbf{M}, m_1 \in m_2 \iff \mathbf{M} \vDash m_1 \in m_2$). Logo, como observado na seção 1, o Teorema do Submodelo Transitivo diz que se iniciarmos com qualquer modelo transitivo de ZFC, então podemos encontrar um modelo transitivo cujo domínio é enumerável (de fato, podemos assumir que esse modelo enumerável é submodelo do modelo com o qual iniciamos).

Suponha, então, que **M** é um modelo transitivo enumerável de ZFC. Isso tem dois efeitos na análise do Paradoxo de Skolem dada na seção anterior. Primeiro, garante que as interpretações em teoria de modelos e português usual de $\Omega(x)$ *coincidam* ao interpretar "$\in$": para $m_1, m_2 \in \mathbf{M}$, $\mathbf{M} \vDash m_1 \in m_2$ se, e somente se, m_1 é realmente um elemento de m_2[23]. Nesse caso, portanto, a explicação do Paradoxo de Skolem tem que envolver a interpretação dos quantificadores. Segundo, o fato de **M** ser transitivo garante que **M** acerte mais do que apenas a noção de pertencimento. Em particular, se f e m habitarem o domínio de **M**, então **M** $\vDash$ "$f : \omega \to m$ é uma bijeção" se, e somente se, f realmente for uma bijeção entre os números naturais e m[24].

Em conjunto, esses fatos ajudam-nos a isolar o que realmente acontece na versão do submodelo transitivo do Paradoxo de Skolem. Considere mais uma vez a fórmula que chamamos $\Omega(x)$. Esta fórmula tem a forma:

$$\Omega(x) \equiv \neg\exists f \text{ “}f : \omega \to x \text{ é uma bijeção”}$$

Sob a interpretação em português comum, essa fórmula diz que o universo da teoria de conjuntos não contém bijeções entre os números naturais e x. Em particular, $\Omega(\hat{m})$ diz que não há bijeções entre os números naturais e $\hat{m}$. Em contraste, a interpretação em teoria de modelos de $\Omega(\hat{m})$ – aquela que é relevante para o fato de **M** $\vDash \Omega[\hat{m}]$ – diz apenas que o domínio de **M** não contém bijeções entre os números naturais e $\hat{m}$[25]. Claramente, essas duas interpretações têm potencial para serem distinguidas.

Essa versão do submodelo transitivo do paradoxo foi amplamente discutida na literatura (McIntosh, 1979; Benacerraf and Wright, 1985b; McCarty and Tennant, 1987). De fato, vários autores sugeriram que a transitividade talvez seja necessária para formular uma versão

[23]Na verdade, nós obtemos algo um pouco mais forte que isso. Se $m_2 \in$ **M**, então para *qualquer* conjunto m_1, $m_1 \in m_2 \Leftrightarrow$ **M** $\vDash m_1 \in m_2$. Assim, modelos transitivos não apenas determina pertencimento corretamente, eles também capturam completamente a noção "x é elemento de m", em que o próprio m é elemento do modelo relevante.

[24]Aqui, a expressão "$f : \omega \to m$ é uma bijeção" é a abreviação (natural) para uma formula muito mais longa de linguagem da teoria de conjuntos formal. A relação entre essa fórmula e a expressão em Português comum "$f : \omega \to m$ é uma bijeção" é similar àquela entre $\Omega(x)$ e "x é não-enumerável". Observe também que, por **M** ser transitivo, $m = \{m' \in$ **M** : **M** $\vDash m' \in m\}$. Logo não há diferença entre dizermos que "$f : \omega \to \{m' \in$ **M** : **M** $\vDash m' \in m\}$ é uma bijeção" e dizermos que "$f : \omega \to m$ é uma bijeção". Por conveniência, portanto, utilizaremos esta última notação pelo restante desse exemplo.

[25]Observe aqui que é apenas por **M** ser transitivo que podemos com confiança construir a interpretação em teoria de modelos de $\Omega(x)$ como se ela dissesse algo sobre bijeções. (Este é o significado da segunda consideração no parágrafo anterior). Se **M** não fosse transitivo, então não haveria nenhuma razão geral para pensarmos que a interpretação em teoria de modelos do porção de $\Omega(x)$ após o quantificador existencial inicial tenha algo a ver com bijeções.

filosoficamente significativa do paradoxo (Benacerraf and Wright, 1985b; Button and Walsh, 2018). Ver McCarty and Tennant (1987) para algumas objeções a esta posição.

2.4 ZFC, Conjuntos Potência e Números Reais

A análise das seções 2.2–2.3 explica em termos gerais de que maneira um modelo enumerável pode satisfazer uma fórmula como $\Omega(x)$ em um elemento particular. Ela pode, porém, ainda deixar uma questão óbvia sem resposta: como é possível um modelo enumerável de ZFC satisfazer tal fórmula? Dado que um modelo arbitrário pode interpretar uma fórmula tal como $\Omega(x)$ de maneira peculiar, de que modo um modelo poderia satisfazer todos os axiomas da teoria de conjuntos e ainda manter essa interpretação peculiar? Não deveria o fato de que **M** satisfaz ZFC garantir que **M** também captura as noções básicas da teoria de conjuntos, tais como enumerabilidade e não-enumerabilidade, corretamente?

A resposta breve para tais questões é esta: modelos enumeráveis "interpretam tão mal" os axiomas da teoria de conjuntos, quanto eles interpretam mal a fórmula $\Omega(x)$. Por agora, permaneçamos com a suposição de que **M** é transitivo e consideremos o conjunto potência[26]:

$$\forall x \exists y \forall z [z \subseteq x \iff z \in y]$$

Em sua interpretação em português usual, este axioma diz que todo conjunto possui um conjunto potência – um conjunto que contém todos e apenas os subconjuntos do conjunto com o qual iniciamos[27]. Em sua interpretação em teoria de modelos, no entanto, o axioma diz algo muito mais fraco. Para todo $x \in$ **M**, o axioma garante que podemos encontrar um $y \in$ **M** que contém exatamente aqueles subconjuntos de x que também habitam **M** (portanto $y = \{z : z \subseteq x \wedge z \in \mathbf{M}\}$). Contudo, se x for infinito, então a maioria dos subconjuntos de x não habitarão o domínio de **M** (afinal, há $2^{\aleph_0}$ subconjuntos de x, enquanto o domínio de **M** é apenas enumerável). Logo, o conjunto y gerado pela interpretação em teoria de modelos do axioma do conjunto potência será muito menor do que o verdadeiro conjunto potência de

[26] Ao formularmos esse axioma, utilizamos "$\subseteq$" como uma abreviação para uma expressão um pouco mais longa na linguagem oficial da teoria de conjuntos. Como a relação de subconjunto é absoluta para modelos transitivos – isto é, como modelos transitivos capturam corretamente a relação "x é subconjunto de y" – essa abreviação é inócua.

[27] Então se $X = \{1, 2, 3\}$, o conjunto potência de X é $\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}$

x (Fraenkel et al., 1973; McCarty and Tennant, 1987; Shapiro, 1991; Hallett, 1994; Giaquinto, 2002; Bays, 2007a).

Neste caso, portanto, diferenças nos modos como a interpretação em teoria de modelos e aquela em português comum lidam com o quantificador $\forall z$ inicial – e, em particular, diferenças relacionadas a quais subconjuntos de x são "vistos" por este quantificador – explica como um modelo enumerável pode satisfazer um axioma que "supostamente" geraria um conjunto não-enumerável. E esse tipo de fenômeno é deveras geral. Em (Resnik, 1966), Michael Resnik observa esse fenômeno por meio do caso dos números reais. Como anteriormente, assuma que **M** é enumerável e que **M** é um modelo transitivo enumerável de ZFC[28]. Então haverá um $r \in$ **M** particular tal que, a menos de algumas abreviações,

$$\mathbf{M} \models \text{“}r \text{ é o conjunto dos números reais.”}$$

Resnik nota que, ainda que **M** satisfaça esta fórmula, R não contém de fato todos os números reais – R contém apenas aqueles números reais que por acaso habitam o domínio de **M**[29]. Assim, o mero fato de R ser enumerável não gera, em nenhum sentido interessante, uma situação paradoxal na qual o conjunto de todos os números reais é também enumerável.

Tomados em conjunto, esses exemplos salientam um fato crucial: as "más interpretações" que explicam de que maneira modelos enumeráveis podem satisfazer uma sentença como $\Omega(\hat{m})$ são de fato bastante sistemáticas. Elas também explicam como esses modelos conseguem satisfazer sentenças tais como "r é o conjunto dos números reais" ou "y é o conjunto potência de ω"; e eles explicam até como esses modelos conseguem satisfazer os axiomas da teoria de conjuntos (por exemplo, o axioma do conjunto potência). Quando se reúnem suficientes interpretações errôneas, elas em conjunto explicam de que maneira é possível um modelo enumerável satisfazer tanto os axiomas da teoria de conjuntos quanto,

[28]Por razões dialéticas, o argumento original de Resnik era focado no sistema de Wang de teoria de conjuntos, ao invés de ZFC. Mas a consideração básica é transferível para o contexto de ZFC, e, para nossos propósitos, é mais sensato discuti-la aqui.

[29]Como no caso do conjunto potência, essa análise se baseia no modo como **M** interpreta um quantificador universal particular. Como **M** é transitivo, **M** captura corretamente os números reais: Se $m \in$ **M**, então **M** $\models$ " m é um numero real" se, e somente se m realmente é um número real. Considere, então, a sentença: $\forall x[x \in R \leftrightarrow$ " x é um numero real"$]$. Na sua interpretação em Português comum, essa sentença diz que R contém todos os números reais. Na sua interpretação em teoria de modelos, no entanto, a sentença diz apenas que R contém todos os números reais *que por acaso habitam* **M**, como se esses fossem os únicos números reais "vistos" pelo quantificador inicial $\forall x$ na sentença.

ao mesmo tempo, manter a interpretação peculiar de $\Omega(x)$ que discutimos nas seções 2.2–2.3. Ao final, portanto, enquanto o teorema de Löwenheim-Skolem pode ainda ser um fato interessante tecnicamente – "uma propriedade nova e inesperada de sistemas formais", nas palavras de van Heijenoort –, o Paradoxo de Skolem não mais deveria aparentar tão paradoxal.

2.5 Quatro considerações finais

Finalizaremos essa discussão da parte matemática do Paradoxo de Skolem com quatro considerações finais. Primeiro, a discussão em 2.3–2.4 teve seu foco no caso do submodelo transitivo do Paradoxo de Skolem. Esse caso é relativamente simples de analisar, e (portanto) é o caso mais discutido na literatura. Mas ele também pode ser um tanto enganoso. Uma boa parte da análise de 2.3–2.4 se baseia no fato de que modelos transitivos capturam muitas coisas "corretamente" sobre o universo da teoria de conjuntos (pertencimento, bijeções, números reais, etc.). Mais importante, se **M** é transitivo e $m \in$ **M**, então $m \in \{m' \in \mathbf{M} : \mathbf{M} \models m' \in m\}$.

Contudo, se **M** não é transitivo, então quase tudo isso desmorona. Bays argumenta que há versões do Paradoxo de Skolem que se baseiam apenas no modo como certos modelos não-transitivos interpretam alguns poucos exemplos específicos da relação de pertencimento em $\Omega(x)$ (Bays, 2007a, seções 4–5). Considerações similares poderiam ser transferidas para a nossa discussão de conjuntos potência e números reais na seção 2.4. Podemos, por exemplo, encontrar um modelo enumerável de ZFC que contém todo o conjunto dos números reais como elemento – o modelo permanece enumerável apenas porque $\mathfrak{R} \neq \{m : \mathbf{M} \models m \in \mathfrak{R}\}$ (Benacerraf and Wright, 1985b; Bays, 2007a, seção 1). Em suma, embora a explicação genérica do Paradoxo de Skolem dada na seção 2.2 – aquela que simplesmente observa que há algumas diferenças entre as interpretações em teoria de modelos e em português comum de $\Omega(x)$ e então atribui o Paradoxo de Skolem a algum tipo de equivocação entre elas – se sustente quando passamos para modelos não-transitivos, nas análises mais detalhadas de 2.3–2.4 tudo se desfaz. No caso geral não-transitivo, portanto, a análise da seção 2.2 pode ser o melhor que podemos fazer ao dar uma explicação para o Paradoxo de Skolem. (O que não quer dizer que não podemos dar explicações mais detalhadas no contexto de qualquer modelo não-transitivo particular).

