

Luiz Henrique de Araújo Dutra  
Alexandre Meyer Luz  
(orgs.)

Linguagem  
Ontologia  
e Ação



RUMOS DA EPISTEMOLOGIA (vol. 10)

**Linguagem, Ontologia e Ação**

**Universidade Federal de Santa Catarina**

Reitor: Álvaro Toubes Prata

**Departamento de Filosofia**

Chefe: Gustavo Caponi

**Programa de Pós-Graduação em Filosofia**

Coordenador: Alessandro Pinzani

**NEL – Núcleo de Epistemologia e Lógica**

Coordenador: Cezar A.Mortari

**GECL – Grupo de Estudos sobre Conhecimento e Linguagem**

Coordenador: Luiz Henrique de A. Dutra

***Principia* – Revista Internacional de Epistemologia**

Editor responsável: Luiz Henrique de A. Dutra

Editores assistentes: Cezar A.Mortari e Jaimir Conte

**VII Simpósio Internacional Principia**

A Filosofia de Nelson Goodman

**Comissão organizadora**

Alberto Cupani  
Alexandre Meyer Luz  
Cezar A.Mortari  
Jaimir Conte

**Comissão científica**

Luiz Henrique de Araújo Dutra (UFSC, presidente)  
Catherine Elgin (Harvard University)  
Gary Hatfield (University of Pennsylvania)  
Oswaldo Chateaubriand Filho (UCRJ)  
Wilson Mendonça (UFRJ)

RUMOS DA EPISTEMOLOGIA (vol. 10)

Luiz Henrique de Araújo Dutra  
Alexandre Meyer Luz  
(orgs.)

# LINGUAGEM, ONTOLOGIA E AÇÃO

NEL - Núcleo de Epistemologia e Lógica  
Universidade Federal de Santa Catarina  
Florianópolis  
2011

© 2011, NEL – Núcleo de Epistemologia e Lógica, UFSC

ISBN: 978-85-87253-16-3 (papel)  
978-85-87253-17-0 (e-book)

Universidade Federal de Santa Catarina  
Centro de Filosofia e Ciências Humanas  
Bloco D, 2º andar, sala 209  
Florianópolis, SC, 88010-970  
(48) 3721-8612  
nel@cfh.ufsc.br  
www.cfh.ufsc.br/~nel

FICHA CATALOGRÁFICA  
Dados Internacionais de Catalogação na Publicação (CIP)

L755 Linguagem, ontologia e ação [recurso eletrônico] / (orgs.) Luiz Henrique de Araújo Dutra, Alexandre Meyer Luz. – Florianópolis: NEL/UFSC, 2011.

(Rumos da epistemologia ; v. 10)

Inclui bibliografia.

Exigência do sistema: conexão com a internet, browser e Adobe Acrobat Reader.

Modo de acesso: World Wide Web.

Trabalhos apresentados no VII Simpósio Internacional Principia, em Florianópolis em agosto de 2011, revistos e ampliados.

ISBN 978-85-87253-17-0

1. Teoria do conhecimento. 2. Lógica - Filosofia. 3. Ontologia - Mecânica quântica. 4. Linguagem e línguas - Filosofia. 5. Linguística. 6. Ética - Política. I. Dutra, Luiz Henrique de Araújo. II. Luz, Alexandre Meyer. III. Série.

CDU 16

CDD 160

(Bibliotecária responsável: Sabrina Leal Araujo – CRB 10/1507)

Reservados todos os direitos de reprodução total ou parcial por  
NEL – Núcleo de Epistemologia e Lógica, UFSC.

Impresso no Brasil

## APRESENTAÇÃO

As primeiras versões dos trabalhos reunidos neste livro – assim como no volume *Temas de filosofia do conhecimento* – foram apresentados no VII Simpósio Internacional *Principia*, realizado em Florianópolis em agosto de 2011. Os textos foram revistos e ampliados por seus autores, e reunidos nas quatro seções temáticas que compõem este volume. No conjunto, esses trabalhos tratam de aspectos da compreensão filosófica e lógica do funcionamento da linguagem e de algumas questões éticas e políticas ligadas à ação. São duas subáreas importantes do pensamento filosófico atual.

A seção I é dedicada a temas de lógica e de filosofia da lógica e contém trabalhos que discutem alguns tópicos especiais desse domínio. Assim como no caso das demais seções, os temas aqui tratados permitem ao especialista aprofundar-se nas discussões desses assuntos e ao iniciante ter uma visão de uma parte da temática relevante nessa área hoje.

A seção II trata de temas específicos de ontologia da mecânica quântica, envolvendo também discussões de caráter lógico e epistemológico. Essa é uma área de debates difíceis para os não especialistas, e os textos dessa seção podem então para a compreensão de algumas dessas questões da física contemporânea que tanto têm desafiado a filosofia e a lógica contemporâneas.

A seção III completa as discussões mais específicas sobre a linguagem, com temas de filosofia da linguagem e de fundamentos da lingüística. Há assuntos mais técnicos e outros mais gerais, mas, no conjunto, esses trabalhos também recobrem uma parte importante das questões neste subdomínio que tanto têm auxiliado a filosofia em geral, particularmente a ontologia e a epistemologia, assim como a própria lógica.

A seção IV trata, por fim, de questões éticas e políticas ligadas e, por si só, oferece uma visão representativa, ainda que parcial, das discussões sobre esses domínios de reflexão filosófica sobre a ação humana.

Desta forma, temos certeza que este livro pode contribuir para uma visão variada dos temas de filosofia da linguagem e lógica, útil não apenas para o especialista, mas também para o iniciante, assim como de alguns temas importantes da reflexão filosófica atual sobre a ação humana em suas dimensões éticas e políticas.

Aproveitamos para agradecer a todos os autores por suas contribuições e pela paciência de reverem mais de uma vez seus textos. Agradecemos também imensamente aos organizadores do já mencionado simpósio *Principia*, em especial, aos colegas Alberto Cupani, César Mortari e Jaimir Conte, assim como a toda a equipe por eles montada e que conduziu tão bem os trabalhos do simpósio.

Agradecemos também, em nome da comissão organizadora do evento o apoio institucional da Universidade Federal de Santa Catarina, de seu Centro de

Filosofia e Ciências Humanas, do Departamento de Filosofia e do Programa de Pós-Graduação em Filosofia da mesma instituição. Agradecemos finalmente às agências governamentais CAPES, CNPq, FINEP e FAPESC pelo apoio financeiro ao evento, em especial, no caso desta publicação, à FAPESC, que também viabilizou a publicação do volume *Temas de filosofia do conhecimento*.

Florianópolis, outubro de 2011.

Luiz Henrique de Araújo Dutra  
Alexandre Meyer Luz

LINGUAGEM, ONTOLOGIA  
E AÇÃO

coleção  
**RUMOS DA EPISTEMOLOGIA**

**Editor:** Luiz Henrique de A. Dutra

**Conselho Editorial:** Alberto O. Cupani

Cezar A. Mortari

Décio Krause

Gustavo A. Caponi

José A. Angotti

Luiz Henrique A. Dutra

Marco A. Franciotti

Sara Albieri



nel@cfh.ufsc.br  
(48) 3721-8612

Núcleo de Epistemologia e Lógica  
Universidade Federal de Santa Catarina

www.cfh.ufsc.br/~nel  
fax: (48) 3721-9751

Criado pela portaria 480/PRPG/96, de 2 de outubro de 1996, o NEL tem por objetivo integrar grupos de pesquisa nos campos da lógica, teoria do conhecimento, filosofia da ciência, história da ciência e outras áreas afins, na própria UFSC ou em outras universidades. Um primeiro resultado expressivo de sua atuação é a revista *Principia*, que iniciou em julho de 1997 e já tem doze volumes publicados, possuindo corpo editorial internacional. *Principia* aceita artigos inéditos, além de resenhas e notas, sobre temas de epistemologia e filosofia da ciência, em português, espanhol, francês e inglês. A Coleção Rumos da Epistemologia é publicada desde 1999, e aceita textos inéditos, coletâneas e monografias, nas mesmas línguas acima mencionadas.

# SUMÁRIO

## I – LÓGICA E FILOSOFIA DA LÓGICA

<b>BRUNO RAMOS MENDONÇA</b>	13
<i>Conhecimento Simbólico na Álgebra da Lógica de Boole e Venn</i>	
<b>CARLOS MANHOLI</b>	26
<i>Uma transformação computável de bases de Horn de 1ª ordem em bases de conhecimento vívidas</i>	
<b>CEZAR A. MORTARI</b>	41
<i>Lógicas de Sylvan e implicação estrita</i>	
<b>KLEIDSON ÊGLICIO CARVALHO DA SILVA OLIVEIRA</b>	53
<i>Análise conceitual do quantificador “poucos” e uma lógica gerada por um desses conceitos.</i>	
<b>WALTER GOMIDE</b>	67
<i>Sobre a noção de “estados de coisas contraditórios”: uma aproximação de noções lógicas a noções métrico-topológicas.</i>	

## II – ONTOLOGIA DA MECÂNICA QUÂNTICA

<b>CAROLINE ELISA MURR</b>	85
<i>O realismo científico Schrödingueriano</i>	
<b>DÉCIO KRAUSE</b>	92
<i>A Calculus of Non-Individuals (Ideas for a quantum mereology)</i>	
<b>JONAS RAFAEL BECKER ARENHART</b>	107
<i>Discussões acerca da versão fraca do Princípio da Identidade dos Indiscerníveis em mecânica quântica</i>	
<b>NEWTON DA COSTA E CHRISTIAN DE RONDE</b>	120
<i>On the Physical Representation of Quantum Superpositions</i>	

### III – FILOSOFIA DA LINGUAGEM E FUNDAMENTOS DA LINGÜÍSTICA

<b>CÍCERO ANTÔNIO CAVALCANTE BARROSO</b> <i>A função de código dos nomes próprios</i>	135
<b>DEBORA FONTOURA DE OLIVEIRA</b> <i>Significado e Percepção: uma fala acerca das sensações</i>	151
<b>LOVANIA ROEHRIG TEIXEIRA</b> <i>Indexicalidade no Português Brasileiro: uma análise semântica baseada em mudanças de contexto</i>	164
<b>RICARDO NAVIA</b> <i>Algunas cuestiones sobre la triangulación davidsoniana como respuesta a las dificultades evidenciadas en el argumento sobre “seguir una regla”</i>	180
<b>VALDETONIO PEREIRA DE ALENCAR</b> <i>A Tese da Designação Rígida</i>	191

### IV – ÉTICA E POLÍTICA

<b>DARLEI DALL’AGNOL</b> <i>Princípios Bioéticos e a Atual Lei de Biossegurança Brasileira</i>	203
<b>FERNANDO CÉSAR COSTA XAVIER</b> <i>O “Grupo de Setembro” e o “11 De Setembro”</i>	212
<b>GABRIEL GARMENDIA DA TRINDADE E LAUREN DE LACERDA NUNES</b> <i>A questão dos deveres indiretos para com os animais não-humanos segundo a filosofia moral kantiana</i>	222
<b>JOSÉ CLÁUDIO MORELLI MATOS</b> <i>Respostas ao dilema malthusiano nas conferências sobre evolução e ética de Huxley e Dewey</i>	232
<b>MARIA EUGENIA BUNCHART</b> <i>O Supremo Tribunal Federal e a anencefalia: uma reflexão sobre a legitimidade democrática do Judiciário à luz de Rawls, Habermas e Nino</i>	245
<b>MILENE CONSENSO TONETTO</b> <i>A “Fórmula do Fim em Si” e a objeção de formalismo à filosofia moral de Kant.</i>	267
<b>REJANE SCHAEFER KALSING</b> <i>Sociabilidade e Direito na Obra A Idéia de uma História Universal de um Ponto de Vista Cosmopolita</i>	276

I

---

LÓGICA E

FILOSOFIA DA LÓGICA



# CONHECIMENTO SIMBÓLICO NA ÁLGEBRA DA LÓGICA DE BOOLE E VENN

BRUNO RAMOS MENDONÇA<sup>1</sup>

*Universidade Federal de Santa Maria (UFSM)*

bruno.ramos.mendonca@gmail.com

## 1. Introdução

O objetivo principal desse artigo é apresentar os tratamentos oferecidos por Boole (1815-1864) e Venn (1834-1923), dois importantes algebristas da lógica do século XIX, para o problema do conhecimento simbólico da lógica. Isso é feito em três etapas. Na seção 2 desse trabalho, apresenta-se como o problema do conhecimento simbólico surge para a tradição da álgebra da lógica do século XIX. Na seção 3, apresentam-se as álgebras da lógica de Boole e Venn, e na seção 4 são introduzidos os tratamentos oferecidos por esses dois autores para o problema do conhecimento simbólico da lógica.

## 2. Conhecimento simbólico na álgebra da lógica

Verifica-se na filosofia recente uma reabilitação do problema do conhecimento simbólico. Esse problema, que remonta aos clássicos, é agora tratado com recurso a uma terminologia técnica tal como a que pode ser encontrada em ESQUISABEL (texto inédito). Símbolos são objetos físicos que estão por outras coisas, representando-as em sua ausência. Conhecimento simbólico é o conhecimento que se pode obter sobre as coisas representadas através da inspeção de seus símbolos. Apelar a esse modo de conhecimento, especialmente em disciplinas formais como matemática e lógica, não apenas é útil como por vezes é necessário: ninguém tem dificuldade em somar números pequenos, contudo a maior parte das pessoas precisa recorrer a um sistema simbólico adequado quando se trata de computar grandes quantidades.

Um símbolo possui ao menos as seguintes funções. Em primeiro lugar, o símbolo exerce uma *função subrogativa*, i.e., está por outra coisa. Por exemplo, o numeral 1 está pelo número 1. Além disso, os símbolos podem exercer uma *função ectética*, no sentido de sensualizar a informação representada: a figura de um triângulo, que subroga o conceito de triângulo, torna sensível a propriedade dos triângulos de ter três lados, já que a figura *assemelha-se* a um objeto geométrico

de três lados. Cabe ressaltar, no entanto, que nem todos os símbolos têm a mesma capacidade ectética: a palavra “triângulo”, por exemplo, é menos ectética que a figura de um triângulo. Por fim, os símbolos possuem ainda uma *função psicotécnica* na medida em que seu uso é essencial para o alcance da maior parte dos conhecimentos que possuímos.<sup>2</sup>

Essa breve exposição do conceito geral de “símbolo” permite que se introduza os diferentes aspectos do problema do conhecimento simbólico. Em primeiro lugar, existe uma dificuldade com relação ao conceito de subrogação. Ao longo da história da filosofia, tem-se mostrado difícil apontar, principalmente com relação aos campos mais abstratos das ciências formais, o que se supõe que os símbolos subrogam. Surge então a questão sobre as condições em que é correto assumir que os símbolos subrogam alguma coisa. Sobre isso, pode-se classificar ao menos três possíveis respostas.<sup>3</sup> A primeira resposta é de cunho *realista*: exige-se que os símbolos, se possuem alguma legitimidade na representação de uma ciência, subroguem conceitos daquela ciência. Desse modo, se é correto usar um símbolo na representação da matemática, por exemplo, então é correto atribuir-lhe significado matemático.