Isso nos leva a uma segunda consideração. O Paradoxo de Skolem depende crucialmente do fato que estamos usando uma axiomatização da teoria de conjuntos em lógica de primeira-ordem. Mais precisamente, depende do fato que estamos usando teoria de modelos de primeira-ordem para interpretar essa axiomatização. Em 1930, Zermelo demonstrou que modelos (de segunda-ordem) de ZFC de segunda-ordem calculam cardinalidades e conjuntos potência corretamente[30]. Em particular, assim, se **M** é modelo de ZFC de segunda-ordem e se $\hat{m} \in$ **M**, então **M** $\vDash$ "$\hat{m}$ é não-enumerável" se, e somente se, $\{m :$ **M** $\vDash m \in \hat{m}\}$ é realmente não-enumerável. Disto decorre que o Paradoxo de Skolem não surge no contexto de segunda-ordem (Zermelo, 1930; Shapiro, 1991).

Esse segundo ponto mostra que o Paradoxo de Skolem se esvai se nossa lógica é suficientemente forte. O terceiro ponto mostra que enfraquecer nossa lógica tem um efeito semelhante. Em (McCarty and Tennant, 1987), Tennant e McCarty mostram que as demonstrações tradicionais do teorema de Löwenheim-Skolem falham em teoria de conjuntos construtivista, além de argumentarem que o teorema por si mesmo é provavelmente inválido construtivamente[31]. Isso significa que não há meios de se gerar o Paradoxo de Skolem dentro do arcabouço da matemática construtivista. Para construtivistas, portanto, assim como para aquelas pessoas que desejam fomentar axiomatizações de segunda-ordem da teoria de conjuntos, o Paradoxo de Skolem não surge.

Juntas, as últimas duas considerações salientam exatamente quão central é a lógica de primeira-ordem para o Paradoxo de Skolem. Do ponto de vista da matemática, isso não deveria ser tão surpreendente. Lindström mostrou que o teorema de Löwenheim-Skolem cumpre um papel fundamental na caracterização da própria lógica de primeira-ordem (Lindström, 1966, 1969; Ebbinghaus, 2007). Dado isso, não deveria uma surpresa que o enigma que se associa mais intimamente com esses teoremas também se mostra ligado fortemente às

[30]Isto é, se **M** é um modelo de ZFC de segunda-ordem, então **M** $\vDash$ "m tem cardinalidade κ" se, e somente se, o conjunto $\{m' \in$ **M** $:$ **M** $\vDash m' \in m\}$ de fato possuir cardinalidade κ. Um fato semelhante, embora ligeiramente mais complicado, ocorre no caso do conjunto potência. Devemos observar aqui que o argumento de Zermelo assume que estamos utilizando modelos padrão para a lógica de segunda-ordem. O argumento falha se admitirmos o uso de modelos de Henkin de primeira-ordem para interpretar o formalismo de segunda-ordem. Para ler mais sobre a diferença entre modelos padrão e modelos de Henkin, veja as seções 2 e 3 na entrada sobre lógica de segunda-ordem e de ordens superiores.

[31]Mais precisamente, eles demonstram que todas as formulações habituais do teorema de Löwenheim-Skolem são independentes de uma certa forme forte de teoria de conjuntos intuicionista e mostram que essas formulações são imediatamente falsificadas por princípios que muitas construtivistas desejariam aceitar. Obviamente, isso não exclui a possibilidade de que alguma variação incomum do teorema venha a ser demonstrável em alguma forma não-standard de teoria de conjuntos "construtivista"; no momento, porém, esse cenário parece distante. Veja pp. 35–36 do artigo de Tennant e McCarty para a visão dos próprios autores sobre o assunto.

peculiaridades da situação em primeira-ordem. Ainda que o paradoxo, como já vimos, não configure uma contradição matemática imediata, ele nos ajuda a entender a natureza e os limites da lógica clássica de primeira-ordem.

Isso nos leva a uma consideração final. A discussão acima explica por quê aquelas de nós que desejam assumir uma atitude genuinamente realista em relação à linguagem da teoria de conjuntos – por exemplo, aquelas de nós que não sentem desconfortos quanto a expressões tais como "a interpretação em português comum de $\Omega(x)$" – deveriam se manter inabaladas pelo Paradoxo de Skolem. É importante enfatizar que esta análise também explica por que o Paradoxo de Skolem não introduz contradições em várias formas de se axiomatizar a teoria de conjuntos, mesmo quando essas axiomatizações são elas mesmas entendidas formalmente ou em teoria de modelos. De um ponto de vista da teoria da demonstração, por exemplo, há uma diferença entre quantificação não-relativizada e quantificação que foi explicitamente relativizada para uma fórmula na nossa linguagem (em que esta fórmula é tal que, intuitivamente, atua para "determinar" o domínio do modelo enumerável de ZFC). Então, não há razões *a priori* para se pensar que uma sentença com quantificadores não-relativizados conflitará com a contraparte completamente relativizada daquela sentença [32]. De modo semelhante, do ponto de vista da teoria de modelos, há uma diferença entre quantificadores que abrangem todo o domínio de um modelo e aqueles que abrangem apenas os "elementos" de um elemento particular do modelo (em que, mais uma vez, este elemento é tal que o modelo maior "pensa" ser um modelo de ZFC). Portanto, ainda que o realismo ingênuo das seções 2.1–2.4 seja útil para fins de exposição, ele não é essencial para a análise subjacente do Paradoxo de Skolem.

3. Problemas Filosóficos

A seção anterior explicou por que o Paradoxo de Skolem não constitui um problema para a matemática. Isto, é claro, não evitou que filósofas argumentassem que o paradoxo constitui um problema para a filosofia. Nesta seção, exploraremos diversas tentativas de se derivar conclusões filosóficas a partir da matemática envolvida no Paradoxo de Skolem. Antes de fazê-lo, porém, são necessárias duas observações cautelares. Primeira, muitas

[32] Por exemplo, não há conflito automático entre uma sentença da forma $\exists x \phi(x)$ e outra da forma $\neg\exists x[\psi(x) \wedge \phi(x)]$. Com efeito, essa consideração é uma versão mais formal da consideração sobre quantificação feita na seção 2.2.

das discussões mais provocativas do Paradoxo de Skolem são bastante breves – se resumindo a pouco mais do que comentários sugestivos feitos de passagem. Logo, muitas das discussões desses comentários terão de ser de algum modo conjeturais. Segunda, muitas discussões críticas do Paradoxo de Skolem focaram apenas em cuidadosamente destrinchar a matemática do paradoxo e depois explicar o por quê de o paradoxo não constituir uma contradição matemática genuína. Como esse material já foi coberto na seção 2, nós não diremos mais nada sobre esses problemas nessa seção.

3.1 Visões de Skolem

No artigo de 1922 em que Skolem originalmente apresentou o Paradoxo de Skolem, ele usou o paradoxo para argumentar a favor de duas conclusões filosóficas: que a teoria de conjuntos não funciona como uma "fundamentação para a matemática" e que axiomatizar a teoria de conjuntos leva a uma "relatividade das noções em teoria de conjuntos" (Skolem, 1922). Essas afirmações, e os argumentos de Skolem a favor delas, atraíram considerável atenção na literatura. Infelizmente, o artigo de Skolem é bastante condensado, sendo pois difícil de determinar a que exatamente se resumem essas afirmações. Atualmente, há três interpretações do artigo de Skolem que possuem algum valor na literatura filosófica.

Começaremos com a afirmação de Skolem que a axiomatização da teoria de conjuntos leva a uma relativização das noções da teoria de conjuntos. Uma maneira de se entender essa afirmação é vê-la diante de um cenário que podemos chamar de uma concepção algébrica ou de teoria de modelos de axiomatização. Nessa concepção os axiomas da teoria de conjuntos caracterizam – ou talvez até definam – noções básicas da teoria de conjuntos, tais como conjunto, pertencimento, e universo da teoria conjuntos. Assim, um universo da teoria de conjuntos é simplesmente um modelo para os axiomas da teoria de conjuntos, um conjunto é simplesmente um elemento em algum universo da teoria de conjuntos e pertencimento refere-se apenas a qualquer relação binária que um universo em particular utilize para interpretar o símbolo "$\in$". Nessa concepção de axiomatização, portanto, os axiomas da teoria de conjuntos não deveriam ser encarados como tentativas de descrever – ou ainda descrever parcialmente – algum "modelo pretendido" da teoria de conjunto previamente dado; pelo contrário, os modelos pretendidos da teoria de conjuntos são simplesmente aqueles modelos que por acaso satisfazem nossas coleções iniciais de axiomas da teoria de conjuntos[33].

Deveríamos enfatizar aqui que essa concepção algébrica da axiomatização era bastante familiar para matemáticas trabalhando à época em que Skolem escreveu seu artigo de 1922. O próprio Skolem foi treinado na escola algébrica de lógica de Schröder, e esta seria, pois, a via natural para ele pensar sobre axiomas. Até pessoas que não foram treinadas na escola de Schröder tinham familiaridade com essa concepção. É esta concepção que está por trás da famosa axiomatização da geometria proposta por Hilbert (sobre a qual Hilbert supostamente afirmou, um tanto notavelmente, que podemos substituir pontos, linhas e planos por mesas, cadeiras e canecas de cerveja, desde que os últimos objetos mantenham os tipos corretos de relações). É também a concepção que está por trás dos resultados do século XIX que afirmam que à aritmética e à análise podem ser dadas axiomatizações categoriais (de segunda-ordem). Finalmente, e mais importante, é esta a concepção que Skolem atribuiu a Zermelo no mesmo artigo que discutimos neste momento e, por conseguinte, Skolem está preocupado em criticar a concepção dos axiomas de Zermelo[34].

Dada essa concepção algébrica de axiomatização, então, Skolem considera os teoremas de Löwenheim-Skolem para argumentar que falta aos axiomas da teoria de conjuntos recursos para assegurar a noção de não-enumerabilidade. Dada qualquer axiomatização da teoria de conjuntos e qualquer fórmula $\Omega(x)$ que deveria capturar a noção de não-enumerabilidade, os teoremas de Löwenheim-Skolem mostram que podemos encontrar um modelo enumerável **M** que satisfaz nossos axiomas. Como na seção 1, portanto, podemos encontrar um elemento $\hat{m} \in$ **M** tal que **M** $\vDash \Omega(\hat{m})$, mas $\{m :$ **M** $\vDash m \in \hat{m}\}$ é apenas enumerável. Assim, desde que as noções básicas de teoria de conjuntos sejam caracterizadas apenas olhando para a teoria de modelos das axiomatizações em primeira-ordem da

[33]Em termos atuais, essa é a concepção de axiomas sobre a qual se baseia uma boa parte da álgebra moderna. Os axiomas da teoria de grupos, por exemplo, caracterizam o que significa uma estrutura matemática ser considerada um grupo, o que é necessário para que um elemento em particular dessa estrutura seja a identidade e o que é considerado o inverso de um elemento particular. Mas não existe um grupo *ur* (primitivo) que os axiomas tentam descrever – há apenas uma classe toda de grupos (igualmente pretendidos) caracterizados pelos axiomas.

[34]Veja pp. 295–296 de (Skolem, 1922) para a caracterização de Skolem dos axiomas de Zermelo. Devemos observar que é uma questão em aberto se Skolem de fato entendeu Zermelo nesse assunto. Em seus escritos mais tardios – incluindo naqueles em que ele diretamente responde ao Paradoxo de Skolem – Zermelo claramente aplica uma concepção algébrica de axiomas (veja, por exemplo, (Zermelo, 1930)). Não é claro, no entanto, se deveríamos aplicar essa concepção a uma leitura da axiomatização original de 1908, à qual Skolem respondia. Veja (Taylor, 1993) e (Ebbinghaus, 2003) para mais sobre essa questão interpretativa.

teoria de conjuntos, muitas dessas noções – e, em particular, as noções de enumerabilidade e não-enumerabilidade – tornar-se-ão inevitavelmente relativas[35].

Isto, então, fornece o conteúdo para a afirmação de Skolem de que axiomatizar a teoria de conjuntos levaria a uma relativização das noções da teoria de conjuntos. É importante aqui distinguir essa afirmação de uma afirmação mais trivial que pode ser imaginada como feita por Skolem. De um ponto de vista, a concepção algébrica de axiomatização leva a uma forma óbvia de relatividade: os elementos considerados conjuntos em um modelo podem não o ser em outro modelo, a relação de pertencimento de um modelo pode ser diferente daquela em outro modelo, e esta última diferença nas relações de pertencimento pode ocorrer até se acontecer de dois modelos partilharem o mesmo domínio. Nessa noção trivial de relatividade, portanto, quase tudo se torna relativo, mesmo noções simples como "x é o conjunto vazio" ou "x é um conjunto unitário". Afinal, um objeto poderia ser um "conjunto unitário" em um modelo, e um "conjunto diádico" em outro conjunto, ou poderia ser "o conjunto vazio" em um modelo e ser omitido completamente do domínio de outro modelo.

É importante enfatizar que a noção de relatividade do próprio Skolem é mais sofisticada do que isso. Permitamos que o elemento específico que atua como "o conjunto vazio" não permanecerá constante quando nos movemos de um modelo da teoria de conjuntos para outro – em que o conjunto vazio do primeiro modelo torne-se, talvez, um conjunto unitário no segundo. De qualquer modo, podemos ainda usar uma fórmula na linguagem da teoria de conjuntos que capture a noção de "x é o conjunto vazio" de uma maneira essencialmente absoluta. Em qualquer modelo dos nossos axiomas, um elemento $\hat{m} \in$ **M** satisfará a fórmula aberta "$\forall y \; y \notin x$" se, e somente se, o conjunto $\{m :$ **M** $\vDash m \in \hat{m}\}$ é realmente vazio. Portanto, há ao menos um sentido em que ainda podemos capturar a noção "x é o conjunto vazio" de dentro do arcabouço algébrico. E essa consideração estende-se mais amplamente – um argumento semelhante poderia ser aplicado a noções como "x é um conjunto unitário" ou "x tem dezessete elementos". Até na concepção algébrica de axiomatização, portanto, há ainda algumas noções da teoria de conjuntos que podemos fixar com grande precisão. O que os teoremas de Löwenheim-Skolem mostram é que, não importa quão ricos sejam os nossos axiomas da teoria de conjuntos (de primeira-ordem), não é possível fixar com esse tipo de técnica a noção "x é não-enumerável". Esse é o resultado por trás de todo

[35] Skolem também sugere, sem demonstrar, que as noções de sequência finita e sequência simplesmente infinita também acabarão sendo relativas. De fato, Skolem estava certo sobre isso, como mostra uma simples aplicação do teorema de compacidade.

discurso de Skolem sobre “relatividade” e um resultado que destaca uma fraqueza genuína na abordagem algébrica para axiomatização da teoria de conjuntos[36].

Em suma, então, a vantagem dessa discussão é esta: se admitirmos uma abordagem puramente algébrica para os axiomas da teoria de conjuntos, então muitas noções básicas da teoria de conjuntos – incluindo as noções de enumerabilidade e não-enumerabilidade – se tornarão relativas. Nas palavras de Skolem: “axiomatizar a teoria de conjuntos leva a uma relatividade das noções da teoria de conjuntos, e essa relatividade está inseparavelmente ligada a toda axiomatização minuciosa e completa” (Skolem, 1922, p. 296). Obviamente, isso ainda deixa aberta a questão de se essas noções são, por assim dizer, absolutamente relativas – de se há algum outro modo não algébrico e não minucioso ou completo de se entender nossos axiomas que não nos levaria ao tipo de relatividade que temos discutido até então. É quando nos voltamos para esta última questão que as diversas interpretações do artigo de Skolem começam a se desfazer.

A interpretação mais tradicional do artigo vê Skolem como a preparar um ataque à teoria de conjuntos. Skolem inicia o artigo notando que os paradoxos clássicos da teoria de conjuntos deveriam nos levar a sermos céticas a entendimentos informais da teoria de conjuntos – do “raciocínio ingênuo com conjuntos”, para utilizar a expressão do próprio Skolem. Dado isto, nossa única opção de fato é retomar a algum tipo de teoria de conjuntos axiomatizada e o único modo respeitável de se entender os axiomas é algebricamente (já que entendê-los intuitivamente resultaria em retornamos à ingenuidade previamente descreditada). O Paradoxo de Skolem, porém, mostra que as noções da teoria de conjuntos são relativas à concepção algébrica de axiomatização. Logo, essas noções são de fato relativas. Brevemente: os paradoxos clássicos mostram que a concepção algébrica de teoria de conjuntos é a melhor concepção que temos e portanto o Paradoxo de Skolem mostra que a noções da

[36]Uma consideração semelhante se aplica ao exemplo de teoria de grupos na nota 31. Claramente, o objeto particular considerado a “identidade” variará quando passamos de grupo a grupo, assim como o escopo dos nossos quantificadores. Mas nós ainda conseguimos utilizar fórmulas na nossa linguagem para capturar noções de teoria de grupos – por exemplo, que um elemento habita no centro de um grupo ou ainda que um elemento possui ordem 17. E se expandirmos nossa linguagem, então podemos capturar ainda mais noções – por exemplo, que o predicado “P” separa um subgrupo normal de um grupo maior que estamos estudando. Claramente, não conseguimos fazer isso para toda noção da teoria de grupos – não existe uma fórmula que capture exatamente a noção “x habita o subgrupo gerado por y e z”. Mas o fato de que não podemos ter sucesso nesse caso é um teorema matemática interessante – não é algo que apenas se segue trivialmente do natureza algébrica dos nossos axiomas. O que Skolem demonstrou, com efeito, é que a noção “x é enumerável” da teoria de conjuntos é mais próxima da noção “x é um subgrupo normal gerado por y e z” do que da noção “x possui ordem 17”.

teoria de conjuntos são inevitavelmente relativas. Essa leitura tradicional é bastante prevalente na concepção popular; variações dela são discutidas em (Hart, 1970; McIntosh, 1979; Muller, 2005; Bellotti, 2006).

A segunda interpretação foca na afirmação de Skolem de que a teoria de conjuntos não fornece uma fundamentação adequada para a matemática. Em particular, Skolem pensa que à teoria de conjuntos faltam os recursos para fornecer uma fundamentação para a aritmética comum – em sua visão, a aritmética é "límpida, natural e não é passível de questionamentos", enquanto a teoria de conjuntos é bem mais problemática. Para mostrar que ela é problemática, Skolem recorre a diferentes maneiras de se interpretar teoria de conjuntos – teoria de conjuntos ingênua, teoria de conjuntos axiomatizada construída com teoria da prova, teoria de conjuntos construída algebricamente, etc.– e ele argumenta que cada um desses entendimentos da teoria de conjuntos é inadequado para fins de fundamentação. Nessa leitura, o Paradoxo de Skolem assume um papel modesto em todo o argumento de Skolem. Ele atua para destacar alguns problemas existentes em uma concepção particular de teoria de conjuntos (a concepção algébrica), mas não atua nos argumentos de Skolem contrários a outras concepções de teoria de conjuntos. Além disso, esses outros argumentos não mostram – ou até afirmam mostrar – que as várias concepções não-algébricas da teoria de conjuntos levam a nenhum tipo de relatividade (ainda que elas tenham, é claro, outros problemas que as tornam inadequadas para fins fundacionalistas)[37]. Versões dessa leitura fundacionalista do artigo de Skolem podem ser encontradas em George (1985) e Benacerraf and Wright (1985b); veja Jané (2001) para algumas críticas a esta linha de interpretação.

A última interpretação do argumento de Skolem está presente em um artigo de Ignacio Jané (2001). A leitura dada por Jané concorda com a interpretação clássica que assume que Skolem prepara um ataque bastante geral à teoria de conjuntos – e, em particular, à noção de conjunto absolutamente não-enumerável. Mas ele concorda com a interpretação fundacionalista ao entender que os ataques são preparados pouco a pouco, com o Paradoxo de Skolem atuando modestamente em uma linha de ataque. Grosseiramente, Jané pensa que Skolem tenta mostrar que não há um modo rigoroso para inicialmente se introduzir a noção de conjunto não-enumerável à matemática. Os paradoxos da teoria de conjuntos

[37]Esse modo de apresentação pode ser um tanto enganoso. O foco primário de Skolem em seu artigo é sobre a axiomatização da teoria de conjuntos de Zermelo, de 1908, e Skolem claramente entende essa axiomatização algebricamente. Assim, o argumento de relatividade contra Zermelo é de fato o argumento central do artigo, ainda que o argumento sirva para um projeto maior anti-fundacionalista e ainda que não leve a argumentos mais gerais em favor da relatividade da teoria de conjuntos.

mostram que não deveríamos ser ingênuas e levar o Teorema de Cantor ao pé da letra – logo, a própria demonstração de Cantor não nos força a aceitarmos conjuntos não-enumeráveis. O Paradoxo de Skolem mostra que adotar um entendimento algébrico da teoria de conjuntos também não nos força a aceitar conjuntos não-enumeráveis, uma vez que sempre podemos interpretar esses axiomas como sendo aplicados a modelos que são apenas enumeráveis.

É óbvio, como observa Jané, que há uma quantidade de estratégias que poderíamos usar para evitarmos essa aplicação do Paradoxo de Skolem: poderíamos usar uma quantidade não-enumerável de de axiomas, forçando nossos modelos a possuírem uma quantidade não-enumerável de domínios, poderíamos recorrer ao teorema Ascendente de Löwenheim-Skolem para mostrarmos que os axiomas de Zermelo também tem uma quantidade não-enumerável de modelos (ver seção 1), ou poderíamos nos mover para uma versão de segunda-ordem dos axiomas de Zermelo e então provarmos que esses axiomas podem ser satisfeitos apenas por modelos com domínios não-enumeráveis (ver seção 2.5). Infelizmente, cada uma dessas estratégias pressupõe que nós já tivéssemos um registro preliminar da noção de um conjunto não-enumerável – por exemplo, para inicialmente caracterizarmos um conjunto não-enumerável de axiomas, para formularmos o teorema Ascendente de Löwenheim-Skolem, ou para demonstrarmos que ZFC de segunda-ordem tem apenas modelos não-enumeráveis. Assim, a princípio nenhuma dessas estratégias pode ser utilizada para introduzir conjuntos não-enumeráveis à matemática. Ao menos, de qualquer modo, é o que Jané considera que Skolem está argumentando.

Essas são, portanto, as três principais interpretações do artigo de Skolem. Sem nos posicionarmos sobre qual dessas interpretações melhor captura as intenções do próprio Skolem, notamos que a maioria das contemporâneas de SKolem interpretaram-no como se estivesse dando algo como o argumento "tradicional" descrito acima e suas respostas ao Paradoxo de SKolem refletem essa interpretação. O próprio Zermelo aceitou a concepção algébrica de seus axiomas, embora ele tenha então insistido que eles deveriam ser interpretados em termos de segunda-ordem e que, assim interpretados, eles não seria vítimas do Paradoxo de Skolem (Zermelo, 1930; Taylor, 1993; Ebbinghaus, 2003). Semelhantemente, Tarski sugeriu que o Paradoxo de Skolem poderia ser neutralizado ao se tratar "$\in$" como uma constante lógica em alguma versão da teoria de tipos (ver as observações publicadas ao fim de (Skolem, 1958)). Mas, enquanto ambas as sugestões permitiriam que matemáticas evitassem o Paradoxo de Skolem, elas dependem de se aceitar peças de um maquinário matematicamente poderoso, os quais Skolem – em qualquer leitura de seu artigo – quase certamente desejaria rejeitar. Dados os fins filosóficos de Skolem, portanto, essas respostas contemporâneas a seu paradoxo não pareceriam muito ameaçadoras (veja (Skolem, 1955) e (Skolem, 1958) para algumas das reflexões do próprio Skolem sobre esses tipos de respostas).