Uma resposta alternativa, de recorte *instrumentalista*, defende que, por vezes, alguns símbolos de um sistema simbólico são usados legitimamente sem que subroguem algum conceito da ciência representada. Esses símbolos representam elementos ideais no sistema. Supor elementos ideais pode ser vantajoso de uma perspectiva prática: por exemplo, conhecimentos científicos que seriam difíceis de serem obtidos sem idealização, tornam-se mais simples de serem alcançados. Porém um elemento ideal não pertence ao domínio estudado, e, portanto, sua suposição deve poder ser eliminada sem acarretar em prejuízo para a obtenção de verdades sobre a ciência. Por fim, tem-se a resposta *formalista* que radicaliza a posição instrumentalista. Nessa versão, nenhum símbolo subroga. À matemática, enquanto disciplina simbólica, não corresponde nenhum conjunto de objetos abstratos, pois tudo que existe são os símbolos matemáticos.

Um segundo aspecto do problema do conhecimento simbólico diz respeito à função ectética dos símbolos. Deseja-se compreender em que consiste a função ectética e por que alguns símbolos são mais ectéticos que outros. Esse problema, que está relacionado ao problema da dicotomia gráfico-linguístico que tem ocupado a comunidade filosófica internacional pelo menos nos últimos vinte anos, não nos ocupará nesse trabalho.<sup>4</sup>

O uso de símbolos para o estudo da lógica é historicamente recente. Apenas no século XIX, lógicas simbólicas passaram a ser sistematicamente apresentadas.<sup>5</sup> Esse processo dependeu tanto de revisão da concepção de lógica da época quanto da aproximação da lógica com a matemática. Essa aproximação foi desenvolvida a partir de pelo menos duas perspectivas. Por um lado, lógicos como Frege (1848-1925) e Peano (1858-1932) aproximaram a lógica à *Análise Matemática*. Desenvolveram assim o que ficou conhecido como *Lógica Matemática*. Um segundo grupo de autores, entre os quais encontram-se Boole (1815-1864) e Venn

(1834-1923) aproximaram a lógica à *Álgebra*, e desenvolveram o que se chamou *Álgebra da lógica*.

A Análise é o ramo da matemática que estuda o domínio dos números reais. Entre suas principais subdisciplinas encontram-se, por exemplo, a álgebra elementar, o cálculo integral e diferencial, as equações diferenciais etc.<sup>6</sup> A lógica matemática foi um dos resultados do processo histórico de rigorização da Análise. Matemáticos do século XIX, buscando dar maior precisão às provas, conceitos e princípios básicos da Análise, num determinado momento passaram a fazer uso de um instrumental lógico no tratamento de seus fundamentos. Esse instrumental é a lógica matemática: lógica entendida como uma disciplina de fundamentos da matemática, através da qual é possível investigar e explicar noções básicas da matemática.<sup>7</sup>

Compreensão bastante diversa da relação entre matemática e lógica pode ser encontrada nos algebristas da lógica. Para esses autores, a lógica é uma disciplina ao lado da matemática sujeita a tratamento matemático algébrico. Essa concepção foi possibilitada pelo desenvolvimento da noção de *Álgebra Abstrata* na Inglaterra do século XIX.

Como aponta DURAND-RICHARD (2000), o contexto acadêmico do século XIX na Inglaterra foi marcado por reformas do currículo universitário. Acadêmicos como G. Peacock (1798-1851), C. Babbage (1791-1871) e F. W. Herschel (1792-1871), entre outros, criaram sociedades de pesquisa e promoveram uma revolução na investigação científica, especialmente em matemática, que se fazia na Inglaterra. O sentido dessas reformas foi eminentemente prático. Com a sociedade inglesa gradualmente passando de uma economia rural a ser uma economia industrial, mostrou-se necessário reconsiderar que tipo de conhecimento seria útil para o futuro.<sup>8</sup>

Certamente o conhecimento matemático possui um interesse prático. O problema é como atribuir sentido a certas áreas da matemática que, apesar de úteis, parecem carecer de significado. Como garantir que os sistemas simbólicos usados na representação dessas áreas da matemática não são meros jogos simbólicos, ainda que úteis? A solução tradicional passava por restringir os artifícios simbólicos àquelas áreas da matemática cujo significado não é problemático. Contudo, para solucionar esse problema, autores como o já mencionado Peacock propuseram uma maneira de legitimar o uso de símbolos matemáticos mesmo em domínios com significado nebuloso.

A álgebra é uma disciplina matemática que trata de propriedades gerais de domínios de entes. Em álgebra é possível, por exemplo, indicar propriedades gerais da soma de quaisquer dois números reais “x” e “y”. A solução de Peacock consiste em propor a existência de duas álgebras. Por um lado, existe a álgebra aritmética, “em que os símbolos são gerais em sua forma, mas específicos em seu valor”.<sup>9</sup> Ou seja, a álgebra aritmética é um exemplo importante de sistema algébrico em que as variáveis percorrem um domínio específico de entes. Por outro lado, existe a álgebra simbólica, em que as variáveis não percorrem um domínio específico de valores.<sup>10</sup> Segundo Peacock, a álgebra simbólica é sugerida

pela álgebra aritmética, mas sua verdade é mais fundamental que a verdade desta.<sup>11</sup> Em resumo, a álgebra simbólica é vista como um sistema universal de matemática, cujos princípios podem ser instanciados nos diversos domínios matemáticos: o domínio dos números naturais, dos reais, da geometria... e da lógica. Cabe notar que nessa perspectiva a lógica não é uma disciplina com relevância especial frente a outras disciplinas matemáticas, consistindo apenas em mais uma disciplina sujeita a tratamento matemático.

Foi Boole quem, dentre os algebristas do século XIX, pela primeira vez aplicou a álgebra e a notação algébrica ao estudo da lógica e desenvolveu uma álgebra da lógica. Seu trabalho foi seguido pelo de lógicos como J. Venn (1834-1923), E. Schröder (1841-1902), entre outros. Mas a concepção da lógica como uma disciplina sujeita a tratamento matemático implica num problema filosófico dos mais difíceis: que relação se mantém entre matemática e lógica? A lógica é uma disciplina matemática?

A resposta a essa questão determina a concepção de lógica proposta por esses autores. FREGE (2009), por exemplo, apresentou uma crítica que se tornou clássica contra os algebristas da lógica:

(...) meu objetivo era diferente do de Boole. Não era meu desejo apresentar uma lógica abstrata através de fórmulas, mas expressar um conteúdo mediante sinais escritos de maneira mais clara e precisa do que seria possível por palavras. Com efeito, desejava produzir não um mero *calculus ratiocinator*, mas uma *lingua characterica* em sentido leibniziano. (FREGE, 2009, p. 68)

Quando consideramos a linguagem formular de Boole como um todo, verificamos que ela se resume a vestir a lógica abstrata com uma roupagem de sinais algébricos. Ela não é adequada para veicular um conteúdo, e também não é esta sua finalidade. Mas esta [veicular um conteúdo] é exatamente a minha intenção (...)

Para tal objetivo, não podia empregar a notação de Boole; pois é inviável ter na mesma fórmula, por exemplo, o sinal + ocorrendo em parte em sentido lógico, e em parte em sentido aritmético. A analogia entre os processos do cálculo lógico e aritmético, de grande valia para Boole só pode trazer equívocos, caso sejam associados. A linguagem por sinais de Boole só é pensável quando inteiramente separada da aritmética. (ibid., p. 72)

Verifica-se assim que entre as duas tradições de simbolização da lógica do século XIX apresenta-se uma diferença a respeito da concepção que se faz da disciplina. Frege decodifica essa diferença apelando à distinção clássica feita por Leibniz entre “lógica como cálculo” e “lógica como linguagem universal” (ou, talvez, apelando a uma interpretação dessa distinção).<sup>12</sup> Além disso, tal diferença dever-se-ia a uma confusão, feita pelos algebristas da lógica, entre o conteúdo matemático dos símbolos algébricos e seu conteúdo lógico.

Ademais, o modo como se entende a relação entre matemática e lógica determina a concepção de conhecimento simbólico da lógica que se pode alcançar por meio de um simbolismo matemático. Nesse trabalho apresentam-se as

respostas a essas questões, propostas pelos algebristas da lógica Venn e Boole. A álgebra da lógica de Boole foi o primeiro sistema de álgebra da lógica proposto dentro dessa tradição. Não se deve confundir sua álgebra com o que hoje conhece-se por álgebra booleana. A álgebra de Boole é um sistema mais primitivo de álgebra da lógica, do qual a álgebra booleana é uma modificação. A álgebra de Boole será, sem grandes modificações, defendida no principal livro de lógica simbólica de Venn, *Symbolic Logic* (1881). No entanto, apesar da semelhança técnica entre os sistemas de Boole e Venn, há uma diferença na interpretação desses sistemas. Enquanto para Venn todos os símbolos que compõem essa representação subrogam conceitos lógicos, para Boole alguns desses símbolos não o fazem.

### 3. Álgebras da lógica de Boole e Venn

Respectivamente, os principais livros de lógica simbólica de Boole e Venn são *An Investigation of The Laws of Thought* (1858) e *Symbolic Logic* (1881). Enquanto a obra de Venn compõe-se de uma historiografia da lógica e de uma exposição de um sistema próprio de álgebra da lógica, a obra de Boole apresenta exclusivamente um sistema de álgebra da lógica e uma teoria lógica da probabilidade.

As apresentações dos sistemas de álgebra da lógica nas duas obras seguem a mesma sistematização. Num primeiro momento apresentam-se o vocabulário e as regras de composição de sentenças logicamente bem formadas. O restante da apresentação concentra-se na introdução das regras de inferência para argumentos válidos. Esse artigo centrar-se-á na exposição da primeira parte desses sistemas – cobrindo o conteúdo de, basicamente, os capítulos 2 e 3 da obra de Venn e do capítulo 2 de Boole.

O que se estudava tradicionalmente como lógica na Inglaterra do começo do século XIX é o que podemos chamar de lógica em sentido amplo: estudava-se a lógica Silogística de Aristóteles, mas, principalmente, estudavam-se temas que hoje seriam classificados como sendo de teoria do conhecimento e filosofia da ciência. A lógica era entendida como uma arte para descobrir-se verdades e era identificada com a metodologia científica.<sup>13</sup> Foi apenas a partir da reforma dos currículos universitários na Inglaterra do século XIX que se revisou o que era estudado e ensinado como lógica. A concepção tradicional da disciplina passou a ser aquela que poderia ser encontrada num manual didático como o de R. Whately (1787-1863), *Elements of Logic*. A obra de Whately, que foi adotada por toda Inglaterra e também nos Estados Unidos, acabou por tornar-se especialmente famosa, recebendo diversas reedições.<sup>14</sup> Essa obra não é especialmente inovadora. Trata-se antes de uma defesa da concepção tradicional de lógica, que restringe o campo da disciplina ao invés de apresentar-lhe novos objetos de estudo.

Em resumo, a teoria da lógica introduzida em *Elements of Logic* possui as seguintes características. Por um lado, a lógica é entendida principalmente como uma ciência formal semelhante à álgebra matemática.<sup>15</sup> Logo, apenas em sentido

derivado a lógica deve ser considerada uma arte. Além disso a lógica de *Elements of Logic* restringiu-se ao estudo da Silogística aristotélica,<sup>16</sup> recusando como tópicos legítimos de estudo temas que hoje seriam considerados como de epistemologia ou de teoria da ciência. Assim, temas como o da indução, então considerados legitimamente lógicos, foram abandonados na exposição de Whately.<sup>17</sup> Essa escolha deve-se à compreensão da lógica como uma disciplina essencialmente formal. Dado que a lógica estuda *apenas* os aspectos formais da argumentação, é inadequado exigir que ela tematize fenômenos específicos da argumentação e, do ponto de vista formal, a argumentação em geral reduzir-se-ia à silogística.

A concepção da lógica como uma ciência formal, da qual a silogística é uma parte muito importante, é subjacente aos sistemas de álgebra da lógica de Venn e Boole. Na introdução de *Symbolic Logic*, por exemplo, Venn adota uma posição conciliadora no conflito entre as novas lógicas simbólicas e a lógica tradicional. Para Venn, a lógica tradicional é insubstituível em importância educacional:

Perfeita clareza de apreensão em todos esses pontos [da lógica tradicional] parece essencial para exatidão do pensamento, e é difícil encontrar qualquer outro meio melhor de adquirir essa clareza do que o estudo de alguns dos manuais padrões de lógica.<sup>18</sup>

Segundo Venn, o que a lógica simbólica propõe é uma “generalização” da lógica tradicional. Por exemplo, enquanto a lógica tradicional estuda silogismos, argumentos com um “termo médio”, na lógica simbólica seria possível estudar inferências com um número variado de termos médios.<sup>19</sup>

Verifica-se, portanto, que elementos estruturais da lógica silogística são preservados na álgebra da lógica de Boole e Venn. Em primeiro lugar, a forma básica de uma proposição, para esses autores, é “A é B”. Algebricamente, qualquer proposição seria expressa por uma equação “A = B”. Assim como para Aristóteles, na álgebra da lógica de Boole e Venn, toda proposição expressa uma relação entre o significado de dois termos, representada pela *cópula* “ser”.<sup>20</sup> Além disso, os termos são entendidos como termos para classes, nomes de conjuntos de coisas.

Para GEACH (1972), a principal característica e o principal defeito da tradição aristotélica de análise proposicional é não fazer uma distinção radical entre sujeito e predicado, tal como a distinção encontrada em Frege entre função e argumento. A lógica dos *Analíticos Anteriores* permite que o sujeito de uma frase sempre possa ser o predicado de outra.<sup>21</sup> Essa crítica vale igualmente para a lógica algébrica de Boole e Venn: uma frase de identidade “A = B” tem o mesmo significado de “B = A”, e os termos “B” e “A” exemplificam o mesmo tipo lógico.

“Termo para classe” é um conceito bastante amplo. Não apenas os substantivos comuns e próprios são termos para classe, mas também adjetivos, e verbos (p. ex., ‘trabalhar’) ao serem transformados em adjetivos (‘ser trabalhador’). Termos para classe podem, assim, representar tanto conjuntos de um único membro (o conjunto ‘Sócrates’), ou nenhum membro (o conjunto vazio ‘0’, p. ex.), ou conjuntos com vários membros, ou conjuntos com todos os membros (o conjunto universo ‘1’, p. ex.).<sup>22</sup>

Levando-se em conta, no entanto, que termos possuem tanto *extensão* quanto *intensão*, deve-se perguntar se as regras que atuam sobre esse sistema simbólico são de natureza intensional ou extensional. A resposta de Boole é a que se segue:

A questão não é de muita importância aqui, já que sua decisão não pode afetar as leis de acordo com as quais os signos são empregados. Eu considero, contudo, que a resposta geral a essa e outras questões semelhantes é que, no processo do raciocínio, signos estão no lugar e exercem o ofício dos conceitos e operações da mente; mas como aqueles conceitos e operações representam coisas, e as conexões e relações das coisas, então signos representam coisas com suas relações e conexões; e, por fim, como signos estão no lugar dos conceitos e operações da mente, eles estão sujeitos às leis daqueles conceitos e operações.<sup>23</sup>

Ou seja, para Boole, tanto a leitura extensional quanto intensional das operações lógicas são adequadas. Contrariamente, Venn não pensa assim. Esse será o foco de sua crítica à álgebra da lógica de Jevons: a leitura intensional do simbolismo não é capaz de explicar as características específicas das proposições de existência, por exemplo.<sup>24</sup>

Além de identidades entre duas classes, podem ser representadas nessa álgebra da lógica, também, identidades entre classes compostas, como em “ $A \text{ e } B = B \text{ e } C$ ”, formadas a partir da aplicação de operações lógicas sobre conjuntos de classes. Segue-se uma breve apresentação dessas operações. Deve-se dizer em primeiro lugar que essas operações são semelhantes às operações lógicas clássicas de conjunção, disjunção, etc. Uma peculiaridade é que sua aplicação sobre classes não gera como resultado proposições, mas classes compostas.<sup>25</sup>

Em primeiro lugar, pode-se construir conjuntos de classes através da operação de *restrição*, análoga à operação lógica clássica de conjunção. Essa operação, que, devido às semelhanças formais com a operação matemática de multiplicação, é representada na álgebra da lógica de Boole e Venn pelo signo “ $\times$ ”, tem suas condições de interpretação apresentadas na tabela (1) abaixo.<sup>26</sup>

(1)

A	B	$A \times B$
1	1	1
0	1	0
1	0	0
0	0	0

As semelhanças entre as operações da álgebra da lógica e as correspondentes – do ponto de vista simbólico – operações matemáticas possuem limites. Um caso clássico é o da comparação entre as operações de restrição e multiplicação. Enquanto a operação lógica de restrição está sujeita à “lei da repetição”, a operação matemática de multiplicação não está. Em termos algébricos, a lei da repetição

regula que “ $y \times y = y$ ” para qualquer valor “ $y$ ”. Isso claramente vale para os valores “0” e “1”, mas não vale numa álgebra que prevê outros valores.