3.2 Ceticismo Skolemita

Ao longo dos anos, houve uma pequena, embora consistente, corrente de filósofas e lógicas que teriam considerado filosoficamente convincente o que se chamou de interpretação tradicional do artigo de Skolem – isto é, convincente enquanto argumento filosófico independente e não apenas enquanto uma interpretação do artigo de Skolem. A posição delas, a qual Michael Resnik denominou a posição “Skolemita”, defende que os teoremas de Löwenheim-Skolem de fato mostram que as noções da teoria de conjuntos são relativas. Na verdade, Skolemitas costumam ir um pouco além, afirmando que, embora um dado conjunto possa ser não-enumerável, “relativo aos meios de expressão de um sistema axiomático”, todo conjunto é enumerável quando considerado de uma perspectiva “absoluta” (Kneale and Kneale, 1962; Goodstein, 1963; Wang, 1962; Fine, 1968; Thomas, 1968, 1971).

Nesta seção, isolaremos a ideia central por trás de alguns desenvolvimentos dessas afirmações Skolemitas, e então consideraremos algumas das respostas a elas as quais apareceram na literatura recente. (Na seção 3.3, consideramos uma nova abordagem interessante da posição Skolemita). Começaremos com o argumento Skolemita. De modo amplo, este argumento se dá em três passos. Primeiro, argumenta-se que a concepção algébrica de teoria de conjuntos é a única concepção respeitável para matemáticas e filósofas contemporâneas adotarem. Segundo, segue-se Skolem ao argumentar que a concepção algébrica da teoria de conjuntos leva uma relatividade das noções da teoria de conjuntos. Finalmente, estende o argumento de Skolem para defender uma forma forte de relatividade mencionada no fim do parágrafo anterior – isto é, aquela em que todo conjunto é enumerável quando observado de uma perspectiva “absoluta”.

Para nossos propósitos, o segundo passo deste argumento já foi considerado em suficiente detalhamento no contexto da nossa discussão de Skolem; então vamos apenas revisar os pontos principais aqui. Na concepção algébrica da teoria de conjuntos, as noções básicas da teoria de conjuntos são caracterizadas ao se olhar para as axiomatizações em teoria de modelos da teoria de conjuntos de primeira-ordem. As noções que permanecem fixas quando nos movimentamos de modelo para modelo – no sentido de “fixas” que discutimos na seção anterior – têm um significado “absoluto”; noções que variam quando nos movemos de modelo para modelo têm um significado "relativo" apenas. Dado isso, os teoremas de Löwenheim-Skolem mostram que as noções de enumerabilidade e não-enumerabilidade irão de fato variar quando nos movimentamos de um modelo a outro. Na concepção algébrica da teoria de modelos, portanto, as noções são apenas "relativas"[38].

Isso nos leva aos passos 1 e 3 no argumento Skolemita. No passo 1 é onde diferentes versões deste argumento mostram maior variabilidade. Em alguns casos, o passo 1 é apenas pressuposto, portanto é difícil ter uma noção de como o argumento subjacente deveria de fato se desenvolver (Kneale and Kneale, 1962; Goodstein, 1963; Wang, 1962). Em outros casos, é sugerido que qualquer rejeição da concepção algébrica – e, em particular, qualquer movimento para simplesmente se levar a sério expressões como "todos os conjuntos" ou "é de fato enumerável" – leva a um retorno a alguma forma inaceitável de "Platonismo" (Fine, 1968; Thomas, 1968, 1971; Klenk, 1976). Ainda em outros casos, Skolemitas seguem a liderança de Skolem e recorrem aos paradoxos da teoria de conjuntos para fomentar sua rejeição ao Platonismo; elas então sugerem que o abandono do Platonismo leva a concepção algébrica da teoria de conjuntos a ser a única opção viável (Klenk, 1976).

Há uma outra estratégia disponível aqui: algumas autoras defendem a posição Skolemita ao utilizarem outros enigmas sobre a interpretação da linguagem matemática – isto é, enigmas outros que não o Paradoxo de Skolem – para motivar o movimento inicial do Platonismo para a concepção algébrica. Então, por exemplo, Klenk argumenta que podemos combinar um dos enigmas clássicos de Benacerraf – aquele apresentado em Benacerraf Benacerraf (1965) – a este tipo de argumento (Klenk, 1976)[39]. Do mesmo modo, Wright recorre a considerações Wittgensteinianas sobre a relação entre significado e uso para motivar uma posição Skolemita restrita (Benacerraf and Wright, 1985a). Finalmente, várias autoras suge-

[38]Devemos aqui enfatizar que esse passo no argumento skolemita beira ser um teorema. Tendo-se um entendimento forte o suficiente da concepção algébrica de teoria de conjuntos e o correto entendimento de "relatividade", a afirmação condicional de que a concepção algébrica implica alguma forma de relatividade da teoria de conjuntos apenas se segue dos teoremas de Löwenheim-Skolem. Àquelas que fariam objeções à relatividade da teoria de conjuntos, portanto, recomenda-se focar sua atenção aos argumentos skolemitas iniciais a favor da concepção algébrica, ao invés de focá-la em seus argumentos subsequentes a favor da relatividade.

[39]O argumento de Klenk é um pouco mais complicado do que essa sentença sugere. Klenk inicia por apresentar vários argumentos a favor do que estamos chamando de concepção algébrica de conjuntos e ela então observa que qualquer um desses argumentos tende a fazer a posição Skolemita parecer de algum modo plausível. Ela também observa, porém, que uma análise como aquela apresentada na seção 2 permitirá uma realista já comprometida a se esquivar das conclusões Skolemitas. Ela termina sugerindo que nós resolvamos esse impasse adotando algum tipo de formalismo sobre teoria de conjuntos.

riram que todo o desenvolvimento da teoria de conjuntos no século XX favorece a abordagem algébrica – afinal, toda a história do tema consistiu em um afastamento de abordagens ingênuas da teoria de conjuntos em direção a uma axiomatização formal (e especialmente uma formalização em primeira-ordem). Ver Klenk (1976) para este tipo de análise.

Voltemos, agora, ao terceiro passo no argumento Skolemita. O teorema matemático que se esconde sob esse passo é claro. Seja $\phi(x)$ uma fórmula que supostamente define um conjunto único – por exemplo, "x é o conjunto potência de ω" ou "x é o conjunto dos números reais"[40]. Então podemos encontrar um modelo $\mathbf{M} \models ZFC$ e um elemento $m \in \mathbf{M}$ tal que $\mathbf{M} \models \phi(m)$ e $\{m' \in \mathbf{M} : \mathbf{M} \models m' \in m\}$ é apenas enumerável. Então, se quisermos conceder que tudo o que é preciso para ser, digamos, o conjunto potência de ω é satisfazer a fórmula relevante que o define em algum modelo da teoria de conjuntos, então podemos dar sentido à afirmação de que ao menos uma instância do conjunto potência de ω é "de fato" enumerável. Se desejamos assumir além disso que é necessário apenas uma bijeção em uma dessas instâncias do conjunto potência de ω para torná-lo "absolutamente" enumerável, então podemos entender a afirmação Skolemita forte sobre enumerabilidade absoluta. É claro, nenhum desses movimentos finais se seguem no sentido estrito a partir da concepção algébrica de axiomatização; ambos são, porém, movimentos que um proponente da concepção algébrica pode muito bem achar convenientes.

Isto, portanto, nos dá a estrutura básica dos diversos argumentos Skolemitas. Antes de nos voltarmos a algumas respostas a esses argumentos que aparecem na literatura recente, é importante sermos claras sobre o papel que o próprio Paradoxo de Skolem pode e não pode desempenhar nesses argumentos. Por vezes, aparenta como se algumas Skolemitas pensassem que os teoremas de Löwenheim–Skolem por si mesmos mostrassem que há um problema com a nossa concepção usual de conjuntos: assim, os teoremas mostram que as noções da teoria de conjuntos são relativas, relatividade é incompatível com nossa concepção usual de conjuntos e então nossa concepção usual de conjuntos deve ser abandonada (Kneale and Kneale, 1962; Goodstein, 1963). Deveria ser claro da seção 2, no entanto, que essa linha de argumentação não tem chances de prosperar. A análise na seção 2 mostra que aquelas de nós que desejam adotar uma atitude ingenuamente realista em relação à teoria de conjuntos – ou até mesmo aquelas que adotam posições mais sofisticadas que se baseiam na concepção iterativa de conjuntos e/ou em alguma forma de estruturalismo de segunda-ordem – não terão problemas com o Paradoxo de Skolem. Daí, o paradoxo por si

[40]Mais formalmente, suponha que $ZFC \models \exists x[\phi(x) \wedge \forall y(\phi(y) \rightarrow x = y)]$.

mesmo não pode nos forçar a abandonar a nossa concepção usual de conjuntos.

Ao invés disso, a Skolemita bem-sucedida precisa seguir a abordagem clássica descrita no começo desta seção. Ela inicia com um argumento independente para a concepção algébrica da teoria de conjuntos – isto é, um argumento que nos levaria a abandonar a concepção usual de conjuntos em favor de uma concepção algébrica e (crucialmente) um argumento que, ele mesmo, não se baseia em questões relacionadas ao Paradoxo de Skolem. Tendo completado esse argumento, a Skolemita pode então passar a utilizar a concepção algébrica de conjuntos (além, é claro, dos teoremas de Löwenheim-Skolem) para defender afirmações sobre a relatividade da teoria de conjuntos feitas nos passos 2 e 3 de seu argumento.

Dois comentários adicionais sobre essa abordagem são necessários. Primeiro, devemos notar que essa abordagem fornece à Skolemita uma resposta para os tipos de argumentos que fizemos na seção 2. Em particular, permite a ela desafiar todos os nossos usos muito ingênuos de expressões como "o entendimento em português usual de '$\in$' ", "os verdadeiros elementos de m", "quantificadores os quais abrangem todo o universo da teoria de conjuntos", etc. Tendo um argumento independente contra a concepção usual de conjuntos, a Skolemita não se impressionará com uma "solução" para o Paradoxo de Skolem que se baseia no emprego ingênuo desses tipos de expressões. Veja (Thomas, 1968, 1971; Klenk, 1976).

Depois, deveríamos notar que, ainda que essa abordagem requeira que a Skolemita inicie com um argumento independente contra nossa concepção usual de conjuntos, ela não é necessária para tornar os teoremas de Löwenheim-Skolem eles próprios completamente supérfluos. Afinal, continua sendo um teorema que noções da teoria de conjuntos tais como enumerabilidade e não-enumerabilidade tornam-se relativas na concepção algébrica. Isso não é algo que aconteça em todas as noções da teoria de conjuntos – por exemplo, "x é o conjunto vazio" ou "x tem dezessete elementos" – e não é algo que apenas ocorra à concepção algébrica de axiomatização.

Tendo dito isso, este é um ponto onde a Skolemita precisa tomar cuidado. A menos que as considerações levantadas no passo 1 de seu argumento estejam intimamente ligadas aos detalhes da concepção algébrica – e ligadas de tal modo que torne esta concepção genuinamente atrativa como um entendimento positivo da teoria de conjuntos – o argumento Skolemita mais amplo está ameaçado por uma certa trivialidade retórica. Afinal, tendo a Skolemita os recursos para nos levar em direção à concepção algébrica da teoria de conjuntos – como no passo 1 deste argumento – então ela também tem os recursos para minar diretamente nossa concepção usual de conjuntos e de fazê-lo sem trazer o próprio Paradoxo de Skolem para a discussão. Se isto está correto, então o argumento Skolemita mais amplo bem poderia se resumir a criticar as noções usuais de teoria de conjuntos por serem "relativas" num contexto retórico em que a Skolemita já apresentara críticas ainda mais fortes a essas noções enquanto defendia o passo inicial em seu argumento. Isto iria além de ser apenas um pouco estranho[41].