Além da operação de restrição, pode-se definir a operação de *agregação*. Essa operação é semelhante à operação lógica clássica de *disjunção*. Na álgebra da lógica de Boole e Venn, a operação de agregação, representada pelo signo “+” devido a uma semelhança formal com a operação matemática de adição, exige uma condição especial de aplicação: a operação não pode ser empregada em proposições nas quais as classes agregadas possuem elementos em comum. Por exemplo, se acerca da classe composta “ $A + B$ ” há uma situação em que é correto atribuir o valor 1 tanto à classe designada por “ $A$ ” quanto à classe designada por “ $B$ ”, então essa classe composta não pode ser expressa dessa forma. A expressão deve ser reformulada de modo que as classes agregadas não possuam membros em comum. Isso pode ser feito, reformulando-a seja como “ $A + (B \times \text{não-}A)$ ”, ou como “ $B + (A \times \text{não-}B)$ ”. As condições de interpretação da expressão reescrita eliminam os casos inconvenientes. Além disso, essa expressão possui o mesmo sentido da expressão original. Isso pode ser verificado na tabela (2) abaixo:

(2)

A	B	Não-A	$B \times \text{não-}A$	$A + B$	$A + (B \times \text{não-}A)$
1	1	0	0	1	1
0	1	1	1	1	1
1	0	0	0	1	1
0	0	1	0	0	0

Na álgebra da lógica de Boole e Venn existem ainda operações *inversas* às operações lógicas já introduzidas. A operação inversa da agregação é chamada *exclusão* e é representada pelo signo “-”, o sinal de subtração sob uma interpretação matemática. Sua interpretação também respeita condições especiais: em “ $A - B$ ” exige-se que o conjunto designado por “ $B$ ” tenha sua extensão contida na extensão de “ $A$ ”. A operação de exclusão tem suas condições de interpretação apresentadas na tabela (3) abaixo:

(3)

C	A	$C \times A$	$C - (C \times A)$
1	1	1	0
0	1	0	0
1	0	0	1
0	0	0	0

A operação inversa da restrição é a “divisão lógica”, representada algebricamente por “ $\div$ ”. Um importante debate foi travado entre os algebristas da lógica a respeito da significação lógica da divisão. Para alguns dos algebristas, o símbolo “ $\div$ ” não subroga nenhum conceito lógico, enquanto para autores como Venn esse símbolo possui significado. A interpretação da divisão lógica é apresentada na tabela (4) abaixo:

(4)

C	A	$C \div A$
1	1	1
0	1	0
1	0	0
0	0	$u$

Se a operação “ $\times$ ” restringe a extensão de duas classes ao que elas possuem em comum, a operação “ $\div$ ” elimina a restrição de uma classe sobre outra: “ $C \div A$ ” retira a restrição de “A” sobre uma classe “C”. O resultado é a geração de uma classe que possui ou uma extensão comum a “C” e “A” (1ª linha) ou uma extensão indefinida  $u$  que não é nem “C” nem “A” (4ª linha). Em termos algébricos, a classe resultante é “ $C \times A + u(\text{não-C} \times \text{não-A})$ ”.

#### 4. Subrogação na álgebra da lógica de Boole e Venn

Os algebristas da lógica disputavam sobre a legitimidade lógica da operação de divisão porque não chegavam a um consenso sobre a aceitabilidade de uma classe indefinida extensionalmente, i.e., uma classe essencialmente indefinida, sobre a qual não se pode dizer se ela possui extensão com nenhum, algum, ou todos os membros.

Boole está entre os autores que não atribuem significado lógico à operação de divisão exatamente por essa razão.<sup>27</sup> Sua posição, no entanto, é controversa, dado que seu sistema de lógica faz uso extensivo da operação de divisão. Aparentemente, Boole estaria envolvido com uma série de contradições em seu sistema de lógica. VAN EVRA (1977) porém rechaça essas críticas através de uma análise apurada da exposição de Boole de sua álgebra da lógica.

De acordo com VAN EVRA, Boole faz uso da operação de divisão apenas na prova de validade de argumentos, nunca na expressão de proposições logicamente bem formadas. A razão para isso é simples: enunciados logicamente bem formados não podem conter símbolos sem significado lógico, como o símbolo de divisão. Mas seu método de prova de validade de argumentos, pelo contrário, é somente um “instrumento” para efetuar cálculos lógicos, que, como tal, pode fazer uso de expressões sem significado lógico.

Além disso, Boole considera ser característico de um método simbólico de prova de validade de argumentos que ele faça uso de procedimentos aos quais não corresponde nenhum princípio lógico.<sup>28</sup> Na álgebra da lógica de Boole, por exemplo, existem princípios inferenciais cuja validade é puramente simbólica, i.e., não importa se os termos não-lógicos representam classes ou números para que o princípio possua vigência. Em suma, a operação de divisão lógica só tem lugar numa das partes da apresentação de seu sistema de lógica simbólica, a saber, na teoria da inferência. Isso porque apenas nesse momento do sistema

Boole considera aceitável o uso de artifícios de representação puramente simbólicos. Um símbolo que não possui significado lógico deve poder ser eliminável da prova de validade de um argumento e não pode estar contido nem nas premissas nem na conclusão.<sup>29</sup>

Venn por sua vez não aceita o uso de símbolos que não possuam significado. Se um símbolo não subroga nenhuma noção lógica, então não é legítimo usá-lo na representação dessa disciplina<sup>30</sup>. No entanto, esse não seria o caso do símbolo de divisão lógica. A operação de divisão é simplesmente inversa à operação de restrição: o uso de ambas seria igualmente legítimo.

Venn sugere que é com base em razões práticas que se determina, de um par de operações inversas, qual é a operação direta e qual é a inversa. Por vezes, a operação inversa é a mais difícil de ser executada.<sup>31</sup> Geralmente, uma operação é considerada inversa por oferecer resultados indeterminados, enquanto seu par oposto gera apenas resultados determinados.<sup>32</sup> Esse seria o caso da operação de divisão lógica: seu resultado é a geração de uma classe ao menos parcialmente indeterminada do ponto de vista extensional. Contudo, nenhuma dessas desvantagens práticas retira o *status* lógico da operação de divisão lógica.<sup>33</sup>

## 5. Conclusão

O debate sobre a significação lógica do símbolo “÷” travado entre os algebristas da lógica do século XIX é um exemplo histórico do debate em torno do problema de conhecimento simbólico. Verifica-se que Boole e Venn representam posições paradigmáticas nessa discussão. Boole defende uma solução tipicamente instrumentalista para o problema: certos símbolos representam elementos puramente ideais no sistema representacional. O apelo a esses elementos está pragmaticamente justificado, mas eles não podem ser confundidos com elementos reais do sistema e devem poder ser eliminados sem prejuízo para a obtenção de verdades sobre a ciência. Venn, por sua vez, representa a posição realista no debate: um sistema simbólico pode ser utilizado na representação de uma teoria científica desde que seus símbolos, sem exceção, subroguem conceitos daquela ciência.

Essas concepções resultam em opiniões diferentes sobre os vínculos entre matemática e lógica. Para Boole, lógica e matemática seriam ciências vinculadas na medida em que ambas respeitam princípios em comum de uma disciplina racional de ordem superior.<sup>34</sup> Venn, por sua vez, reconhece apenas uma analogia formal entre as duas ciências, que, por acaso, permite que elas sejam representadas semelhantemente.<sup>35</sup> Todos os princípios que vigoram na teoria lógica são lógicos.

## Referências

- AUDI, R. (Ed.) *Cambridge Dictionary of Philosophy*, Cambridge: Cambridge University Press, 1999.
- BOOLE, G. *An Investigation of the Laws of Thought*. New York: Dover Publications, 1958.
- DURAND-RICHARD, M-J. “Logic versus Algebra: English Debates and Boole’s Mediation”. In: *A Boole Anthology*, Netherlands: Kluwer Academic Publishers, 2000, pp. 139-166.
- ESQUISABEL, O. “Representing and Abstracting: an Analysis of Leibniz’s Concept of Symbolic Knowledge”. In: LASSALLE CASANAVE, A. et al. *Symbolic Knowledge from Leibniz to Husserl* (em preparação).
- FREGE, G. “Sobre a Finalidade da Conceitografia (1882-1883)”, Trad. Paulo Alcoforado. In: ALCOFORADO, P. (Org.) *Lógica e Filosofia da Linguagem*. São Paulo: EDUSP, 2009. pp. 67-80.
- GEACH, P. T. “History of the Corruptions of Logic”. In: *Logic Matters*. Berkeley: University of California Press, 1980. pp. 44-61.
- GRATTAN-GUINNESS, I. “Mathematical Turns in Logic”. In: *Handbook in the History of Logic*, Vol. 3, Amsterdam: Elsevier, 2004, pp. 545-556.
- \_\_\_\_\_. *The Search for Mathematical Roots 1870-1940*. Princeton: Princeton University Press, 2000.
- JAQUETTE, D. “Boole’s Logic”. In: *Handbook in History and Philosophy of Logic*, Vol. 4, Amsterdam: Elsevier, 2008, pp. 331-379.
- LASSALLE CASANAVE, A. “Avatars of Symbolic Knowledge in Kant”. In: LASSALLE CASANAVE, A. et al. *Symbolic Knowledge from Leibniz to Husserl* (em preparação).
- PECKHAUS, V. “Calculus Ratiocinator versus Characteristica Universalis? The Two Traditions in Logic, Revisited”. In: *History and Philosophy of Logic*, Vol. 25, 2004.
- SHIMOJIMA, A. “The Grafic Linguistic Distinction: Exploring Alternatives”. In: *Artificial Intelligence Review*, Netherlands: Kluwer Academic Publishers, 2001, pp 5-27.
- VAN EVRA, J. “A Reassessment of George Boole’s Theory of Logic”. In: *Notre Dame Journal of Formal Logic*, Vol. 18, nº 3, 1977, pp. 363-377.
- \_\_\_\_\_. “Richard Whately and logical Theory”. In: *Handbook in History and Philosophy of Logic*, Vol. 4, Amsterdam: Elsevier, 2008a, pp. 75-91.
- \_\_\_\_\_. “John Venn and Logical Theory”. In: *British Logic in the Nineteenth Century*, Vol. 4, Amsterdam: Elsevier, 2008b, pp. 507-513.
- VENN, J. *Symbolic Logic*. London: Macmillan, 1881.

## Notas

<sup>1</sup> Mestrando pelo Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Maria (PPGF - UFSM), sob orientação do prof. Dr. Frank Thomas Sautter. Bolsista pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

<sup>2</sup> ESQUISABEL (texto inédito) pp. 25-31.

<sup>3</sup> Uma apresentação dessas respostas pode ser encontrada em LASSALLE CASANAVE (texto inédito, pp. 4-6).

<sup>4</sup> Cf. SHIMOJIMA (2001) para uma exposição detalhada das diversas posições nesse debate.

<sup>5</sup> O primeiro uso da expressão “lógica simbólica” deu-se no livro *Symbolic Logic* (1881) de VENN, que será estudado nesse trabalho. Cf. GRATTAN-GUINNESS (2004, p. 545).

<sup>6</sup> AUDI (1999, “Mathematical Analysis”, pp. 540-1).

<sup>7</sup> GRATTAN-GUINNESS, op. cit., 546.

<sup>8</sup> DURAND-RICHARD (2000, p. 141).

<sup>9</sup> “... where symbols are general in their form, but specific in their value”. (PEACOCK, 1833 apud Ibid., p. 156).

<sup>10</sup> DURAND-RICHARD, loc. cit.

<sup>11</sup> DURAND-RICHARD, loc. cit.

<sup>12</sup> Cf. PECKHAUS (2004) para uma apresentação do modo como essa distinção foi recebida pelos lógicos do século XIX.

<sup>13</sup> VAN EVRA (2008a, pp. 78-80)

<sup>14</sup> Ibid., pp. 77-8.

<sup>15</sup> Ibid., p. 84

<sup>16</sup> Ibid., p. 85.

<sup>17</sup> Ibid., p. 87.

<sup>18</sup> “Perfect clearness of apprehension on all these points seems essential to accuracy of thought, and it is difficult to find any better means of acquiring this clearness than the study of some of the ordinary logical manuals”. (VENN, 1881, p. xxv)

<sup>19</sup> Ibid., p. xx.

<sup>20</sup> Ibid., p. 29-31; JAQUETTE (2008, p. 333).

<sup>21</sup> GEACH (1972, p. 47).

<sup>22</sup> BOOLE (1858, pp. 27-28).

<sup>23</sup> “The question is not of great importance here, as its decision cannot affect the laws according to which signs are employed. I apprehend, however, that the general answer to this and such like questions is, that in the processes of reasoning, signs stand in the place and fulfil the office of the conceptions and operations of the mind; but that as those conceptions and operations represent things, and the connexions and relations of things, so signs represent things with their connexions and relations; and lastly, that as signs stand in the place of the conceptions and operations of the ind, they are subject to the laws of those conceptions and operations” (Ibid., p. 26).

<sup>24</sup> GRATTAN-GUINNESS (2000, p. 61).

<sup>25</sup> VENN (1881, p. 32). A exposição que se segue das operações lógicas *restrição*, *agregação* e *exclusão* é fiel à exposição de Venn e Boole dessas operações nos capítulos 2 de suas respectivas obras. Venn, no capítulo 3 de *Symbolic Logic* trata especificamente da operação de divisão lógica. Já Boole, que como se verá a seguir não considera à divisão lógica uma operação lógica legítima, não oferece tratamento lógico a essa operação.

<sup>26</sup> A atribuição de 1 a uma das classes significa que a extensão daquela classe pertence ao universo das coisas que existem. A atribuição de 0 a uma das classes significa que aquela classe não possui membros, i.e., é coextensional com a classe vazia.

<sup>27</sup> VAN EVRA (1977, p. 373).

<sup>28</sup> VAN EVRA (1977, p. 367).

<sup>29</sup> *Ibid.*, p. 368.

<sup>30</sup> VENN (1881, pp. xiii-xiv).

<sup>31</sup> *Ibid.*, p. 70, em nota.

<sup>32</sup> *Ibid.*, pp. 69-70.

<sup>33</sup> *Ibid.*, pp. 71-73.

<sup>34</sup> GRATTAN-GUINNESS (2000, p. 53-4).

<sup>35</sup> VENN (1881, p. x).

# UMA TRANSFORMAÇÃO COMPUTÁVEL DE BASES DE HORN DE 1ª ORDEM EM BASES DE CONHECIMENTO VÍVIDAS

CARLOS MANHOLI

*Professor do Departamento de Filosofia da Universidade Estadual de Londrina – UEL*  
manholi@uel.br

**Resumo:** O presente trabalho está organizado do seguinte modo: começamos apresentando a noção de uma base de conhecimento vívida, e demonstramos dois teoremas relacionados com as propriedades de tais bases. O teorema 1 afirma que as bases vívidas são consistentes e completas, e o teorema 2 afirma que as consequências lógicas de tais bases podem ser computadas em tempo polinomial. Isto feito, tecemos alguns comentários sobre a baixa expressividade das bases vívidas, e sobre como essa desvantagem resulta das mesmas características de tais bases que as tornam computacionalmente tratáveis. Na sequência, introduzimos uma classe de bases de conhecimento mais expressivas que as bases vívidas, às quais chamamos de bases de Horn de 1ª ordem restritas, bem como uma transformação T capaz de obter uma base vívida de uma qualquer dessas bases. Depois, provamos dois teoremas relativos às propriedades dessa transformação. O corolário 4 – que resulta imediatamente do teorema 3 – afirma que a transformação T obtém uma base vívida a partir de uma base de Horn de 1ª ordem restrita sem que haja perda semântica, e o teorema 5 afirma que T é computável. Por fim, mostramos que o algoritmo capaz de computar T, apresentado na prova do teorema 5, tem todavia complexidade exponencial, e terminamos apresentando um outro algoritmo capaz de computar T, e que, embora também possua complexidade exponencial, se mostra preferível ao anterior em algumas situações.

**Palavras-chave:** bases de conhecimento, expressividade, computabilidade, complexidade computacional.

Em seu livro *Knowledge representation and reasoning*, R. Brachman e H. Levesque definem o conceito de uma base de conhecimento vívida como um conjunto  $\Gamma$  de fórmulas de uma linguagem  $L(S)$  qualquer para o cálculo de predicados de 1ª ordem clássico com igualdade – construída com base em um conjunto de símbolos  $S$  sem ocorrências de símbolos funcionais – que contém um conjunto finito  $\Delta$  de fórmulas atômicas fechadas sem igualdade, que vamos chamar de *núcleo* da base de conhecimento vívida, e satisfaz as seguintes condições: a) se uma fórmula atômica fechada sem igualdade  $\alpha \in L(S)$  é tal que  $\Gamma \models \alpha$ , então  $\sim \alpha \in \Gamma$ ; b) toda fórmula  $\sim \bar{c}_i = \bar{c}_j \in \Gamma$ , para quaisquer constantes distintas  $\bar{c}_i, \bar{c}_j \in S^1$ ; c) a fórmula  $\forall x (x = \bar{c}_1 \vee x = \bar{c}_2 \vee \dots \vee x = \bar{c}_n) \in \Gamma$ , sendo que  $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n$  é uma sequência de todas as constantes individuais em  $S^2$ .

Portanto,  $\Delta$  contém apenas fórmulas capazes de representar proposições atômicas, que afirmam que os objetos de uma dada sequência estão em uma

dada relação, e  $\Gamma$  contém essas fórmulas, e mais a negativa de qualquer fórmula atômica que não seja uma consequência lógica de  $\Gamma$ , fórmulas que afirmam que constantes individuais distintas denotam indivíduos distintos, e uma fórmula que afirma que qualquer indivíduo  $x$  no domínio de um modelo para  $\Gamma$  é denotado por alguma das constantes individuais em  $S$ . É fácil verificar que  $\Gamma$  é consistente, e completo no sentido de que, para qualquer fórmula fechada  $\alpha \in L(S)$ ,  $\Gamma \models \alpha$  ou  $\Gamma \models \sim \alpha$ , conforme segue<sup>3</sup>:

TEO. 1: Seja  $\Gamma$  uma base de conhecimento vívida com núcleo  $\Delta$ ,  $L(S)$  a linguagem formal de  $\Gamma$ , e  $\alpha$  uma fórmula fechada qualquer de  $L(S)$ ; segue-se que  $\Gamma$  é consistente, e completo no sentido de que  $\Gamma \models \alpha$  ou  $\Gamma \models \sim \alpha$ .

Prova: i)  $\Gamma$  é consistente: suponha-se que  $\Gamma$  tenha sido obtido a partir de  $\Delta$  da seguinte maneira: a) acrescenta-se a  $\Delta$  a fórmula  $\forall x (x = \bar{c}_1 \vee x = \bar{c}_2 \vee \dots \vee x = \bar{c}_n)$ , para uma sequência  $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n$  de todas as constantes individuais em  $S$ , resultando o conjunto  $\Delta'$ ; b) acrescenta-se a  $\Delta'$  cada fórmula  $\sim \bar{c}_i = \bar{c}_j$ , para todo par  $(\bar{c}_i, \bar{c}_j)$  de constantes individuais distintas de  $S$ , resultando o conjunto  $\Gamma_0$ ; c) toma-se uma sequência  $\alpha_1, \alpha_2, \dots$ , de todas as fórmulas atômicas fechadas de  $L(S)$  - exceção feita para fórmulas da forma  $\bar{c}_i = \bar{c}_j$ , para todo par  $(\bar{c}_i, \bar{c}_j)$  de constantes individuais de  $S$  - e acrescenta-se a fórmula  $\sim \alpha_k$  a  $\Gamma_k$ , caso  $\Gamma_{k-1} \models \alpha_k$ , para todo

ordinal finito  $k$ ; vamos mostrar que  $\Gamma = \Gamma_\omega = \bigcup_{i < \omega} \Gamma_i$ : 1)  $\Delta \subset \Gamma_\omega$  por construção; 2) a fórmula  $\sim \bar{c}_i = \bar{c}_j \in \Gamma_\omega$  por construção, para todo par  $(\bar{c}_i, \bar{c}_j)$  de constantes individuais distintas de  $S$ ; 3) a fórmula  $\forall x (x = \bar{c}_1 \vee x = \bar{c}_2 \vee \dots \vee x = \bar{c}_n) \in \Gamma_\omega$  por construção, para uma sequência  $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n$  de todas as constantes individuais em  $S$ ; 4) considere-se a estrutura  $E = (D_E, I_E)$ , tal que: a) para qualquer fórmula atômica fechada  $\Phi^m \bar{c}_1 \dots \bar{c}_m \in L(S)$ , a  $n$ -upla  $(I_E(\bar{c}_1), \dots, I_E(\bar{c}_m)) \in I_E(\Phi^m)$  sse  $\Phi^m \bar{c}_1 \dots \bar{c}_m \in \Delta$ ; b)  $I_E(\bar{c}_i) \neq I_E(\bar{c}_j)$  para todo par  $(\bar{c}_i, \bar{c}_j)$  de constantes individuais distintas de  $S$ ; c) para todo  $x \in D_E$ ,  $x = I_E(\bar{c})$ , para uma dada constante individual  $\bar{c}$  de  $S$ ; claramente,  $E \models \Gamma_0$ , e  $E \models \alpha_1$  sse  $\alpha_1 \in \Delta$ ; agora, suponha-se que  $E \models \Gamma_k$ , e que  $E \models \alpha_{k+1}$  sse  $\alpha_{k+1} \in \Delta$ ; obviamente,  $E \models \Gamma_{k+1}$ : de fato, se  $\alpha_{k+1} \in \Delta$ , então  $\Gamma_k \models \alpha_{k+1}$  e  $\Gamma_{k+1} = \Gamma_k$ , e portanto  $E \models \Gamma_{k+1}$  *ex hypothesis*; já se  $\alpha_{k+1} \notin \Delta$ , temos por hipótese que  $E \models \sim \alpha_{k+1}$ , e daí que  $\Gamma_k \models \sim \alpha_{k+1}$ , e logo que  $\Gamma_{k+1} = \Gamma_k \cup \{\sim \alpha_{k+1}\}$ ; como  $E \models \sim \alpha_{k+1}$ , temos que  $E \models \sim \alpha_{k+1}$ ; daí e do fato de que  $E \models \Gamma_k$ , temos que  $E \models \Gamma_{k+1}$ ; agora, suponha-se que  $\alpha_{k+2} \in \Delta$ ; obviamente,  $E \models \alpha_{k+2}$ ; já se  $\alpha_{k+2} \notin \Delta$ , temos que  $E \models \sim \alpha_{k+2}$ ; segue-se que  $E \models \sim \alpha_{k+2}$  sse  $\alpha_{k+2} \in \Delta$ ; assim, temos por indução matemática sobre  $i$  que  $E \models \Gamma_i$ , e que  $E \models \alpha_{i+1}$  sse  $\alpha_{i+1} \in \Delta$ , para todo  $i \geq 0$ ; disso se segue que  $E \models \Gamma_\omega$  e que, para cada  $i > 0$ , se  $\alpha_i \notin \Delta$ , então  $\Gamma_{i-1} \models \sim \alpha_i$ , e, portanto,  $\Gamma_i = \Gamma_{i-1} \cup \{\sim \alpha_i\}$ ; daí temos que, para cada  $i > 0$ , se  $\alpha_i \notin \Delta$ , então  $\sim \alpha_i \in \Gamma_\omega$ ; assim, se  $\alpha_i \notin \Delta$ , então  $\Gamma_\omega \models \sim \alpha_i$ ; como  $E \models \Gamma_\omega$ , temos que  $\Gamma_\omega$  é consistente, e logo se  $\alpha_i \notin \Delta$ , então  $\Gamma_\omega \models \sim \alpha_i$ ; obviamente, se  $\alpha_i \in \Delta$ , então  $\Gamma_\omega \models \alpha_i$ ; segue-se que  $\Gamma_\omega \models \alpha_i$  sse  $\alpha_i \in \Delta$ ; daí temos que se  $\Gamma_\omega \models \sim \alpha_i$ , então  $\sim \alpha_i \in \Gamma_\omega$ ; assim,  $\Gamma_\omega = \Gamma$  por 1-4 e, como  $E \models \Gamma_\omega$ ,  $E \models \Gamma$ , donde se segue que  $\Gamma$  é consistente.

ii)  $\Gamma$  é completo: vamos provar a completude de  $\Gamma$  por indução completa sobre o número  $k$  de operadores e quantificadores em  $\alpha$ :

Base:  $n = 0$

$\alpha$  é atômica; se  $\alpha \in \Delta$ ,  $\Gamma \models \alpha$ ; se  $\alpha \notin \Delta$  e  $\alpha$  não tem ocorrências da constante predicativa '=', segue-se por i) que  $\Gamma \models \alpha$ , e daí que  $\sim \alpha \in \Gamma$ , e logo que  $\Gamma \models \sim \alpha$ ; já se  $\alpha$  contém ocorrências da constante '=', então  $\alpha$  tem a forma  $\bar{c} = \bar{c}$ , para uma dada constante individual  $\bar{c}$  de  $S$ , ou a forma  $\bar{c}_i = \bar{c}_j$ , para um par  $(\bar{c}_i, \bar{c}_j)$  qualquer de constantes individuais distintas de  $S$ ; no primeiro caso, obviamente  $\Gamma \models \alpha$ , já que  $\alpha$  é verdadeira em todas as estruturas, e portanto  $\alpha$  é em todos os modelos de  $\Gamma$ ; no segundo caso,  $\sim \alpha \in \Gamma$  por definição, e logo  $\Gamma \models \sim \alpha$ .

Hipótese indutiva (HI): ii) vale para todo  $j < n$

Passo de indução<sup>6</sup>:

Caso 1:  $\alpha = \sim \beta$ : por HI,  $\Gamma \models \beta$  ou  $\Gamma \models \sim \beta$ ; no último caso,  $\Gamma \models \alpha$ ; no primeiro,  $E \models \beta$  dada toda estrutura  $E$  tal que  $E \models \Gamma$ , e daí, como  $E \models \beta$  sse  $E \models \sim \sim \beta$ , temos que  $E \models \sim \sim \beta$  dada toda estrutura  $E$  tal que  $E \models \Gamma$ , isto é, temos que,  $\Gamma \models \sim \sim \beta$ , o que significa que  $\Gamma \models \sim \alpha$ .

Caso 2:  $\alpha = \beta \rightarrow \gamma$ : por HI,  $\Gamma \models \beta$  ou  $\Gamma \models \sim \beta$ , e  $\Gamma \models \gamma$  ou  $\Gamma \models \sim \gamma$ ; se  $\Gamma \models \beta$  e  $\Gamma \models \sim \gamma$ , então temos que  $\Gamma \models \sim (\beta \rightarrow \gamma)$ , isto é, temos que  $\Gamma \models \sim \alpha$ ; nos outros três casos, temos que  $\Gamma \models \beta \rightarrow \gamma$ , o que quer dizer que  $\Gamma \models \alpha$ .

Caso 3:  $\alpha = \forall \xi \beta[\xi]$ , onde  $\xi$  é uma variável individual qualquer em  $S$ , e  $\beta[\xi]$  é uma fórmula de  $L(S)$  contendo ao menos uma ocorrência dessa variável<sup>7</sup>: por HI, para qualquer constante  $\bar{c} \in S$ , ou  $\Gamma \models \beta[\xi/\bar{c}]$  ou  $\Gamma \models \sim \beta[\xi/\bar{c}]$ , onde  $\beta[\xi/\bar{c}]$  é a fórmula que resulta de  $\beta[\xi]$  por substituição de cada ocorrência da variável  $\xi$  pela constante  $\bar{c}$ ; como  $\forall x (x = \bar{c}_1 \vee x = \bar{c}_2 \vee \dots \vee x = \bar{c}_n) \in \Gamma$ , e logo  $\Gamma \models \forall x (x = \bar{c}_1 \vee x = \bar{c}_2 \vee \dots \vee x = \bar{c}_n)$ , temos que  $\Gamma \models \forall \xi (\xi = \bar{c}_1 \vee \xi = \bar{c}_2 \vee \dots \vee \xi = \bar{c}_n)$ ; pois bem, se  $\Gamma \models \beta[\xi/\bar{c}]$  para toda constante individual  $\bar{c}$  de  $S$ , segue-se daí e de que  $\Gamma \models \forall \xi (\xi = \bar{c}_1 \vee \xi = \bar{c}_2 \vee \dots \vee \xi = \bar{c}_n)$  que  $\Gamma \models \forall \xi \beta[\xi]$ , isto é, segue-se que  $\Gamma \models \alpha$ ; já se  $\Gamma \models \sim \beta[\xi/\bar{c}_i]$  para alguma constante individual  $\bar{c}_i$  de  $S$ , temos, por HI, que  $\Gamma \models \sim \beta[\xi/\bar{c}_i]$ ; segue-se daí que  $E \models \sim \beta[\xi/\bar{c}_i]$  dada toda estrutura  $E = (D_E, I_E)$  tal que  $E \models \Gamma$ , e logo  $E \models \beta[\xi/\bar{c}_i]$  dadas tais estruturas; mas, nesse caso, dadas tais estruturas,  $\nu^{\sigma_E}(\beta[\xi]) = 0$  para ao menos uma atribuição  $\tau_E$   $\xi$ -variante de uma atribuição  $\sigma_E$  qualquer, a saber, uma atribuição  $\tau_E$  tal que  $\tau_E(\xi) = I_E(\bar{c}_i)$ ; segue-se que  $\nu^{\sigma_E}(\forall \xi \beta[\xi]) = 0$  dada qualquer atribuição  $\sigma_E$ , e daí que  $E \models \sim \forall \xi \beta[\xi]$  dada toda estrutura  $E$  tal que  $E \models \Gamma$ ; segue-se disso que  $E \models \sim \forall \xi \beta[\xi]$  dadas essas estruturas, e logo que  $\Gamma \models \sim \forall \xi \beta[\xi]$ , isto é, que  $\Gamma \models \sim \alpha$ .