Para evitar esse tipo de estranheza, pensamos que a Skolemita deveria estruturar seu argumento menos como uma crítica às nossas noções de teoria de conjuntos e mais como uma análise construtiva de uma concepção algébrica da teoria de conjuntos. Isto é, ela deveria focar primariamente em defender a concepção algébrica da teoria de conjuntos como uma concepção independentemente plausível de teoria de conjuntos (passo 1) e ela deveria então apresentar a relatividade da teoria de conjuntos simplesmente como uma nova e surpreendente consequência dessa concepção positiva (passos 2–3).Essa estratégia argumentativa abre espaço para que os teoremas de Löwenheim-Skolem realizem de fato algum trabalho filosófico – por exemplo, como descrito há dois parágrafos. Também dá ao passo 1 um foco mais preciso – e mais construtivo. Nessa leitura, o passo 1 atua principalmente para destacar as virtudes positivas da concepção algébrica; criticar as noções usuais da teoria de conjuntos é (no máximo) uma preocupação secundária[42]. (Veja a seção 3.3 para mais sobre esse tipo de consideração).

Isso nos leva à crítica do argumento Skolemita que aparece na literatura recente. É importante mencionar três formas gerais de crítica. Primeira, várias autoras responderam ao argumento Skolemita apenas vagarosa e cuidadosamente desembrulhando a matemática envolvendo os teoremas de Löwenheim-Skolem a fim de mostrar que os teoremas por si mesmos não causam problemas mesmo para entendimentos ingênuos da teoria de con-

[41]Com efeito, nossa preocupação aqui é que os argumentos iniciais no passo 1 fazem todo o trabalho filosófico para a Skolemita e que os teoremas de Löwenheim-Skolem apenas pegam carona (e, talvez, fazem um papel de levantar uma bandeira para promover o argumento Skolemita). Conforme indicado no texto principal, nós pensamos que há ao menos uma leitura do argumento Skolemita que evita esse tipo de trivialização e que permite aos teoremas de Löwenheim-Skolem fazer algum trabalho genuíno. Dito isso, esse é um argumento um tanto delicado de se fazer e não fica claro que nenhuma Skolemita real o tenha feito com sucesso.

[42]Obviamente, isso também deixa o passo 1 substancialmente mais difícil para a Skolemita. Suspeitamos que muitas das considerações discutidas na nossa exposição inicial do passo 1 – por exemplo, o kit de emergência de argumentos contra o "platonismo" que aparece em muitos desenvolvimentos da posição Skolemita – serão inadequadas nessa nova leitura do argumento.

juntos (Resnik, 1966; Benacerraf and Wright, 1985b; Bays, 2007a). Ainda que esse tipo de resposta seja efetiva contra a versão simplista do argumento Skolemita que discutimos há seis parágrafos, ela pouco contribui contra os argumentos mais sofisticados que consideramos no momento – por exemplo, argumentos que agora começam com uma crítica independente desses entendimentos ingênuos [43]. Dado isto, e dado que nós já discutimos esse tipo de resposta com algum detalhe na seção 2, não diremos mais nada sobre ele aqui.

Segunda, várias autoras responderam ao argumento Skolemita ao criticarem diretamente a concepção algébrica da teoria de conjuntos e defendendo entendimentos mais usuais e intuitivos da linguagem da teoria de conjuntos (Myhill, 1951; Resnik, 1969; Hart, 1970; Benacerraf and Wright, 1985b). Há três questões que devemos destacar aqui. Primeira, é difícil ver como a concepção algébrica forneceria uma versão geral da linguagem matemática, dado que a própria concepção parece pressupor uma teoria intuitiva de antemão, na qual os resultados em teoria de modelos são formulados e demonstrados (por exemplo, os teoremas de Löwenheim-Skolem). Essa questão é exacerbada quando nos atentamos para o terceiro passo no argumento Skolemita, uma vez que este passo parece requerer tanto uma versão absoluta dos números naturais, quanto uma versão absoluta de enumeração, a fim de formular sua concepção de "enumerabilidade absoluta" (veja Resnik (1969); Benacerraf and Wright (1985b); e Shapiro (1991); veja Thomas (1971); Klenk (1976); e Bellotti (2006) para algumas preocupações sobre essa linha de argumentação).

Observe aqui que essas considerações iniciais parecem advocar contra o uso de qualquer crítica completamente geral ao realismo matemático a fim de direcionar as pessoas à concepção algébrica dos axiomas. Superficialmente, afinal, qualquer crítica suficientemente geral do realismo se aplicaria tanto ao próprio modelo da Skolemita quanto à teoria de conjuntos clássica. É duvidoso, portanto, se a Skolemita pode realmente recorrer a, digamos, simples preocupações sobre "platonismo" ou sobre nosso acesso epistêmico a objetos matemáticos para motivar uma posição Skolemita absoluta. Em suma: o próprio fato de argumentos Skolemitas basearem-se em teoremas matemáticos substanciais parece forçar a Skolemita a aceitar que algumas partes da matemática não estão sujeitas à relatividade Skolemita. (Além das referências no parágrafo anterior, veja (Bays, 2001; Bellotti, 2005; Bays, 2007b) para uma discussão desse tipo de consideração no contexto do argumento sobre

[43] Para fazer jus, muitas das apresentações do argumento Skolemita se parecem bastante com o simples argumento discutido há alguns parágrafos, logo essa resposta técnica ainda tem um papel importante para esclarecer a dialética geral.

teoria de modelos de Putnam).

Claramente, este primeiro argumento abre a possibilidade de que a teoria de conjuntos seja um caso especial – isto, ainda que algumas áreas da matemática como teoria de números e análise devessem ser entendidas absolutamente, teoria de conjuntos, como teoria de grupos e topologia, deveria ainda ser entendida algebricamente. Infelizmente, há diversas diferenças óbvias entre a prática da teoria de conjuntos e aquela das áreas mais claramente algébricas, como a teoria de grupos. Assim, por exemplo, matemáticas tendem a tratar os axiomas da teoria de conjuntos como sendo menos rígidos que aqueles da teoria de grupos ou topologia. Em teoria de conjuntos, matemáticas às vezes levantam a questão se os axiomas de ZFC estão corretos – isto é, elas falam como se houvesse uma noção intuitiva de conjunto em relação à qual os axiomas de ZFC podem ser checados e descobertos insuficientes. Em teoria de grupos e topologia, em contraste, simplesmente não faz sentido algum falar em "noções intuitivas" que poderiam divergir da noção especificada pelos axiomas relevantes[44]. Numa mesma linha, teóricas de conjuntos às vezes debatem se deveríamos adicionar novos axiomas aos axiomas padrão da teoria de conjuntos – por exemplo, axiomas de grandes cardinais, ou axiomas como $\mathbf{V} = \mathbf{L}$ ou até mesmo apenas axiomas como $Con(ZFC)$. Em contraste, ninguém sonharia em fazer adições aos axiomas da teoria de grupos ou da topologia. Nesse sentido, então, uma abordagem algébrica à teoria de conjuntos é revisionária da prática de teoria de conjuntos de um modo que a abordagem algébrica à teoria de grupos não o é.

Finalmente, mesmo se aceitarmos uma concepção algébrica de teoria de conjuntos – talvez porque tenhamos um comprometimento maior com algum tipo de filosofia estruturalista da matemática – não é claro por que este comprometimento requer de nós que limitemo-nos a axiomatizações em primeira-ordem da teoria de conjuntos. Afinal, muitos do exemplos mais bem-sucedidos da abordagem algébrica à axiomatização – por exemplo, os resultados do século XIX de que a aritmética e a análise podem ter axiomatizações categóricas – baseiam-se no uso de uma lógica de segunda-ordem ao fundo. E, como observamos na seção 2, versões de segunda-ordem de ZFC não geram o Paradoxo de Skolem. Então, não é suficiente que Skolemitas defendam uma abordagem algébrica à axiomatização da teoria de conjuntos, elas precisam mostrar que uma abordagem algébrica de primeira-ordem é

[44]Claramente, algumas filósofas seguem uma linha semelhante sobre ZFC. Mas, no caso de ZFC, é ao menos uma questão em aberto se essa é a linha correta de se seguir. No caso da teoria de grupos ou topologia, ela é a única linha que de fato faz algum sentido.

o caminho certo a ser seguido. Veja Hart (1970) e Shapiro (1991) para desenvolvimentos desta linha de argumentação.

Chega, portanto, de falar sobre as críticas gerais da concepção algébrica de axiomatização e seu papel no argumento Skolemita. Nós nos voltamos agora para uma objeção mais direcionada ao terceiro passo deste argumento. Para fins de argumento, vamos conceder que a Skolemita mostrou que nossas noções de teoria de conjuntos são relativas e que, para todo tipo de conjunto que nós podemos definir com uma fórmula, há uma instância deste tipo de conjunto que é apenas enumerável. Então, há uma instância enumerável o conjunto potência de ω, uma instância dos números reais, etc[45]. Ainda assim, não é claro por que isto mostra que todo conjunto é "absolutamente" enumerável. Afinal, assim como o teorema Descendente de Löwenheim-Skolem[46] mostra que que podemos encontrar instâncias enumeráveis de todos esses conjuntos, o axioma Ascendente de Löwenheim-Skolem mostra que nós podemos também encontrar instâncias não-enumeráveis.

Dado isso, diversas críticas sugerem que as Skolemitas encaram dois problemas explanatórios e que, até o momento, nenhuma delas foi capaz de resolvê-los. Primeiro, a Skolemita precisa explicar como podemos identificar conjuntos através de diferentes modelos – isto é, por que deveríamos considerar que os diversos objetos que satisfazem "x é o conjunto potência de ω" em diferentes modelos da teoria de conjuntos são "o mesmo conjunto". Observe que alguma identificação do tipo é essencial se a Skolemita iniciará com a demonstração de enumerabilidade de um desses objetos e então usar essa demonstração para argumentar pela enumerabilidade absoluta de todos os outros (Resnik, 1966). Segundo, a Skolemita precisa explicar sua preferência por conjuntos enumeráveis. Ainda que a Skolemita possa identificar "instâncias" enumeráveis e não-enumeráveis de um determinado conjunto, ela precisa explicar por que essa identificação leva à conclusão de que todos os conjuntos são "absolutamente enumeráveis", ao invés de levar à conclusão de que todos os conjuntos são "absolutamente não-enumeráveis" (Resnik, 1966; Benacerraf and Wright,

[45]Recorde, mais uma vez, que nenhuma dessas conclusões são forçadas sobre a realista. A realista tem uma explicação genuinamente boa para dizer porque essas diversas "instâncias" são, na melhor das hipóteses, aproximações enumeráveis do verdadeiro conjunto potência de ω ou do verdadeiro conjunto dos números reais. Assim, estamos fazendo algumas concessões bastante generosas para seguir a Skolemita mesmo até esse ponto. Veja a seção 2.4 para um desenvolvimento mais detalhado desse tipo de consideração.

[46]Ainda que a palavra "descendente" não apareça no texto original, fica claro que o autor faz menção a esse axioma em contraposição ao axioma Ascendente. Assim, a tradução se reserva o direito de adicioná-la aqui para manter o sentido do texto.

1985b).

Essas são, portanto, as principais críticas à posição Skolemita que apareceram na literatura. Tratá-las por completo requereria, infelizmente, que nós mergulhássemos profundamente em questões relacionadas, por exemplo, ao status do nosso entendimento informal da linguagem da teoria de conjuntos, a legitimidade da quantificação de segunda-ordem e condições de identidade associadas a objetos matemáticos em filosofias estruturalistas da matemática. Explorar essas questões nos distanciaria demais do Paradoxo de Skolem. Para uma discussão recente de algumas das literaturas aqui relevantes, veja (Bellotti, 2006).