O interesse que uma base de conhecimento vívida possui reside no fato de que é possível desenvolver um mecanismo inferencial que, quando aplicado a uma tal base de conhecimento, não só é capaz de calcular o conjunto de suas consequências lógicas, como ainda é capaz de fazê-lo em tempo polinomial<sup>8</sup>.

TEO. 2: Seja  $\Gamma$  uma base de conhecimento vívida de núcleo  $\Delta$ , e  $\alpha$  uma fórmula qualquer de  $L(S)$ ; há um algoritmo computável em tempo polinomial que recebe o par  $(\Gamma, \alpha)$  como entrada, e devolve 1 caso  $\Gamma \models \alpha$ , e 0 caso contrário.

Prova: indução completa sobre o número  $n$  de operadores e quantificadores em  $\alpha$ .

Base:  $n = 0$

$\alpha$  é atômica; se  $\alpha$  possui a forma  $\bar{c} = \bar{c}$ , para uma dada constante individual  $\bar{c}$  de  $S$ , o mecanismo inferencial (MI) devolve 1; já se  $\alpha$  tem a forma  $\bar{c}_i = \bar{c}_j$ , para um

par  $(\bar{c}_i, \bar{c}_j)$  qualquer de constantes individuais distintas de  $S$ , MI devolve 0; nos demais casos, MI simplesmente percorre  $\Delta$ , e devolve 1 se  $\alpha \in \Delta$ , 0 se  $\alpha \notin \Delta$ ; essa verificação demanda no máximo  $m + 1$  passos, onde  $m$  é o número de fórmulas em  $\Delta$ ; logo, se  $f$  é a função que determina o tempo  $t$  que MI demora para verificar se  $\alpha$  se segue de  $\Gamma$ , em função de  $m$ , então se  $p(m) = m + 1$ , temos que, para  $k = 1$  e  $k_0 = 0$ ,  $f(m) \leq k p(m)$ , para qualquer  $m \geq k_0$ <sup>9</sup>.

Hipótese indutiva (HI): TEO. 2 vale para todo  $j < n$

Passo de indução<sup>10</sup>:

Caso 1:  $\alpha = \sim\beta$ : tendo computado em tempo polinomial se  $\Gamma \models \beta$ , o que deve ocorrer por HI, MI devolve 0 se  $\Gamma \models \beta$ , 1 caso contrário.

Caso 2:  $\alpha = \beta \ \& \ \gamma$ : tendo computado em tempo polinomial se  $\Gamma \models \beta$  e se  $\Gamma \models \gamma$ , o que deve ocorrer por HI, MI devolve 1 se  $\Gamma \models \beta$  e  $\Gamma \models \gamma$ , 0 caso contrário.

Caso 3:  $\alpha = \forall \xi \beta[\xi]$ : por HI, MI computa em tempo polinomial se  $\Gamma \models \beta[\xi/\bar{c}]$ , para uma constante individual qualquer  $\bar{c} \in S$ ; seja  $r$  o número de constantes individuais em  $S$ ; seja  $p(m)$  o polinômio tal que  $f(m) \leq k p(m)$ , para qualquer  $m \geq k_0$ , para  $k$  e  $k_0$  dados, sendo  $f$  a função que determina o tempo  $t$  que MI demora para verificar se  $\Gamma \models \beta[\xi/\bar{c}]$ ; então, se o polinômio  $q(m) = r p(m)$ , teremos que  $g(m) \leq k q(m)$ , para qualquer  $m \geq k_0$ , onde  $g$  é a função que determina o tempo  $t$  que MI demora para verificar se  $\Gamma \models \beta[\xi/\bar{c}]$  para toda constante individual  $\bar{c} \in S$ , o que significa que essa verificação pode ser feita por MI em tempo polinomial; daí, se  $\Gamma \models \beta[\xi/\bar{c}]$  para toda constante individual  $\bar{c} \in S$ , MI devolve 1, do contrário, MI devolve 0.

Esse fato sobre as bases de conhecimento vívidas é certamente notável, considerando-se: a) a indecidibilidade do cálculo de predicados de 1ª ordem (CP1), que tem como consequência que sistemas baseados em CP1<sup>11</sup>, como é o caso daqueles baseados em resolução, não podem determinar em todos os casos, após um número finito de passos, se uma fórmula arbitrária é uma consequência lógica de uma base de conhecimento (um conjunto de fórmulas) arbitrária, e, b) considerando-se que mesmo na lógica proposicional o tempo que um mecanismo inferencial qualquer vai tomar para decidir se uma fórmula arbitrária é consequência lógica de uma base de conhecimento arbitrária cresce exponencialmente conforme cresce o tamanho da base de conhecimento em questão<sup>12</sup>.

No entanto, está claro que uma base de conhecimento vívida é limitada do ponto-de-vista da sua capacidade expressiva, já que há fatos sobre o mundo que não podem ser designados por proposições representáveis pelas fórmulas atômicas que são admitidas no núcleo de tais bases de conhecimento, como é o caso de leis universais como ‘todos os planetas são esferóides’<sup>13</sup>. Na realidade, é bem conhecida a relação existente entre a expressividade de uma base de conhecimento, a capacidade inferencial de um dado mecanismo inferencial, e o grau de complexidade computacional – ou mesmo a decidibilidade – que vai adquirir a tarefa de se computar as consequências lógicas dessa base de conhecimento utilizando tal mecanismo inferencial. Para manter um sistema decidível e computacionalmente tratável, é necessário reduzir a expressividade

das bases de conhecimento às quais o mecanismo inferencial de tal sistema é aplicado, ou reduzir as capacidades inferenciais do mecanismo em questão. Falando a esse respeito no livro que mencionamos mais acima, Brachman e Levesque defendem o interessante ponto de vista de que o trabalho principal a ser desenvolvido na área de representação do conhecimento consiste em explorar os pontos de equilíbrio entre expressividade e capacidade inferencial, que proporcionem os melhores resultados em termos da tratabilidade computacional de um dado sistema<sup>14</sup>.

Na sequência, vamos apresentar os resultados da exploração de um desses pontos, com bases de conhecimento mais expressivas que as bases vívidas, e um procedimento capaz de transformá-las em bases vívidas, sem perda semântica, isto é, sem que se percam as informações veiculadas por aquelas fórmulas na base de conhecimento que não são permitidas em bases vívidas.

Para sermos mais específicos, nossas bases de conhecimento são bases vívidas acrescidas de uma quantidade finita de fórmulas universais da forma  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ , onde  $\xi_i$ , para  $1 \leq i \leq k$  e  $k$  sendo um número natural  $\geq 1$ , é uma variável individual qualquer;  $\beta$  e  $\alpha_j$ , para  $0 \leq j \leq m$  e  $m$  sendo um número natural  $\geq 0$ , são fórmulas atômicas de  $L(S)$  sem igualdade, cujas variáveis livres estão na sequência  $\xi_1, \dots, \xi_k$ . Nos casos em que  $m = 1$  e  $m = 0$ , respectivamente, temos fórmulas das formas  $\forall \xi_1 \dots \forall \xi_k (\alpha \rightarrow \beta)$  e  $\forall \xi_1 \dots \forall \xi_k \beta$ . Exemplos dessas fórmulas podem ser:  $\forall x \forall y \forall z ((\text{menor}(x, y) \& \text{menor}(y, z)) \rightarrow \text{menor}(x, z))$  e  $\forall x (\text{planeta}(x) \rightarrow \text{esferóide}(x))$ . Em notação clausal essas fórmulas têm a forma  $[\sim \alpha_1, \dots, \sim \alpha_m, \beta]$ <sup>15</sup>, e são conhecidas como cláusulas de Horn. Desse modo, uma qualquer de nossas bases de conhecimento pode ser descrita como um conjunto  $\Gamma$  de fórmulas, obtido a partir de um conjunto finito  $\Delta$  de fórmulas representáveis em notação clausal por cláusulas de Horn positivas sem igualdade e livres de símbolos funcionais – que vamos chamar de núcleo de  $\Gamma$  – mediante o acréscimo a  $\Delta$ , cujo *dicionário* contém todas as constantes individuais ou predicativas que ocorrem em  $\Delta$ : i) da fórmula  $\forall x (x = \bar{c}_1 \vee x = \bar{c}_2 \vee \dots \vee x = \bar{c}_n)$ , sendo  $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n$  uma sequência de todas as constantes individuais no dicionário de  $\Delta$ , ii) da fórmula  $\sim \bar{c}_i = \bar{c}_j$ , para todo par  $(\bar{c}_i, \bar{c}_j)$  de constantes individuais distintas no dicionário de  $\Delta$ , e iii) de  $\sim \alpha$ , para toda fórmula atômica fechada  $\alpha \in L(S)$ , tal que  $\Gamma \models \alpha$ . Vamos chamar às nossas bases  $\Gamma$  de *bases de Horn de 1ª ordem restritas*.

Agora considere-se uma base de conhecimento  $\Gamma$  com as características mencionadas. Vamos chamar de  $T$  à seguinte transformação aplicada a  $\Gamma$ : para toda fórmula  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , acrescenta-se  $\beta[\xi/\bar{c}]$  a  $\Gamma$ , desde que  $\alpha_i[\xi/\bar{c}] \in \Gamma$  para  $1 \leq i \leq m$ , e  $\beta[\xi/\bar{c}] \notin \Gamma$ , sendo  $\alpha_i[\xi/\bar{c}]$  a fórmula que resulta de  $\alpha_i$  pela substituição de cada ocorrência da variável  $\xi_j$ ,  $1 \leq j \leq k$ , por uma dada constante individual  $\bar{c}$ , as mesmas variáveis sendo substituídas pelas mesmas constantes, e sendo  $\beta[\xi/\bar{c}]$  cada fórmula que resulta de  $\beta$  pela mesma substituição de variáveis por constantes que permite a obtenção de  $\alpha_i[\xi/\bar{c}]$  a partir de  $\alpha_i$ ,  $1 \leq i \leq m$ , com as variáveis de  $\beta$  que não ocorrem em nenhuma das fórmulas  $\alpha_i$  sendo substituídas por cada uma das constantes  $\bar{c}$  que ocorrem no dicionário de  $\Delta \subset \Gamma$ ; repete-se esse procedimento para todas as substituições  $\xi/\bar{c}$  possíveis<sup>16</sup>; vamos

dizer que dessa ação resulta a base de conhecimento  $\Gamma_1$ ; repete-se o procedimento, resultando daí a base de conhecimento  $\Gamma_2$ ; repete-se o procedimento, até que se obtenha a base  $\Gamma_n = \Gamma_{n-1}$ . A base de conhecimento  $T(\Gamma)$  resulta de  $\Gamma_n$  por supressão de todas as fórmulas do tipo  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ . Claramente,  $T(\Gamma)$  é uma base vívida cujas consequências lógicas são as mesmas que as de  $\Gamma$ , isto é, para qualquer fórmula  $\alpha \in L(S)$ ,  $\Gamma \models \alpha$  sse  $T(\Gamma) \models \alpha$ , o que confirma nossa afirmação anterior de que a transformação  $T$  da base de conhecimento  $\Gamma$  na base vívida  $T(\Gamma)$  preserva o conteúdo de  $\Gamma$ .

TEO. 3: Seja  $\Gamma$  uma de nossas bases de Horn de 1ª ordem restritas, então uma estrutura  $E$  é modelo de  $\Gamma$  sse  $E$  é modelo de  $T(\Gamma)$ .

Prova: i) suponha-se que  $E \models \Gamma$ ; agora, suponha-se que  $E \models \Gamma_j$ ; então temos que  $E \models \gamma$ , dada toda fórmula  $\gamma \in \Gamma_j$ ; isso inclui o fato de que  $E \models \forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ , para qualquer fórmula dessa forma em  $\Gamma_j$ ; isso quer dizer que a) para todas as fórmulas em questão,  $\nu^{\sigma_E}(\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)) = 1$  dada toda atribuição  $\sigma_E$ ; dada alguma substituição  $[\xi/\bar{c}]$ , se  $\alpha_i[\xi/\bar{c}] \in \Gamma_j$ , para  $1 \leq i \leq m$ , então b)  $\nu^{\sigma_E}(\alpha_i[\xi/\bar{c}]) = 1$  dada toda atribuição  $\sigma_E$ ; de a) segue-se obviamente que  $\nu^{\sigma_E}((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)[\xi/\bar{c}]) = 1$  dada toda atribuição  $\sigma_E$ , e daí e de b) segue-se que  $\nu^{\sigma_E}(\beta[\xi/\bar{c}]) = 1$  dada toda atribuição  $\sigma_E$ , e logo que  $E \models \beta[\xi/\bar{c}]$ ; como já tínhamos por hipótese que  $E \models \Gamma_j$ , e as fórmulas  $\beta[\xi/\bar{c}]$  são as únicas eventualmente acrescentadas a  $\Gamma_j$  na obtenção de  $\Gamma_{j+1}$ , segue-se que  $E \models \Gamma_{j+1}$ ; daí, como  $E \models \Gamma = \Gamma_0$ , temos por indução matemática sobre  $j$  que  $E \models \Gamma_j$  para todo número natural  $j$ , e logo que  $E \models \Gamma_n$ ; agora, como  $T(\Gamma) \subset \Gamma_n$ , segue-se que  $E \models T(\Gamma)$ ; ii) agora, suponha-se que  $E \models T(\Gamma)$ ; como  $\Gamma_{n+1} = \Gamma_n = \Gamma_{n-1}$ , temos que, para toda fórmula  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta) \in \Gamma$ , se  $\alpha_i[\xi/\bar{c}] \in \Gamma_n$ ,  $1 \leq i \leq m$ , para uma substituição  $[\xi/\bar{c}]$  qualquer, então  $\beta[\xi/\bar{c}] \in \Gamma_n$ ; daí, como  $T(\Gamma)$  resulta de  $\Gamma_n$  pela eliminação das fórmulas da forma  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ , temos que se  $\alpha_i[\xi/\bar{c}] \in T(\Gamma)$ ,  $1 \leq i \leq m$ , então  $\beta[\xi/\bar{c}] \in T(\Gamma)$ ; e como por hipótese  $E \models T(\Gamma)$ , temos que se  $\alpha_i[\xi/\bar{c}] \in T(\Gamma)$ ,  $1 \leq i \leq m$ , então  $E \models \alpha_i[\xi/\bar{c}]$ ,  $1 \leq i \leq m$ , donde se segue que  $E \models (\alpha_1 \& \dots \& \alpha_m)[\xi/\bar{c}]$ ; mas, nesse caso, vimos acima que  $\beta[\xi/\bar{c}] \in T(\Gamma)$ , donde também se segue que  $E \models \beta[\xi/\bar{c}]$ ; portanto, se  $E \models (\alpha_1 \& \dots \& \alpha_m)[\xi/\bar{c}]$ , então  $E \models \beta[\xi/\bar{c}]$ , e logo temos que  $E \models ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)[\xi/\bar{c}]$ ; ora, como as substituições  $\xi/\bar{c}$  envolvem todas as  $k$ -uplas de constantes em  $S$ , e como  $E \models \forall x (x = \bar{c}_1 \vee x = \bar{c}_2 \vee \dots \vee x = \bar{c}_n)$ , já que essa fórmula ocorre em  $T(\Gamma)$ , é um exercício simples verificar que  $E \models \forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ ; como essas são as únicas fórmulas que foram eliminadas de  $\Gamma_n$  para a obtenção de  $T(\Gamma)$ , e como  $E \models T(\Gamma)$ , segue-se que  $E \models \Gamma_n$ ; e como  $\Gamma \subseteq \Gamma_n$ , temos que  $E \models \Gamma$ .