3.3 O Multiverso

Durante a última década, o teórico de conjuntos Joel Hamkins tem argumentado em favor de uma concepção de teoria de conjuntos que se parece surpreendentemente com a visão Skolemita tradicional (ainda que as motivações de Hamkins parece vir mais da própria teoria de conjuntos do que de literatura filosófica tradicional). Hamkins nota que, enquanto teóricas de conjuntos desenvolveram mais ferramentas e mais poderosas, a fim de construir e comparar diferentes modelos de forcing em teoria de conjuntos, teoria de modelos internos, imersões de grandes cardinais, etc. – eles têm sido cada vez menos capazes de tratar qualquer modelo em particular como canônico. Ao invés disso, a teoria de conjuntos tem se concentrado em comparar diferentes modelos de teoria de conjuntos, em lugar de separar um modelo como privilegiado. Hamkins argumenta, portanto, que teóricas de conjuntos deveriam aceitar o que ele chama de concepção "multiversista" de teoria de conjuntos – uma concepção na qual nenhum modelo da teoria de conjuntos é privilegiado e o propósito da teoria de conjuntos é simplesmente explorar as relações entre os diversos modelos.

Essa concepção multiversista está claramente relacionada à concepção algébrica discutida nas seções 3.1–3.2. Além disso, ela satisfaz um dos *desiderata* chave que isolamos na seção 3.2. Hamkins defende o multiverso como uma concepção independentemente plausível de teoria de conjuntos e argumenta que a motivação para aceitá-lo vem de dentro da prática matemática. (Isto é, Hamkins não argumenta que, porque as extensões via forcing são possíveis, nós estamos presas com uma relatividade da teoria de conjuntos; pelo contrário, ele argumenta que, porque as extensões de forcing são naturais, deveríamos acatar a relatividade da teoria de conjuntos). Nesse sentido, algo como o multiverso poderia muito bem constituir o modo "correto" de se desenvolver uma concepção algébrica.

Além disso, a concepção multiversista leva naturalmente a tipos de conclusão com as quais Skolemitas tradicionais tenderiam a concordar. Seja a um conjunto em algum modelo **M** (em que **M** habita em algum lugar do multiverso). Logo **M** possui uma extensão via forcing, **M**$[G]$, em que a é apenas enumerável. Isso fornece um verniz natural à afirmação Skolemita que “todo conjunto é enumerável de alguma perspectiva”. Semelhantemente, o viés Skolemita em favor de enumerabilidade (veja seção 3.2) pode ser explicado pelo fato de que, se a é enumerável em um modelo **M**, então ele permanece enumerável em todas as extensões daquele modelo. Em contraste, conjuntos não-enumeráveis podem sempre serem tornados enumeráveis ao passarmos para uma extensão via forcing apropriada. Para mais sobre o multiverso, veja Hamkins 2011[47] e (Hamkins, 2012). Para mais críticas, veja Koellner 2013 (em *Outras referências da Internet*[48]).

3.4 O Argumento Modelo-teórico de Putnam

Recentemente, a versão mais amplamente discutida do Paradoxo de Skolem veio (em uma versão) do chamado “argumento modelo-teórico contra o realismo” de Hilary Putnam. O objetivo geral de Putnam no argumento modelo-teórico é mostrar que a nossa linguagem é semanticamente indeterminada – que não há nenhum fato a que se referem os termos e predicados da nossa linguagem. No caso da teoria de conjuntos, portanto, ele quer mostrar que não há um único universo da teoria de conjuntos o qual nossos quantificadores percorram e não há uma única relação à qual a palavra “pertencimento” se refira. Nos termos do próprio Putnam, não há um único “modelo pretendido” para a linguagem da teoria de conjuntos.

Nas primeiras páginas de seu artigo de 1980, “Models and Reality”, Putnam argumenta que há ao menos um modelo pretendido para a linguagem da teoria de conjuntos que satisfaz

[47] N. da T.: A referência utilizada pelo autor para esse artigo devolve a mesma entrada que (Hamkins, 2012). A decisão de manter a leitura é para se aproximar do original.

[48] N. da T.: Ainda que o artigo final tenha sido recentemente atualizado, a referência dada aqui leva a uma página inexistente. Assim, mantive a linha de texto para manter a tradução, mas, infelizmente não há uma correspondência para essa referência.

o axioma da teoria de conjuntos **V** $=$ **L**[49]. Para demonstrá-lo, Putnam começa por assumir que há apenas duas coisas que poderiam atuar fixando o modelo pretendido para a linguagem da teoria de conjuntos. Primeira, há o que Putnam chama "restrições teóricas". Essas incluem os axiomas padrão da teoria de conjuntos, assim como os princípios e teorias de outros ramos da ciência. Segunda, há "restrições operacionais". Essas são apenas as diversas observações empíricas e medições que fazemos no curso da investigação científica.

Dadas essas suposições, Putnam argumenta que encontrar um modelo pretendido que satisfaz **V** $=$ **L** requer simplesmente encontrar um modelo de ZF$+$**V** $=$ **L** que satisfaça as restrições teóricas e operacionais relevantes. Sua estratégia para encontrar esse modelos se baseia no seguinte teorema:

Teorema 1 *ZF mais **V** $=$ **L** tem um ω-modelo que contém qualquer conjunto enumerável de números reais dado.*

Aqui, o fato de que este modelo satisfaz ZFC supostamente garante que ele satisfaz todas as restrições teóricas que vem da própria teoria de conjuntos, enquanto a riqueza de ZFC garante que o modelo também tem os recursos pra codificar nossas melhores teorias científicas (e assim satisfazer todas as restrições teóricas que vêm das ciências naturais). Finalmente, o fato de que esse modelo contém um conjunto arbitrário de números reais garante que ele pode codificar todas as diversas observações e medições que constituem a as nossas "restrições observacionais"[50]. Logo, se Putnam estiver correto ao pensar que

[49]Nós provavelmente deveríamos dizer um pouco mais sobre o contexto local deste argumento. O objetivo de Putnam nessa seção de seu artigo é mostrar que a indeterminação de referência leva a uma indeterminação no valor de verdade para sentenças como $V = L$, em suas palavras, essas sentenças "não possuem valor de verdade determinado … elas são apenas verdadeiras em alguns modelos pretendidos e falsas em outros". Como resultado, não "faz sentido" pensar que " '$V = L$' é de fato falsa, mesmo que ela seja consistente com a teoria de conjuntos" (p.5). Agora, porque Putnam assume estar argumentando contra Gödel – quem pensava que havia apenas um único "modelo pretendido" da teoria de conjuntos e que $V = L$ era falsa neste modelo –, Putnam não sente a necessidade de argumentar a favor de um modelo pretendido que satisfaça $V \neq L$ (ele assume que Gödel lhe concederá a existência de tal modelo). Logo, Putnam pensa que se ele puder simplesmente encontrar um modelo pretendido que satisfaça $V = L$, então ele terá concluído seu argumento; em suas palavras, ele terá demonstrado que a famosa " 'relatividade das noções da teoria de conjuntos' [de Skolem] estende-se para o valor de verdade de $V = L$' " (p. 8). Para ler mais sobre a matemática desse exemplo – por exemplo, definições de V e de L – veja (Jech, 1978) e (Kunen, 1980).

[50]Para colocar todo o conteúdo acima nos termos de Putnam, seja OP um conjunto de números reais que codifica todas as medições que seres humanos farão em toda sua existência. Então Putnam escreve:

os modelos pretendidos da teoria de conjuntos são fixados somente pela estrutura formal das nossas teorias científicas – incluindo nossos axiomas explícitos da teoria de conjuntos – e pelas medições físicas que calhou de fazermos, então esse teorema gera um modelo pretendido em que **V** $=$ **L** é verdadeiro.

Essa versão do argumento modelo-teórico tem três conexões com o Paradoxo de Skolem. Primeira, o próprio Putnam apresenta o argumento como um desenvolvimento natural do paradoxo. No início de seu artigo, Putnam fornece um rápido rascunho do Paradoxo de Skolem e então sugere que sua análise de **V** $=$ **L** surge ao se tomarem os argumentos de Skolem e "estendê-los mais ou menos na direção que ele [Skolem] pareceu indicar" (p.1). Segunda, como evidenciada pelas passagens citadas na note de rodapé 44, as conclusões gerais de Putnam se adequam bem aos entendimentos Skolemitas mais recentes do Paradoxo de Skolem – veja, por exemplo, sua conclusão de que **V** $=$ **L** não possui "valor de verdade determinado" (p. 5) ou ainda que " 'a relatividade das noções da teoria de conjuntos' [de Skolem] se estende a uma relatividade do valor de verdade de **V** $=$ **L**" (p. 8). Finalmente, e mais importante, a demonstração do teorema de Putnam se baseia crucialmente nos teoremas de Löwenheim-Skolem. (Muito grosseiramente, Putnam inicia por aplicar *a* **L** o teorema Descendente de Löwenheim-Skolem, a fim de provar que seu teorema é verdadeiro *em* **L**; ele então utiliza a noção de absoluto de Shoenfield para refletir o teorema de volta para **V**[51]).

O argumento de Putnam recebeu diversos tipos de críticas na literatura. No front técnico, Bays argumenta que o uso de Putnam do teorema Descendente de Löwenheim-Skolem é ilegítimo, já que sistemas padrão da teoria de conjuntos não nos permitem aplicar esse teorema a uma classe própria como **L**. De fato, ainda que deixássemos os detalhes da demonstração de Putnam de lado, considerações Gödelianas mostram que o teorema de Putnam não pode de modo algum ser demonstrado em ZFC (já que o teorema implica a consistência de

> Agora, suponha que formalizamos toda a linguagem da ciência dentro da teoria de conjuntos ZF mais $V = L$. Qualquer modelo de ZF que contenha um conjunto abstrato isomorfo a OP pode ser estendido a um modelo para essa linguagem formalizada de ciência que é padrão com respeito a OP; logo ...podemos encontrar um modelo para toda a linguagem da ciência que satisfaz 'tudo é construível' e que atribui o valor correto a todas magnitudes físicas. (p. 7)

[51]São necessários mais alguns detalhes aqui. Seja x a coleção enumerável de números reais. A demonstração de Putnam começa por observar que, no caso especial em que permitimos que nossos ω-modelos sejam enumeráveis, podemos codificar tanto o modelo e x por reais únicos. Nesse caso, portanto, o teorema pode ser formulado como uma sentença Π_2 da forma: (Para todo real s)(existe um real M) tal que $(\ldots M, s \ldots)$. Deste ponto, Putnam argumenta conforme se segue:

ZFC). É claro, se Putnam desejar utilizar um teoria de fundo mais forte para demonstrar seu teorema – por exemplo, ZFC+"existe um cardinal inacessível" – então ele pode evitar esses tipos de críticas. Nesse caso, porém, não fica claro por que ainda deveríamos considerar que o modelo que resulta do teorema de Putnam satisfaz nossas restrições teóricas. Afinal, qualquer pessoa que aceite os novos axiomas utilizados na demonstração revisada de Putnam terá restrições teóricas que vão de algum modo para além de ZFC+**V** $=$ **L** – por exemplo, suas restrições teóricas podem muito bem incluir o axioma "há um cardinal inacessível". Veja Bays (2001) para a formulação original de Bays dessa objeção; veja Velleman (1998) e Gaifman (2004) para algumas formulações alternativas; veja Bellotti (2005) e Bays (2007b) para uma discussão crítica; e veja capítulo 3 esp. § 3.3.3) de (Hafner, 2005) para discussão de uma consideração semelhante sobre o uso da transitividade por Putnam.