COR. 4: Para qualquer fórmula  $\alpha \in L(S)$ ,  $\Gamma \models \alpha$  sse  $T(\Gamma) \models \alpha$ .

Prova: Suponha-se que  $\Gamma \models \alpha$ , então  $E \models \alpha$  dada toda estrutura  $E$  tal que  $E \models \Gamma$ ; segue-se pelo teorema 3 que  $E \models \alpha$  a dada toda estrutura  $E$  tal que  $E \models T(\Gamma)$ , e logo que  $T(\Gamma) \models \alpha$ ; supondo-se que  $T(\Gamma) \models \alpha$ , segue-se por argumento similar que  $\Gamma \models \alpha$ .

As consequências lógicas de  $T(\Gamma)$ , uma vez que essa é uma base de conhecimento vívida, são computáveis em tempo polinomial. Além disso, como vamos demonstrar na sequência, a transformação  $T$  é computável, isto é, há um

número natural  $n$  tal que  $\Gamma_n = \Gamma_{n-1}$ , para qualquer base de conhecimento  $\Gamma$  do tipo especificado acima.

TEO. 5: Seja  $\Gamma$  uma qualquer de nossas bases de Horn de 1ª ordem restritas, e  $\Gamma_1, \Gamma_2, \dots$  a série de bases de conhecimento mencionada mais acima, na descrição da transformação  $T$ ; temos que há um número natural  $n$  tal que  $\Gamma_n = \Gamma_{n-1}$ .

Prova: Suponha-se que a transformação  $T$  aplicada a  $\Gamma$  seja implementada com base no seguinte algoritmo  $\mathbb{A}$ :  $\mathbb{A}$  recebe  $\Gamma$ , e extrai de  $\Gamma$  uma sequência  $(\bar{c}_1, \dots, \bar{c}_c)$  de todas as  $c$  constantes individuais que ocorrem em  $\Gamma$ ; depois disso, para cada fórmula do tipo  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta) \in \Gamma$ ,  $\mathbb{A}$  toma todas as  $k$ -uplas ordenadas  $(\bar{c}^{i_1}, \dots, \bar{c}^{i_k})$ , para  $1 \leq i \leq c$ , e obtém as fórmulas  $\alpha_j[\xi_j/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}]$ ,  $1 \leq j \leq m$ , e  $\beta[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}]$ , a partir de  $\alpha_j$  e  $\beta$ , respectivamente, mediante a substituição de cada ocorrência da variável  $\xi_p$  pela constante individual  $\bar{c}^{i_p}$ ,  $1 \leq p \leq k$ ; em seguida  $\mathbb{A}$  percorre  $\Gamma$ , e acrescenta  $\beta[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}]$  a  $\Gamma$ , caso cada  $\alpha_j[\xi_j/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}] \in \Gamma$ ,  $1 \leq j \leq m$ , e  $\beta[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}] \notin \Gamma$ ; como todas as substituições possíveis de variáveis por constantes são verificadas por  $\mathbb{A}$ , após esse procedimento teremos que, se  $\alpha_j[\xi/\bar{c}] \in \Gamma$ ,  $1 \leq j \leq m$ , então  $\beta[\xi/\bar{c}] \in \Gamma$ , sendo  $\beta[\xi/\bar{c}]$  a fórmula que resulta de  $\beta$  pela mesma substituição de variáveis por constantes que permitiu a obtenção de  $\alpha_j[\xi/\bar{c}]$  a partir de  $\alpha_j$ , com as variáveis que ocorrem em  $\beta$  mas não ocorrem em  $\alpha_j$  substituídas por toda constante individual em  $(\bar{c}_1, \dots, \bar{c}_c)$ , ou seja, teremos que  $\mathbb{A}$  obteve  $\Gamma_1$ ; sabemos que  $\mathbb{A}$  pode obter  $\Gamma_1$  em um número finito de passos devido aos fatos de que: a) há uma quantidade finita de fórmulas da forma  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$  e, b) como a sequência  $(\bar{c}_1, \dots, \bar{c}_c)$  é finita, também deve sê-lo o conjunto de todas as  $k$ -uplas ordenadas  $(\bar{c}^{i_1}, \dots, \bar{c}^{i_k})$ , para qualquer número natural  $k$ ; ora, ao obter  $\Gamma_1$  de  $\Gamma$ ,  $\mathbb{A}$  acrescenta a  $\Gamma$  uma quantidade finita  $r \geq 0$  de fórmulas  $\beta[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}]$ , mas não acrescenta nenhuma nova fórmula universal a  $\Gamma$ , e nem acrescenta qualquer constante individual nova a  $\Gamma$ , de modo que, para obter  $\Gamma_2$  de  $\Gamma_1$ ,  $\mathbb{A}$  pode repetir o procedimento que permitiu a obtenção de  $\Gamma_1$  a partir de  $\Gamma$ , verificando se  $\alpha_j[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}] \in \Gamma_1$ ,  $1 \leq j \leq m$ , para cada fórmula da forma  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , dadas as mesmas substituições de variáveis por constantes utilizadas na obtenção de  $\Gamma_1$  a partir de  $\Gamma$ ; mas, nesse caso, o número de fórmulas  $\beta[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}]$  que podem ser acrescentadas a  $\Gamma$  é finito; assim, ao obter  $\Gamma_q$  de  $\Gamma_{q-1}$ , para um dado número natural  $q$ ,  $\mathbb{A}$  acrescenta ou não fórmulas  $\beta[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}]$  a  $\Gamma_{q-1}$ ; se  $\mathbb{A}$  não acrescenta nenhuma dessas fórmulas a  $\Gamma_{q-1}$ , então  $\Gamma_q = \Gamma_{q-1}$ , e  $\mathbb{A}$  pára; do contrário  $\mathbb{A}$  deve prosseguir, e repetir o procedimento inicial para  $\Gamma_q$ ; entretanto, no limite,  $\mathbb{A}$  terá obtido uma base  $\Gamma_q$ , para um dado número natural  $q$ , tal que a fórmula  $\beta[\xi_1/\bar{c}^{i_1}, \dots, \xi_k/\bar{c}^{i_k}] \in \Gamma_q$  para todas as  $k$ -uplas  $(\bar{c}^{i_1}, \dots, \bar{c}^{i_k}) \in \{\bar{c}_1, \dots, \bar{c}_c\}^k$ , dada cada uma das fórmulas da forma  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ ; nesse caso,  $\Gamma_{q+1}$  será necessariamente igual a  $\Gamma_q$ , e  $\mathbb{A}$  pára.

Muito embora seja computável, como acabamos de demonstrar, a obtenção de  $T(\Gamma)$  a partir de  $\Gamma$  por  $\mathbb{A}$  não apresenta complexidade polinomial. Na realidade, já a obtenção de  $\Gamma_1$  a partir de  $\Gamma_{i-1}$  tem complexidade exponencial no número máximo  $k$  de variáveis de uma fórmula do tipo  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ . De fato, é um exercício simples mostrar que o número de passos que  $\mathbb{A}$  deve

executar para obter  $\Gamma_i$  de  $\Gamma_{i-1}$ , para algum  $i$  tal que  $1 \leq i \leq n$ , não será maior que  $ac^k + b$ , onde  $a = u [(m + 1)t + m + 2]$ , e  $b = t + u$ , sendo  $m$  o número máximo de condições  $\alpha_i$  em uma fórmula do tipo  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ ,  $u$  o número de tais fórmulas em  $\Gamma$ ,  $t$  o número de fórmulas atômicas em  $\Gamma$ ,  $c$  o número de constantes individuais no dicionário da base de conhecimento, e  $k$  o número máximo de variáveis em uma fórmula do tipo  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ . Supondo  $m$ ,  $u$ ,  $t$  e  $c$  constantes,  $a$  e  $b$  também o serão, donde se segue que  $f(k) = b$  é  $O(1)$ . Como  $ac^k \leq (ac)^k$  para  $k \geq 0$ , temos, fazendo  $ac = d$ , que  $ac^k$  é  $O(d^k)$ . Como  $f(k) = b$  é  $O(1)$ , e logo é também  $O(d^k)$ , temos, pela regra da soma, que  $g(k) = ac^k + b$  é  $O(d^k)$ .

Do fato de que  $\mathbb{A}$  toma ao menos tempo exponencial para obter  $T(\Gamma)$  a partir de  $\Gamma$  não se segue, é claro, que seja impossível obter um outro algoritmo que faça o mesmo trabalho em tempo polinomial. Como desconfiamos que um tal algoritmo não existe, a continuação óbvia do trabalho apresentado aqui consiste em procurar obter uma redução polinomial de um problema NP-completo para o problema de se computar  $T(\Gamma)$  a partir de  $G^{17}$ . Uma vez que se tenha conseguido tal redução, teremos uma garantia de que a desconfiança mencionada acima está correta, exceto se pudermos mostrar que o problema de se computar  $T(\Gamma)$  a partir de  $\Gamma$  está em NP, e ainda que  $P = NP^{18}$ .

Entretanto, mesmo que a obtenção de  $T(\Gamma)$  a partir de  $\Gamma$  seja um problema NP-completo ou mesmo NP-hard, o fato de essa redução ser computável já nos põe em vantagem com relação à questão de se determinar se uma dada fórmula atômica se segue de uma base de Horn de 1ª ordem irrestrita, que é sabidamente indecidível. Além disso, parece-nos que, para fins de aplicação em sistemas baseados em conhecimento, o que mostramos aqui evidencia que bases de Horn restritas podem ser interessantes, sobretudo no caso de se estar representando conhecimento com pouca matemática envolvida. De fato, nessas situações as fórmulas que são utilizadas para representar regras costumam apresentar uma sequência não muito longa de quantificadores, minimizando o problema com a intratabilidade de nossa transformação, uma vez que a mesma é exponencial precisamente no número máximo  $k$  de variáveis quantificadas nas fórmulas universais que representam regras em nossas bases de conhecimento, conforme mostramos mais acima.

Por fim, é possível substituir  $\mathbb{A}$  por outros algoritmos que, embora ainda tenham complexidade exponencial, parecem se comportar melhor que  $\mathbb{A}$  em boa parte dos casos. Por exemplo,  $\mathbb{A}$  utiliza um procedimento conhecido como *backward chaining*, que tem complexidade exponencial mesmo quando aplicado a bases de conhecimento que só contém fórmulas de uma linguagem para a lógica proposicional clássica<sup>19</sup>. O seguinte algoritmo, que vamos chamar de  $\mathbb{B}$ , usa o procedimento conhecido como *forward chaining*, e, como veremos mais adiante, há ao menos um algoritmo baseado em  $\mathbb{B}$  cujo desempenho é melhor que o de  $\mathbb{A}$  em alguns casos:  $\mathbb{B}$  recebe uma de nossas bases de conhecimento  $\Gamma$  e, para cada fórmula do tipo  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ ,  $\mathbb{B}$  percorre o conjunto  $T \subset \Gamma$  das fórmulas atômicas de  $\Gamma$ , em busca de instâncias de  $\alpha_i$ ; para cada instância

$\alpha_j[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}]$  encontrada,  $\mathbb{B}$  verifica se  $\beta[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}] \in \Gamma$ ; se esse não for o caso;  $\mathbb{B}$  verifica se  $\alpha_2[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}] \in \Gamma$ ; se isso for o caso,  $\mathbb{B}$  verifica se  $\alpha_3[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}] \in \Gamma$ , e assim por diante, até chegar, eventualmente, a  $\alpha_m[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}]$ ; se  $\alpha_i[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}] \in \Gamma$ , para  $1 \leq i \leq m$ , então  $\mathbb{B}$  acrescenta  $\beta[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}]$  a  $\Gamma$ ; ao final desse procedimento,  $\mathbb{B}$  terá obtido  $\Gamma_i$ ;  $\mathbb{B}$  repete o mesmo procedimento, até que não haja mais instâncias de  $\beta$  a acrescentar a  $\Gamma_i$ , para qualquer das fórmulas universais em  $\Gamma$ , de modo que  $\Gamma_{i+1} = \Gamma_i = \Gamma_n$ ; isso feito  $\mathbb{B}$  elimina de  $\Gamma_n$  cada fórmula universal, obtendo  $T(\Gamma)$ .

Se para ao menos um  $j$ ,  $1 \leq j \leq m$ ,  $\alpha_j$  contiver todas as  $k$  variáveis quantificadas em uma fórmula  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ , e isso for o caso para cada uma das fórmulas desse tipo em  $\Gamma$ , então  $\mathbb{B}'$  é capaz de computar  $\Gamma_i$  a partir de  $\Gamma_{i-1}$  em tempo polinomial, sendo  $\mathbb{B}'$  um algoritmo idêntico a  $\mathbb{B}$ , com a exceção de que  $\mathbb{B}'$ , para cada universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , começa verificando se  $\alpha_1$  contém todas as  $k$  variáveis quantificadas em  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ , e, em caso negativo, passa para a 1ª posição a primeira fórmula  $\alpha_j$  que tenha essa propriedade, se houver uma tal fórmula em  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$ . Para cada universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ ,  $\mathbb{B}'$  toma no máximo  $m$  passos para fazer a verificação mencionada; mais 1 passo para fazer, se for o caso, a troca de posição mencionada; mais  $t$  passos para percorrer as  $t$  fórmulas atômicas em  $\Gamma_i$  à busca de uma instância  $\alpha_i[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}]$  de  $\alpha_i$ ; mais  $t$  passos para verificar se  $\beta[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}] \in \Gamma_i$ ; mais  $t$  passos para verificar se  $\alpha_j[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}] \in \Gamma_i$ , eventualmente para cada  $j$  tal que  $2 \leq j \leq m$ , caso todas essas instâncias sejam encontradas; mais 1 passo, eventualmente, para acrescentar  $\beta[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}]$  a  $\Gamma_i$ , totalizando  $(m+1)t + m + 2$  passos. Isso deve ser feito para cada instância de  $\alpha_i$  encontrada em  $\Gamma_i$ , das quais obviamente não pode haver mais que  $t$ , resultando em um limite de  $(m+1)t^2 + (m+2)t$  passos para cada universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma^{20}$ ; após encontrar todas as instâncias de  $\alpha_i$  em  $\Gamma_i$ ,  $\mathbb{B}'$  ainda deverá percorrer o conjunto das fórmulas atômicas em  $\Gamma_i$  para verificar que não há mais nenhuma dessas instâncias, totalizando  $(m+1)t^2 + (m+2)t + t = (m+1)t^2 + (m+3)t$  passos para cada universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ . Considerando todas as  $u$  universais  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , temos que  $\mathbb{B}'$  computa  $\Gamma_{i+1}$  a partir de  $\Gamma_i$  em não mais que  $u((m+1)t^2 + (m+3)t)$  passos, sendo  $m$  o maior número de fórmulas atômicas  $\alpha_j$  nas universais  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta) \in \Gamma$ . Desenvolvendo essa expressão, temos  $(m+1)ut^2 + (m+3)ut$ , e daí  $(mu+u)t^2 + (mu+3u)t$ . Supondo  $u$  e  $m$  constantes, temos, para  $a = mu+u$  e  $b = mu+3u$ , que  $\mathbb{B}'$  toma não mais que  $at^2 + bt$  passos para computar  $\Gamma_{i+1}$  a partir de  $\Gamma_i$ . Como  $f(t) = bt$  é  $O(t)$ , e portanto também  $O(t^2)$ , e  $g(t) = at^2$  é  $O(t^2)$ , temos pela regra da soma que  $h(t) = at^2 + bt$  é  $O(t^2)$ , ou seja, a computação por  $\mathbb{B}'$  de  $\Gamma_{i+1}$  a partir de  $\Gamma_i$  tem complexidade quadrática, e portanto polinomial, dependente do número  $t$  de fórmulas atômicas em  $\Gamma_i$ , quando aplicado a uma base de conhecimento  $\Gamma_i$  em que cada universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  satisfaz a condição mencionada mais acima.

Vamos agora considerar o problema completo da computação de  $T(\Gamma)$  a partir de  $\Gamma$ . Na pior hipótese, em que  $\Gamma$  não contém nenhuma das instâncias de  $\beta$ ,

para qualquer das  $u$  universais  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ ,  $T(\Gamma)$  contém todas essas instâncias, e cada passagem de  $\Gamma_{i-1}$  a  $\Gamma_i$ ,  $1 \leq i \leq n$ , só acrescenta uma única dessas instâncias a  $\Gamma_{i-1}$ , teremos que  $n \leq uc^r$ , onde  $c$  é o número de constantes individuais em  $\Gamma$ , e  $r$  a maior aridade de  $\beta$  nas universais  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ . Assim,  $\mathbb{B}'$  computa  $T(\Gamma)$  a partir de  $\Gamma$  em não mais que  $uc^r$  ( $at^2 + bt$ ), onde  $a$  e  $b$  são definidos como acima, e  $t$  é o número de fórmulas atômicas em  $\Gamma_n$ <sup>21</sup>. Supondo  $t$  fixo, e fazendo  $d = uat^2 + ubt$ , temos que a complexidade do problema de se obter  $T(\Gamma)$  a partir de  $\Gamma$  utilizando  $\mathbb{B}'$  é dada pela função  $dc^r$ . A computação de  $T(\Gamma)$  a partir de  $\Gamma$  por  $\mathbb{B}'$  tem, portanto, complexidade exponencial na maior aridade de  $\beta$  nas universais  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ .

Assim,  $\mathbb{B}'$  se mostra superior a  $\mathbb{A}$  em pelo menos dois aspectos. Primeiro,  $\mathbb{B}'$  computa  $\Gamma_i$  a partir de  $\Gamma_{i-1}$  em tempo polinomial, diferente de  $\mathbb{A}$ , que o faz em tempo exponencial. Além disso, embora  $\mathbb{B}'$  compute  $T(\Gamma)$  a partir de  $\Gamma$  em tempo exponencial, o tempo para realizar tal computação cresce exponencialmente em função da maior aridade de  $\beta$  nas universais  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , e não em função de  $k$ , como ocorre com  $\mathbb{A}$  já na computação de  $\Gamma_i$  a partir de  $\Gamma_{i-1}$ . Em sistemas baseados em conhecimento, envolvendo ou não matemática, constantes predicativas com aridade maior que 3 são ao que parece ainda mais raras que longas sequências de quantificadores.

Todavia, essa superioridade de  $\mathbb{B}'$  só está presente quando, em cada universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ ,  $\alpha_j$  contiver todas as variáveis  $\xi_1, \dots, \xi_k$ , para ao menos um  $j$  tal que  $1 \leq j \leq m$ . Vamos mostrar agora, por meio de exemplos, por que isso é o caso. Considere-se a base de conhecimento  $\Gamma$  contendo o seguinte núcleo:  $\Delta = \{\forall x \forall y \forall z \forall w ((Fx \& Gz \& Ryw \& Szxyw) \rightarrow Hwx), Fa, Fb, Ga, Gb, Raa, Rab, Rba, Rbb, Sbbab, Saaab\}$ . Recebendo  $\Gamma$  como entrada,  $\mathbb{B}'$  reorganiza as fórmulas atômicas no antecedente da única universal em  $\Delta$ , obtendo a fórmula  $\forall x \forall y \forall z \forall w ((Szxyw \& Fx \& Gz \& Ryw) \rightarrow Hwx)$ ; isto feito,  $\mathbb{B}'$  percorre o conjunto  $T$  das fórmulas atômicas de  $\Delta$ , procurando por instâncias de  $\alpha_1 = Szxyw$ , encontrando  $\alpha_1[x/b, y/b, z/a, w/b] = Sbbab$ ; ato contínuo,  $\mathbb{B}'$  percorre  $T$  procurando por  $\beta[x/b, y/b, z/a, w/b] = Hbb$ , e não encontra; na sequência  $\mathbb{B}'$  percorre  $T$  procurando por  $\alpha_2[x/b, y/b, z/a, w/b] = Fb$ , que é encontrada; depois,  $\mathbb{B}'$  percorre  $T$  em busca de  $\alpha_3[x/b, y/b, z/a, w/b] = Ga$ , que é encontrada;  $\mathbb{B}'$  então percorre  $T$  em busca de  $\alpha_4[x/b, y/b, z/a, w/b] = Rbb$ , que também é encontrada;  $\mathbb{B}'$  então acrescenta  $\beta[x/b, y/b, z/a, w/b] = Hbb$  a  $\Delta$ , e reinicia o processo todo, percorrendo  $T$  à procura de outras instância de  $\alpha_1$ , encontrando  $\alpha_1[x/a, y/a, z/a, w/b] = Saaab$ ; como  $Hba \notin \Delta$ , e  $Fa, Ga, Rab \in \Delta$ , ao fim do processo  $\mathbb{B}'$  acrescenta  $\beta[x/a, y/a, z/a, w/b] = Hba$  a  $\Delta$ ; ao percorrer  $T$  novamente em busca de novas instâncias de  $\alpha_1$ ,  $\mathbb{B}'$  não encontra nada, e como não há outras universais em  $\Delta$ , temos então que  $\mathbb{B}'$  obteve  $\Gamma_1$ , contendo o núcleo  $\Delta_1 = \{\forall x \forall y \forall z \forall w ((Fx \& Gz \& Ryw \& Szxyw) \rightarrow Hwx), Fa, Fb, Ga, Gb, Raa, Rab, Rba, Rbb, Sbbab, Saaab, Hbb, Hba\}$ . Como as novas fórmulas atômicas acrescentadas a  $\Delta$  para a obtenção de  $\Delta_1$  não são instâncias de nenhuma fórmula que ocorre no antecedente de nenhuma universal em  $\Gamma$  (e portanto em  $\Gamma_1$ , já que  $\mathbb{B}'$  não acrescenta universais a  $\Gamma_i$  para a obtenção de  $\Gamma_{i+1}$ ), temos que, ao

atuar sobre  $\Gamma_1$ ,  $\mathbb{B}'$  não acrescentará nenhuma nova fórmula atômica a  $\Delta_1$ , obtendo portanto  $\Gamma_2 = \Gamma_1$ ; na sequência  $\mathbb{B}'$  elimina a única universal em  $\Gamma_2$ , obtendo a base vívida  $T(\Gamma)$  com o núcleo  $\Delta' = \{Fa, Fb, Ga, Gb, Raa, Rab, Rba, Rbb, Sbbab, Saaab, Hbb, Hba\}$ .

No exemplo acima,  $\mathbb{B}'$  se comportou exatamente do modo descrito há pouco, de modo que a medida de complexidade que fizemos obviamente se aplica. Entretanto, considere-se agora a base de conhecimento  $\Gamma$  com o núcleo  $\Delta = \{\forall x\forall y\forall z\forall w ((Ryw \& Gz \& Fx) \rightarrow Kwxz), Fa, Fb, Ga, Gb, Raa\}$ , que não possui nenhuma fórmula atômica no antecedente de sua única universal na qual ocorram todas as variáveis quantificadas ali. Ao receber  $\Gamma$  como entrada,  $\mathbb{B}'$  toma sua única universal, verifica que não há em seu antecedente nenhuma fórmula atômica com as características mencionadas, e então percorre o conjunto  $T$  das fórmulas atômicas de  $\Delta$ , à procura por instâncias de  $\alpha_1 = Ryw$ , encontrando  $\alpha_1[x/? , y/a, z/? , w/a]$ ; escrevemos  $x/?$  e  $z/?$  para expressar o fato de que, ao tomar uma dada instância de  $\alpha_1$ ,  $\mathbb{B}'$  teve acesso aos valores de  $y$  e  $w$  em tal instância, mas não aos valores de  $x$  e  $z$ , uma vez que não há ocorrências desas variáveis em  $\alpha_1$ ;  $\mathbb{B}'$  então percorre  $\Delta$  em busca de  $\beta[x/? , y/a, z/? , w/a]$ ; como os valores de  $x$  e  $z$  não estão especificados, podemos modificar  $\mathbb{B}'$ , admitindo que procure por fórmulas da forma  $Kaxz$ , com  $x$  e  $z$  substituídas por quaisquer constantes individuais no dicionário de  $\Delta$ ; como  $\mathbb{B}'$  não encontra uma tal fórmula,  $\mathbb{B}'$  percorre  $\Delta$  em busca  $\alpha_2[x/? , y/a, z/? , w/a] = Gz$ , com  $z$  substituída por qualquer constante individual;  $\mathbb{B}'$  encontra  $Ga$ ; como agora  $z$  foi especificada, faz sentido que nossa modificação de  $\mathbb{B}'$  passe a percorrer  $\Delta$  à procura por  $\alpha_3[x/? , y/a, z/a, w/a]$ , encontrando  $Fa$ ; isto feito, como agora  $x$  também foi especificada,  $\mathbb{B}'$  acrescenta  $\beta[x/a, y/a, z/a, w/a] = Kaaa$  a  $\Delta$ ; se  $\mathbb{B}'$  prosseguir para a próxima instância de  $\alpha_1$ , não vai encontrar nada, e como não há outras universais em  $\Delta$ ,  $\mathbb{B}'$  vai agir como se tivesse obtido  $\Gamma_1$ , repetindo todo o processo com a base de conhecimento de núcleo  $\{\forall x\forall y\forall z\forall w ((Ryw \& Gz \& Fx) \rightarrow Kwxz), Fa, Fb, Ga, Gb, Raa, Kaaa\}$ ; como essa base de conhecimento difere de  $\Gamma$  apenas por uma fórmula atômica que não é instância de nenhuma fórmula atômica no antecedente de sua única universal,  $\mathbb{B}'$  vai obter essa mesma base novamente, e agirá como se tivesse obtido  $\Gamma_n$ , eliminando a universal e liberando como saída a base de conhecimento vívida de núcleo  $\{Fa, Fb, Ga, Gb, Raa, Kaaa\}$ . Mas essa base de conhecimento não é  $T(\Gamma)$ .

Assim  $\mathbb{B}'$  não funciona corretamente com bases de conhecimento cujas universais não satisfazem a condição de que estivemos tratando, mesmo com as modificações que introduzimos de modo que  $\mathbb{B}'$  pudesse lidar com variáveis de valor não especificado<sup>22</sup>. Mas onde está o defeito em  $\mathbb{B}'$  que determina seu mau funcionamento nesses casos? Claramente, deveríamos utilizar um algoritmo  $\mathbb{B}''$  que, depois de acrescentar  $Kaaa$  a  $\Delta$ , retornasse à busca de  $\alpha_3[x/? , y/a, z/a, w/a]$ , de modo a verificar se não há em  $\Delta$  outras instâncias de  $Fx$  além de  $Fa$ ; ao fazê-lo,  $\mathbb{B}''$  encontra  $Fb$ , e acrescenta a  $\Delta$   $b[x/b, y/a, z/a, w/a] = Kaba$ , verificando que ainda não temos que  $Kaba \in \Delta$ ; como só temos as constantes  $a$  e  $b$  no dicionário de  $\Delta$ ,  $\mathbb{B}''$  então não precisa retornar à busca de  $\alpha_3[x/? , y/a, z/a, w/a]$ , mas deve retornar à procura por  $\alpha_2[x/? , y/a, z/? , w/a]$ , de modo a verificar se não há em  $\Delta$

outras instâncias de  $Gz$  além de  $Ga$ ; encontrando  $Gb$ ,  $\mathbb{B}''$  percorre  $\Delta$  em busca de  $\alpha_3[x/?, y/a, z/b, w/a]$ , encontra  $Fa$  e acrescenta  $Kaab$  a  $\Delta$ , após verificar que ainda não temos que  $Kaab \in \Delta$ ; depois,  $\mathbb{B}''$  deve retornar à busca por  $\alpha_3[x/?, y/a, z/b, w/a]$ , de modo a verificar se não há em  $\Delta$  outras instâncias de  $Fx$  além de  $Fa$ ;  $\mathbb{B}''$  encontra  $Fb$  e acrescenta  $Kabb$  a  $\mathbb{B}$ , após verificar que ainda não temos que  $Kabb \in \Delta$ . Procedendo assim,  $\mathbb{B}''$  obtém  $\Gamma_1$ , com o núcleo  $\Delta_1 = \{\forall x\forall y\forall z\forall w ((Ryw \& Gz \& Fx) \rightarrow Kwxyz), Fa, Fb, Ga, Gb, Raa, Kaaa, Kaba, Kaab, Kabb\}$ . Atuando sobre  $\Gamma_1$ ,  $\mathbb{B}''$  obtém  $\Gamma_2 = \Gamma_1 = \Gamma_n$ , e elimina  $\forall x\forall y\forall z\forall w ((Ryw \& Gz \& Fx) \rightarrow Kwxyz)$  de  $\Gamma_n$  de modo a obter a base vívida  $T(\Gamma)$  com o núcleo  $D' = \{Fa, Fb, Ga, Gb, Raa, Kaaa, Kaba, Kaab, Kabb\}$ .