Button (2011) argumenta que, ainda que esse tipo de crítica técnica se posicione contra a versão do argumento de Putnam que explicitamente invoca o teorema Descendente de Lowenheim-Skolem, há formulações alternativas dele que podem evitar essa crítica. Em particular, Button observa que mesmo teorias muito fracas podem demonstrar teoremas como: "se ZFC é consistente então ZFC tem um modelo enumerável". Como qualquer proponente de ZFC deve aceitar que ZFC é consistente, essas teorias fracas são suficientes para se obter diversas variações do argumento de Putnam. Veja (Button, 2011) para desenvolvi-

---

> Considere esta sentença no modelo interno $V = L$. Para todo s no modelo interno – isto é, para todo s em L – existe um modelo – a saber, o próprio L – que satisfaz '$V = L$' e contém s. Pelo teorema Descendente de Löwenheim-Skolem, existe um modelo enumerável que é elementarmente equivalente a L e contém s. (Estritamente falando, não precisamos aqui apenas do teorema Descendente de Löwenheim-Skolem, mas da construção do 'Invólucro de Skolem' que é utilizada para se demonstrar o teorema). Pelo trabalho de Gödel, esse submodelo enumerável está ele próprio em L, assim como, fato facilmente verificável, também estão os reais que o codificam. Logo, a sentença Π_2 acima é verdadeira no modelo interno $V = L$.
>
> Mas Shoenfield demonstrou que sentenças Π_2 são absolutas: se uma sentença Π_2 é verdadeira em L, ela deve ser verdadeira em V. Então a sentenças acima é verdadeira em V. (p. 6)

Para discussões mais detalhadas do lado matemático deste argumento, veja Bays (2001); Bellotti (2005); Bays (2007b); e Hafner (2005). Veja Jech (1978) e Kunen (1980) para o pano de fundo relevante em teoria de conjuntos (por exemplo, sobre L e noção de absoluto de Shoenfield).

mento dessa consideração. Veja (Bellotti, 2005) e (Bays, 2007a) para a discussão de uma consideração mais ou menos semelhante.

Permanecendo na linha técnica, diversas autoras observaram a tensão no modo como o argumento de Putnam lida com a noção de finitude. Por um lado, Putnam precisa usar essa noção para caracterizar seu modelo como um ω-modelo e (até mesmo) para dar sentido às definições formais de uma linguagem de primeira-ordem e da relação de satisfação de primeira-ordem[52]. Por outro lado, Putnam não pode permitir que opositoras a seu argumento utilizem essa noção para especificar o que elas pensam fazer um modelo pretendido. Se suas oponentes pudessem usar essa noção, então elas poderiam definir a noção de um modelo sendo "bem-fundado" e isto seria suficiente para descartar os modelos gerados pelo teorema de Putnam. Nesse sentido, portanto, o argumento de Putnam parece se basear numa assimetria sem motivação entre os tipos de maquinários técnicos que ele mesmo utiliza e os tipos de maquinários que ele expõe a críticas. Veja (Bays, 2001) e (Bellotti, 2005) para desenvolvimentos dessa consideração; veja seção 3.4 de (Hafner, 2005) para algumas reflexões críticas.

De um lado mais puramente filosófico, muitas autoras criticaram a suposição de Putnam de que apenas satisfazer uma formalização em primeira-ordem de nossas restrições teóricas é suficiente para que um modelo seja "pretendido". Então, por exemplo, Hacking argumentou que deveríamos realmente estar comprometidos com uma formulação de teoria de conjuntos de segunda-ordem e que o teorema-chave de Putnam não se aplica para tais formulações (Hacking, 1983). Outras argumentaram que o modelo pretendido para a teoria de conjuntos necessita ser transitivo e que, mais uma vez, não há razões para crer que o modelo produzido pelo teorema de Putnam seja transitivo (Bays, 2001). Finalmente, como mencionado no parágrafo anterior, muitas autoras sugeriram que um modelo pretendido para a teoria de conjuntos deveria ao menos ser bem-fundado, mas não há razões para crer que o modelo de Putnam seja bem-fundado (Bellotti, 2005).

---

[52]Aqui, um ω-modelo é apenas um modelo que captura corretamente os números naturas – isto é, um modelo em que os "números naturais" do modelo são isomorfos aos números reais de fato. Putnam necessita da noção de finitude para capturar linguagens de primeira-ordem, porque as sentenças de uma linguagem de primeira-ordem pode ser de comprimento arbitrário finito, mas não podem ser infinitas. Além disso, o fato que ZFC utiliza esquemas de axiomas para capturar Substituição e Separação significa que realmente precisamos da generalidade aqui – esses esquemas não podem sequer ser formulados sem que utilizemos a noção de fórmula arbitrária. Uma consideração semelhante se aplica ao modo como a definição de satisfação de primeira-ordem faz uso de recursão.

A resposta de Putnam para esse tipo de objeção é interessante. Muito grosseiramente, Putnam sugere que quaisquer condições impostas sobre modelos pretendidos as quais outras filósofas podem vir a propor – por exemplo, aquelas mencionadas no parágrafo anterior – deveriam elas mesmas serem formalizadas em termos de primeira-ordem e tratadas como novas restrições teóricas. Quando essas novas restrições teóricas são colocadas de volta no argumento de Putnam, ele irá gerar mais uma vez um modelo que "satisfaz" essas restrições. Assim, ao simplesmente adotar uma leitura particularmente flexível da frase "restrições teóricas", Putnam garante que quase todas as condições impostas sobre modelos pretendidos podem simplesmente ser trazidas de volta para seu argumento original (Putnam, 1980, 1983, vii-xii)[53].

Este argumento – geralmente chamado de "apenas mais teoria" – recebeu uma atenção enorme na literatura. A resposta mais comum a ele envolve traçar uma distinção entre *descrever* as propriedades de um modelo que o fazem um modelo pretendido e simplesmente *adicionar novas sentenças* para que um modelo as satisfaça. Dito de outro modo, envolve distinguir entre *mudar* a semântica na qual nossos axiomas são interpretados – por exemplo, restringindo a classe de estruturas que contam como modelos para a nossa linguagem e/ou fortalecendo a noção de satisfação que liga sentenças a modelos – e simplesmente *adicionar novos axiomas* para serem interpretados utilizando a mesma semântica antiga. A resposta então passa a argumentar que propostas como aquelas discutidas há dois parágrafos – por exemplo, que modelos pretendidos deveriam ser transitivos ou bem-fundados ou satisfazerem ZFC de segunda-ordem – deveriam ser entendidos como pertencentes à parte das descrições nessa distinção, ao invés de pertencer à parte da "adição de sentenças" (ainda que seja nesta onde o argumento "apenas mais teoria" de Putnam insiste resolutamente em colocá-las).

Putnam, por sua vez, argumentou que esse tipo de resposta comete petição de princípio contra seu argumento geral. O argumento de Putnam, afinal, se ocupa da questão se nossa linguagem matemática tem algum significado determinado e a resposta que estamos consi-

[53]Isso reflete a leitura padrão do argumento de Putnam na literatura (Devitt (1984), capítulo 11; Lewis (1984); and (1991); Cleve (1992); e Crispin Wright (1997)). Recentemente, no entanto, várias comentadoras desafiaram essa interpretação. Elas argumentam que a conversa de Putnam sobre "apenas mais teoria" meramente deveria destacar a inadequação teórica de diversas teorias de referência particulares; não deveria providenciar um argumento geral do tipo rascunhado acima (Anderson, 1993; Douven, 1999; Haukioja, 2001). Veja (García-Carpintero, 1996) e (Bays, 2008) (§2) para algumas reflexões críticas nessas interpretações revisionistas do argumento de Putnam.

derando parece assumir simplesmente que tem esse tal significado quando a resposta usa frases como "transitivo", "bem-fundado" ou "conjunto potência completo de **M**" para descrever sua noção de "modelo pretendido". Em suma: enquanto a determinação da linguagem matemática for um problema, será uma petição de princípio a questão para se fazer uso livre dessa linguagem para descrever o modelo pretendido da teoria de conjuntos. Ao menos, de qualquer forma, Putnam tenta argumentar.

Como indicado acima, esse aspecto do argumento de Putnam gerou uma enorme literatura. Veja (Devitt, 1984), capítulo 11; (Lewis, 1984; and, 1991; Cleve, 1992; e Crispin Wright, 1997; Chambers, 2000; Bays, 2001, 2008) para algumas críticas representativas do argumento de Putnam. Veja (Putnam, 1983), vii–xii e (Putnam, 1989) para a resposta de Putnam. Veja (Anderson, 1993; Douven, 1999; Haukioja, 2001; Kroon, 2001) para algumas defesas recentes desse aspecto do argumento de Putnam.

4. Conclusão

Fechamos essa entrada com uma breve retomada dos dois caminhos principais que tentamos enfatizar. Primeiro, de um ponto de vista puramente matemático, não há conflitos entre o teorema de Cantor e os Teoremas de Löwenheim-Skolem. Há uma solução técnica ao Paradoxo de Skolem que explica porque os teoremas de Löwenheim-Skolem não apresentam problemas tanto para formas ingênuas do realismo em teoria de conjuntos, quanto para diversas formas de teoria de conjuntos axiomatizadas. Logo, não há chances de se utilizar apenas os teoremas de Löwenheim-Skolem para gerar conclusões Skolemitas substanciais. É claro, há ainda algumas questões técnicas interessantes que circundam o Paradoxo de Skolem. Por exemplo, podemos olhar como o paradoxo atua no contexto de modelos particulares de primeira-ordem; podemos examinar o grau com que diversos tipos de lógicas que não são de primeira-ordem são suscetíveis ao paradoxo; e podemos tentar isolar precisamente as propriedades da lógica de primeira-ordem que permitem que o paradoxo se aplique a ela. Cada um desses tópicos claramente se relaciona ao Paradoxo de Skolem e cada um deles levanta questões sobre a relação entre teoria de modelos e teoria de conjuntos que valem a pena serem exploradas. Quando considerado isoladamente, o Paradoxo de Skolem não oferece ameaças à teoria de conjuntos clássica.

Segundo, se chegarmos ao Paradoxo de Skolem com dúvidas prévias sobre a teoria de conjuntos clássica – por exemplo, os tipos de dúvidas que existem por trás de algumas das reconstruções mais sofisticadas do argumento original de Skolem, os tipos de dúvidas que existem por trás das versões mais plausíveis do passo 1 no argumento Skolemita ou ainda os tipos de dúvidas sobre a determinação da semântica que existe por trás do argumento modelo-teórico de Putnam – então é bem possível que guiemos a discussão do Paradoxo de Skolem em direção a algum tipo de conclusão filosófica interessante. Claramente ainda haverão desafios aqui: precisamos da conta do status das teorias de fundo nas quais demonstramos os teoremas de Löwenheim-Skolem, explicar o significado especial das axiomatizações da teoria de conjuntos em primeira-ordem, e talvez precisamos explicar como podemos identificar elementos olhando para diferentes modelos da teoria de conjuntos. A princípio, porém, esses tipos de usos sofisticados do Paradoxo de Skolem não são evitados pela solução técnica ao paradoxo mencionada no parágrafo anterior. Se colocarmos filosofia suficiente na nossa análise do Paradoxo de Skolem, então deveríamos esperar obter ao menos um pouco de filosofia.

Referências Bibliográficas

Barry Taylor and. 'just more theory': A manoeuvre in putnam's model-theoretic argument for antirealism. *Australasian Journal of Philosophy*, 69(2):152–166, 1991. doi: 10.1080/00048409112344601. URL https://doi.org/10.1080/00048409112344601.

David Leech Anderson. What is the model-theoretic argument? *Journal of Philosophy*, 90 (6):311–322, 1993.

Calixto Badesa. *The Birth of Model Theory: Lowenheim's Theorem in the Frame of the Theory of Relatives*. Princeton University Press, 2004.

Timothy Bays. On putnam and his models. *Journal of Philosophy*, 98(7):331, 2001. doi: 10.2307/2678439.

Timothy Bays. The mathematics of skolem's paradox. In *Philosophy of Logic*, pages 615–648. Elsevier, 2007a.

Timothy Bays. More on putnam's models: A reply to bellotti. *Erkenntnis*, 70(2):283, 2007b.

Timothy Bays. Two arguments against realism. *Philosophical Quarterly*, 58(231):193–213, 2008. doi: 10.1111/j.1467-9213.2007.505.x.

Luca Bellotti. Putnam and constructibility. *Erkenntnis*, 62(3):395–409, 2005. doi: 10.1007/s10670-004-0602-7.

Luca Bellotti. Skolem, the skolem 'paradox'and informal mathematics. *Theoria*, 72(3):177–

212, 2006.

Paul Benacerraf. What numbers could not be. *The philosophical review*, 74(1):47–73, 1965.

Paul Benacerraf and Crispin Wright. Skolem and the skeptic. *Aristotelian Society Supplementary Volume*, 59(1):85–138, 1985a. doi: 10.1093/aristoteliansupp/59.1.85.

Paul Benacerraf and Crispin Wright. Skolem and the skeptic. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 59:85–137, 1985b.

Tim Button. The metamathematics of putnam?s model-theoretic arguments. *Erkenntnis*, 74 (3):321–349, 2011. doi: 10.1007/s10670-011-9270-6.

Tim Button and Sean Walsh. *Philosophy and model theory*. Oxford University Press, 2018.

Timothy Chambers. A quick reply to putnam's paradox. *Mind*, 109(434):195–197, 2000. doi: 10.1093/mind/109.434.195.

James Van Cleve. Semantic supervenience and referential indeterminacy. *Journal of Philosophy*, 89(7):344–361, 1992. doi: 10.2307/2940764.

M. Devitt. Realism & truth. 1984.

Igor Douven. Putnam?s model-theoretic argument reconstructed. *Journal of Philosophy*, 96 (9):479–490, 1999. doi: 10.2307/2564709.

Bob Hale e Crispin Wright. Putnam's model-theoretic argument against metaphysical realism. In Bob Hale e Crispin Wright e Alexander Miller, editor, *A Companion to the Philosophy of Language*, pages 703–733. Wiley-Blackwell, 1997.

H-D Ebbinghaus. Löwenheim-skolem theorems. In *Philosophy of Logic*, pages 587–614. Elsevier, 2007.

Heinz-Dieter Ebbinghaus. Zermelo: Definiteness and the universe of definable sets. *History and Philosophy of Logic*, 24(3):197–219, 2003.

Arthur I. Fine. Quantification over the real numbers. *Philosophical Studies*, 19(1-2):27–32, 1968. doi: 10.1007/bf00463056.

Abraham Adolf Fraenkel, Yehoshua Bar-Hillel, and Azriel Levy. *Foundations of set theory*, volume 67. Elsevier, 1973.

Haim Gaifman. Non-standard models in a broader perspective. *Nonstandard models of arithmetic and set theory, Contemporary mathematics*, 361:1–22, 2004.

Manuel García-Carpintero. The model-theoretic argument: Another turn of the screw. *Erkenntnis*, 44(3):305–316, 1996. doi: 10.1007/bf00167660.

Alexander George. Skolem and the löwenheim-skolem theorem: A case study of the philosophical significance of mathematical results. *History and Philosophy of Logic*, 6(1):75–89,

1985. doi: 10.1080/01445348508837077.

Marcus Giaquinto. *The Search for Certainty: A Philosophical Account of Foundations of Mathematics*. Oxford University Press UK, Oxford, England, 2002.

R. L. Goodstein. The significance of incompleteness theorems. *British Journal for the Philosophy of Science*, 14(55):208–220, 1963. doi: 10.1093/bjps/xiv.55.208.

Ian Hacking. Representing and intervening. *British Journal for the Philosophy of Science*, 35 (4):381–390, 1983.

Johannes Hafner. *From Metamathematics to Philosophy: A Critical Assessment of Putnam's Model-Theoretic Arguments*. PhD thesis, University of California at Berkeley, 2005.

Michael Hallett. Putnam and the skolem paradox. In Peter Clark and Bob Hale, editors, *Reading Putnam*, pages 66–97. Blackwell, 1994.

Joel David Hamkins. The set-theoretic multiverse. *Review of Symbolic Logic*, 5(3):416–449, 2012. doi: 10.1017/s1755020311000359.

W. D. Hart. Skolem's promises and paradoxes. *Journal of Philosophy*, 67(4):98–109, 1970. doi: 10.2307/2024030.

Jussi Haukioja. Not so quick: A reply to chambers. *Mind*, 110(439):699–702, 2001. doi: 10.1093/mind/110.439.699.

Ignagio Jané. Reflections on skolem's relativity of set-theoretical concepts. *Philosophia Mathematica*, 9(2):129–153, 2001.

Thomas J. Jech. *Set Theory*. Academic Press, New York, NY, USA, 1978.

Stephen Cole Kleene. *Mathematical Logic*. Dover Publications, Mineola, N.Y., 1967.

Virginia Klenk. Intended models and the löwenheim-skolem theorem. *Journal of Philosophical Logic*, 5(4):475–489, 1976. doi: 10.1007/bf02109439.

William Kneale and Martha Kneale. The development of logic. *Philosophy*, 40(151):79–83, 1962.

Peter koellner. Hamkins on the multiverse, 2013. URL https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=26fb06dc48cfe92d32ddbd8fd963047687b4459d. manuscript.

Frederick Kroon. Chambers on putnam's paradox. *Mind*, 110(439):703–708, 2001. doi: 10.1093/mind/110.439.703.

Kenneth Kunen. *Set Theory: An Introduction to Independence Proofs*. North-Holland, 1980.

Michael Levin. Putnam on reference and constructible sets. *British Journal for the Philosophy of Science*, 48(1):55–67, 1997. doi: 10.1093/bjps/48.1.55.

David Lewis. Putnam?s paradox. *Australasian Journal of Philosophy*, 62(3):221–236, 1984. doi: 10.1080/00048408412340013.

Per Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32(3):186–195, 1966. doi: 10.2307/2270871.

Per Lindström. On extensions of elementary logic. *Theoria*, 35(1):1–11, 1969. doi: 10.1111/j.1755-2567.1969.tb00356.x.

Leopold Löwenheim. On possibilities in the calculus of relatives. *1967), From Frege to Gödel. A Source Book in Mathematical Logic*, 1931:228–251, 1879.

Charles McCarty and Neil Tennant. Skolem's paradox and constructivism. *Journal of Philosophical Logic*, pages 165–202, 1987.

Clifton McIntosh. Skolem's criticisms of set theory. *Noûs*, 13(3):313–334, 1979. doi: 10.2307/2215103.

F. A. Muller. Deflating skolem. *Synthese*, 143(3):223–253, 2005. doi: 10.1007/s11229-005-0800-0.

J Myhill. On the ontological significance of the löwenheim-skolem theorem. w: M. white (ed.) academic freedom, logic and religion, 1951.

Hilary Putnam. Models and reality. *Journal of Symbolic Logic*, 45(3):464–482, 1980. doi: 10.2307/2273415.

Hilary Putnam, editor. *Realism and Reason*. Cambridge University Press, New York, 1983.

Hilary Putnam. *Model Theory and the 'factuality' of Semantics*, pages 213—231. Blackwell, 1989.

Michael David Resnik. On skolem's paradox. *Journal of Philosophy*, 63(15):425–438, 1966. doi: 10.2307/2024063.

Michael David Resnik. More on skolem's paradox. *Noûs*, 3(2):185–196, 1969. doi: 10.2307/2216264.

Stewart Shapiro. *Foundations without foundationalism: A case for second-order logic*, volume 17. Clarendon Press, 1991.

Th Skolem and Jens Erik Fenstad. *Selected Works in Logic*. Oslo,, 1970.

Thoraf Skolem. Some remarks on axiomatised set theory. In J. Van Heijenoort, editor, *Iteheijenoort*, pages 290–301. Harvard University Press, 1922.

Thoralf Skolem. A critical remark on foundational research. pages 581–6. 1955.

Thoralf Skolem. Une relativisation des notions mathématiques fondamentales. In E. J. Fenstad, editor, *Selected Works in Logic (1970)*, pages 633–8. Universitetsforlaget, 1958.

R Gregory Taylor. Zermelo, reductionism, and the philosophy of mathematics. *Notre Dame journal of formal logic*, 34(4):539–563, 1993.

William J. Thomas. Platonism and the skolem paradox. *Analysis*, 28(6):193–6, 1968. doi: 10.1093/analys/28.6.193.

William J. Thomas. On behalf of the skolemite. *Analysis*, 31(6):177–86, 1971. doi: 10.1093/analys/31.6.177.

Dirk Van Dalen and Dirk Van Dalen. *Logic and structure*, volume 3. Springer, 1994.

Jean Van Heijenoort. *From Frege to Gödel: A source book in mathematical logic, 1879–1931*, volume 9. Harvard University Press, 1967.

Daniel J Velleman. Review of levin's" putnam on reference and constructible sets'(1997). *Mathematical Reviews*, 98, 1998.

Hao Wang. *A Survey of Mathematical Logic*. North-Holland Publishing Company, Amsterdam, 1962.

Ernst Zermelo. Über grenzzahlen und mengenbereiche: Neue untersuchungen Über die grundlagen der mengenlehre. *Fundamenta Mathematicæ*, 16:29–47, 1930.

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

Alan Renê Antezana é doutorando nas áreas de Lógica e Metafísica no Centro de Lógica, Epistemologia e História da Ciência na Universidade Estadual de Campinas (CLE/UNICAMP), sendo atualmente bolsista FAPESP sob a orientação do prof. Dr. Marcelo Esteban Coniglio. Possui mestrado em Filosofia nas mesmas áreas pela Universidade de Brasília (PPGFIL/UnB), tendo recebido bolsa CAPES sob a orientação do prof. Dr. Alexandre Costa-Leite. Foi também diagramador e membro de corpo editorial do periódico Revista de Filosofia Moderna e Contemporânea (PPGFIL/UnB).

Daniel Alves da Silva Lopes Diniz é doutorando em Filosofia (linha de pesquisa: Lógica) na Universidade Estadual de Campinas (ingresso em 2020). Sua tese avalia a viabilidade de interpretar lógica possibilística como uma teoria da vagueza. É mestre em Filosofia (linha de pesquisa: Lógica) (2020), e bacharel em Linguística (2016) pela mesma instituição.

Renato Semaniuc Valvassori é doutorando pela Universidade Estadual de Campinas (UNICAMP). Mestre e graduado pela mesma instituição, realizou em 2023 um estágio de pesquisa na Universidade de Turim. Possui experiência nas áreas de metafísica e filosofia da linguagem. Atualmente, sob orientação do prof. Dr. Marco Ruffino, é bolsista FAPESP e dedica-se ao estudo de filosofia da ficção.

Rafael Cavalcanti de Souza é doutorando pela Universidade Estadual de Campinas (PPGFil-UNICAMP). Mestre em Filosofia pela mesma instituição e programa, e graduado em Filosofia pela Universidade Federal de Pernambuco. Tem experiência em filosofia da ciência e matemática na Grécia antiga, com ênfase em Aristóteles e Platão. Foi bolsista no mestrado e, atualmente, é bolsista FAPESP no doutorado, sob orientação do prof. Dr. Lucas Angioni.

Mahan Vaz cursa doutorado em Filosofia pela UNICAMP em cotutela com a Ruhr-Universität Bochum, na área de Lógica, sob orientação de Marcelo Coniglio e Heinrich Wansing, com bolsas da FAPESP e DAAD. É mestre pela UNICAMP, o qual cursou com bolsa da CAPES, sob orientação de Giorgio Venturi, com período na Université Paris Cité. Possui graduação em Filosofia também pela UNICAMP, com bolsa de iniciação científica da FAPESP, sob orientação de Giorgio Venturi. Atualmente pesquisa lógicas modais paraconsistentes, em particular semânticas não-determinísticas para tais lógicas. Seus interesses incluem lógica, lógica matemática, filosofia da ciência, epistemologia, ética formal e lógica deôntica.

Editora UFPel

DISSERTATIO FILOSOFIA