O procedimento de retorno que acrescentamos a  $\mathbb{B}'$  para obter  $\mathbb{B}''$  é bem conhecido, e é chamado de *backtracking*. Assim, há como especificar  $\mathbb{B}''$  de modo preciso, e mostrar que, dado um mesmo input,  $\mathbb{B}''$  devolve o mesmo output que  $\mathbb{A}^{23}$ . Contudo, considere-se, por exemplo, o caso em que cada fórmula atômica no antecedente de cada universal em  $\Gamma$  possui apenas uma variável, e cada uma dessas fórmulas possui uma variável diferente das que ocorrem nas demais ( $\forall x\forall y\forall z\forall w ((Fx \& Gy \& Hz \& Jw) \rightarrow Kwxyz)$ , por exemplo). Em tal situação,  $\mathbb{B}''$  deverá: executar, eventualmente, as mesmas  $m + 1$  ações iniciais de  $\mathbb{B}'$ ; percorrer as  $t$  fórmulas atômicas de  $\Gamma_i$  em busca de instâncias de  $\alpha_j$ ,  $1 \leq j \leq m$ ,  $c^k$  vezes, sendo  $c$  o número de constantes no dicionário de  $D$ ; depois de cada uma dessas  $c^k$  verificações,  $\mathbb{B}''$  deverá percorrer as  $t$  fórmulas atômicas de  $\Gamma_i$  em busca da instância correspondente de  $\beta$ , e, no caso de não encontrá-la, acrescentar a mesma a  $\Gamma_i$ . Portanto,  $\mathbb{B}''$  deve executar no máximo  $2tc^k + c^k + m + 1$  passos, para cada universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , de modo a obter  $\Gamma_{i+1}$  a partir de  $\Gamma_i$ , resultando em um limite de  $2utc^k + uc^k + (m + 1)u$  passos, sendo  $k$  o número de variáveis da universal em  $\Gamma$  com a maior quantidade delas, e  $m = k$  o maior número de fórmulas  $\alpha_j$  no antecedente das universais em  $\Gamma$ . Assim, nos casos sob consideração, supondo  $t, u, c$  constantes, e fazendo  $2ut = a$ , temos que a complexidade do problema de se obter  $\Gamma_i$  a partir de  $\Gamma_{i-1}$  por meio de  $\mathbb{B}''$  é dada pela função  $ac^k + uc^k + (k + 1)u = (a + u)c^k + (k + 1)u$ . Fazendo  $a + u = b$ , obtemos  $f(k) = bc^k + uk + u$ . Como  $bc^k \leq (bc)^k$  para  $k \geq 0$ , temos, fazendo  $bc = d$ , que  $bc^k$  é  $O(d^k)$ , e como,  $uk$  e  $u$  são  $O(d^k)$ , temos, pela regra da soma, que  $f(k)$  é  $O(d^k)$ . Logo, a computação de  $\Gamma_i$  a partir de  $\Gamma_{i-1}$  por meio de  $\mathbb{B}''$ , nos casos sob consideração, tem complexidade exponencial no número máximo  $k$  de variáveis em uma universal  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , tal como ocorre com  $\mathbb{A}$ .

## Referências

- AHO, Alfred & ULLMAN, Jeffrey. *Foundations of computer science*. New York: Computer Science Press, 1995.
- BRACHMAN, Ronald & LEVESQUE, Hector. *Knowledge representation and reasoning*. San Francisco, CA: Morgan Kaufmann, 2004.

GAREY, Michael & JOHNSON, David. *Computers and intractability: A guide to the theory of NP-completeness*. New York: W. H. Freeman and Company, 1997.

LEWIS, Harry & PAPANITRIOU, Christos. *Elements of the theory of computation*. 2.ed. Upper Saddle River, NJ: Prentice Hall, 1998.

MENDELSON, Elliott. *Introduction to mathematical logic*. New York: Van Nostrand Reinhold, 1964.

## Notas

<sup>1</sup> Estamos mantendo inalterada a definição de Brachman e Levesque para o conceito de uma base de conhecimento vívida. Entretanto, a característica de tais bases especificada no item b) da definição dada acima teria já sido incluída no item a) da mesma definição, caso omitíssemos dali a expressão ‘sem igualdade’. De fato, uma vez que não possui fórmulas com ocorrências da constante predicativa ‘=’, o núcleo  $\Delta$  de uma base de conhecimento vívida obviamente não vai acarretar nenhuma fórmula atômica com igualdade, exceção feita para fórmulas da forma  $\bar{c} = \bar{c}$ , para qualquer constante individual  $\bar{c}$  de S. Nesse caso, como  $\Delta \models \bar{c}_i = \bar{c}_j$ , para qualquer par  $(\bar{c}_i, \bar{c}_j)$  de constantes individuais distintas de S, o item a) da definição de uma base de conhecimento vívida, modificado do modo mencionado, já garante que  $\sim \bar{c}_i = \bar{c}_j \in \Gamma$ , para qualquer par  $(\bar{c}_i, \bar{c}_j)$  de constantes individuais distintas de S, tornando desnecessário o item b) da mesma definição.

<sup>2</sup> Obviamente, S deve conter um número finito de constantes individuais (cf. BRACHMAN & LEVESQUE, 2004, p. 337).

<sup>3</sup> Nas demonstrações a seguir supomos familiaridade com noções semânticas relativas à lógica de predicados de 1ª ordem clássica, tais como estruturas, modelos, atribuições, valorações e consequência lógica, bem como com teoremas envolvendo essas noções, tais como os que podem ser encontrados em manuais de lógica elementar como MENDELSON, 1964.

<sup>4</sup> Como  $\alpha_{k+1}$  é fechada, ou  $E \models \alpha_{k+1}$  ou  $E \not\models \alpha_{k+1}$ .

<sup>5</sup> Como na sequência se demonstra que  $\Gamma_w = \Gamma$ , temos que  $\Gamma \models \alpha_i$  sse  $\alpha_i \in \Delta$ .

<sup>6</sup> Supondo como primitivos os operadores lógicos ‘ $\rightarrow$ ’ e ‘ $\sim$ ’, e o quantificador ‘ $\forall$ ’.

<sup>7</sup> L(S) pode ser definida de um modo tal que uma fórmula da forma  $\forall \xi \beta[\xi]$  deva ter ao menos uma ocorrência da variável  $\xi$  em  $\beta[\xi]$ . Do contrário, algumas poucas alterações óbvias na prova acima serão necessárias.

<sup>8</sup> Ou seja, a função  $f(n)$  que determina o tempo  $t$  que o mecanismo inferencial demora para verificar se uma dada fórmula  $\alpha$  se segue da base de conhecimento vívida  $\Gamma$ , em função do tamanho  $n$  dessa base de conhecimento ( $n$  deve ser um inteiro não negativo), é tal que há um polinômio  $p(n)$ , um número real positivo  $k$ , e um número natural  $k_0$  tais que  $f(n) \leq k p(n)$ , para todo  $n \geq k_0$ . Diz-se nesse caso que  $f(n)$  é  $O(p(n))$ , e que o algoritmo que implementa o mecanismo inferencial em questão exibe complexidade polinomial e é computacionalmente tratável. Já se  $f(n)$  for  $O(g(n))$ , para alguma função exponencial  $g(n)$ , e se o fato de que  $f(n)$  é  $O(h(n))$ , para uma função  $h(n)$  qualquer, implicar que  $g(n)$  é  $O(h(n))$ , diz-se que o mesmo exibe complexidade exponencial e é computacionalmente intratável.

<sup>9</sup> A função  $f(m) = 1$  - que determina, em função de  $m$ , o tempo de execução da parte do algoritmo que verifica a forma de  $\alpha$  - é obviamente  $O(1)$ , isto é, essa parte do algoritmo é executada em algum tempo constante independente do número  $m$  de fórmulas em  $\Delta$ . Já a função  $g(m) = m$ , que determina, em função de  $m$ , o tempo de execução da parte do mesmo algoritmo que, para cada fórmula  $\beta_i$  em  $\Delta = \{\beta_1, \beta_2, \dots, \beta_m\}$ , verifica se  $\beta_i = \alpha$ , para  $1 \leq i \leq m$ , é  $O(m)$ . Como  $f(m)$  também é  $O(m)$ , já que, para  $k = 1$  e  $k_0 = 1$ ,  $f(m) \leq k m$ , para todo  $m \geq k_0$ , segue-se pela assim-chamada regra da soma (cf. AHO & ULLMAN,

1995, pp. 105-106) que a função  $g(m) + f(m)$ , ou seja,  $m + 1$ , é  $O(m)$ . Note-se que, por exemplo, quando dizemos que o mecanismo inferencial toma no máximo  $m$  passos para verificar se  $\alpha \in \Delta = \{\beta_1, \beta_2, \dots, \beta_m\}$ , estamos supondo que MI toma um passo para verificar se  $\beta_1 = \alpha$ , isto é, que MI faz essa verificação em algum tempo constante independente do número  $m$  de fórmulas em  $\Delta$ , tal como dissemos ser o caso da verificação da forma de  $\alpha$ . Entretanto, em uma implementação real de MI, usando uma linguagem de programação como C, por exemplo, a verificação de se  $\beta_1 = \alpha$  pode ter de tratar  $\beta_1$  e  $\alpha$  como strings de caracteres, e nesse caso será necessário especificar o comprimento de  $\beta_1$  e  $\alpha$ , e, se essas fórmulas tiverem o mesmo comprimento, compará-las caractere a caractere. A rigor, portanto, MI faz a verificação em questão em um tempo que não depende do número  $m$  de fórmulas em  $\Delta$ , mas que não é constante, uma vez que depende do comprimento de  $\alpha$ . Por uma questão de simplicidade, contudo, vamos supor que seja constante o tempo que MI toma para verificar a forma de uma fórmula, ou se uma dada fórmula pertence a um dado conjunto de fórmulas.

<sup>10</sup> Supondo como primitivos os operadores lógicos '&' e '~', e o quantificador '∨'.

<sup>11</sup> Estamos entendendo que um *sistema de representação do conhecimento*, ou simplesmente um *sistema*, é uma terna ordenada  $(L, BC, I)$ , de uma linguagem formal  $L$ , uma base de conhecimento  $BC \subseteq L$ , e um mecanismo inferencial  $I$ , que, no caso de sistemas baseados em um determinado sistema formal  $F$ , consiste em um algoritmo que recebe como entrada uma fórmula a qualquer de  $L$ , e produz todas as deduções a partir de  $BC$  em  $F$ , de modo a verificar se a se segue de  $BC$  em  $F$  (isto é, se  $BC \models_F \alpha$ ). Genericamente, dizemos que um sistema baseado em um sistema formal qualquer para CP1, como resolução, por exemplo, é um sistema baseado em CP1. Obviamente, se  $F$  é um sistema formal correto, então um sistema de representação do conhecimento  $S = (L, BC, I)$  baseado em  $F$  determina que  $\alpha$  é uma consequência lógica de sua base de conhecimento  $(BC \wedge a)$  sempre que é capaz de determinar que a se segue dessa base de conhecimento em  $F$  ( $BC \models_F \alpha$ ).

<sup>12</sup> O número  $n$  de variáveis proposicionais na base de conhecimento, nesse caso.

<sup>13</sup> Um efeito óbvio dessa limitação é que, se quisermos incluir em uma base de conhecimento vívida a informação de que todos os membros de um conjunto possuem uma dada propriedade  $P$ , teremos que incluir nessa base de conhecimento as fórmulas que representam, para cada um desses indivíduos, a proposição de que o mesmo possui a tal propriedade  $P$ . Isso será trabalhoso no caso de conjuntos finitos com muitos elementos e, evidentemente, impossível no caso de conjuntos infinitos.

<sup>14</sup> Cf. BRACHMAN & LEVESQUE, 2004, p. 328.

<sup>15</sup> Uma cláusula de Horn pode ter  $m = 0$ , e apresentar a forma  $[\beta]$ . Uma fórmula atômica fechada  $\alpha$ , posta em notação clausal, resulta na cláusula  $[\alpha]$ , que é uma cláusula de Horn com  $m = 0$ , e na qual não há ocorrências de variáveis. Uma cláusula de Horn também pode ter  $m \geq 0$  e omitir  $\beta$ : essas cláusulas têm a forma  $[\sim\alpha_1, \dots, \sim\alpha_m]$ , e são chamadas de *cláusulas de Horn negativas*. Uma cláusula negativa notável, com  $m = 0$ , é  $[\ ]$ , a cláusula vazia. Se tomarmos um conjunto irrestrito de cláusulas de Horn, a determinação de se uma fórmula atômica se segue de tal conjunto não é um problema computável (cf. BRACHMAN & LEVESQUE, 2004, p. 94).

<sup>16</sup> Claramente, uma qualquer dessas substituições  $\xi/\bar{c}$  pode também ser representada pela notação  $\xi_1/\bar{c}_1, \xi_2/\bar{c}_2, \dots, \xi_k/\bar{c}_k$ , para uma dada  $k$ -upla ordenada de constantes individuais em  $S$ , sendo  $k$  o número de variáveis na fórmula  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  a cujas componentes atômicas a substituição em questão é aplicada. Portanto, o procedimento descrito acima deve ser repetido, dada cada fórmula  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , para todas as  $k$ -uplas ordenadas de constantes individuais em  $S$ .

<sup>17</sup> Como computar  $T(\Gamma)$  a partir de  $\Gamma$  não é um problema de decisão, a redução polinomial em questão deverá ter como alvo, de fato, o problema de se determinar se uma dada instância  $\beta[\xi_1/\bar{c}^{i1}, \dots, \xi_k/\bar{c}^{ik}]$  de  $\beta$  em uma fórmula do tipo  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$  - para uma dada  $k$ -upla ordenada  $[\bar{c}^{i1}, \dots, \bar{c}^{ik}]$  de constantes individuais no núcleo  $\Delta$  de  $\Gamma$  - é uma consequência lógica de  $\Gamma$ .

<sup>18</sup> Sobre as classes de problemas  $P$  e  $NP$ , e sobre o famoso problema  $P = NP$ , cf. LEWIS & PAPADIMITRIOU, 1998, pp. 275-300. Sobre problemas  $NP$ -completos e reduções polinomiais, cf. LEWIS & PAPADIMITRIOU, 1998, pp. 301-333. Sobre problemas  $NP$ -hard, cf. GAREY & JOHNSON, 1997, pp. 109-120.

<sup>19</sup> Cf. BRACHMAN & LEVESQUE, 2004, p. 91-94.

<sup>20</sup> Os  $m + 1$  primeiros passos não precisam, obviamente, ser repetidos para cada instância de  $\alpha_1$  em  $\Gamma_i$ , de modo que, em vez de  $(m + 1)t^2 + (m + 2)t$ , a medição correta da complexidade da parte de  $\mathbb{B}'$  considerada até aqui é dada pela função  $(m + 1)t^2 + t + m + 1$ . Entretanto, como  $(m + 2)t > t + m + 1$  para  $t > 1$ , e  $(m + 2)t = t + m + 1$  para  $t = 1$ , temos que se a parte em questão de  $\mathbb{B}'$  não toma mais de  $(m + 1)t^2 + t + m + 1$  passos para ser executada, também não tomará mais que  $(m + 1)t^2 + (m + 2)t$  passos para tanto.

<sup>21</sup> Vamos desprezar o fato de que, na verdade,  $n \leq uc^r + 1$ , já que, depois de eventualmente acrescentar as  $uc^r$  fórmulas atômicas a  $\Gamma$  para obter  $\Gamma_{uc^r}$ ,  $\mathbb{B}'$  deve ainda atuar sobre  $\Gamma_{uc^r}$  de modo a verificar que  $\Gamma_{uc^r+1} = \Gamma_n = \Gamma_{uc^r}$ . Além disso, desprezamos também os  $u$  passos que  $\mathbb{B}'$  deve executar sobre  $\Gamma_n$  para eliminar as fórmulas universais de modo a obter  $T(\Gamma)$ . Os passos adicionais decorrentes desses fatos não alteram a conclusão de que a computação de  $T(\Gamma)$  a partir de  $G$  por  $\mathbb{B}'$  é exponencial na maior aridade de  $\beta$  nas universais  $\forall \xi_1 \dots \forall \xi_k ((\alpha_1 \& \dots \& \alpha_m) \rightarrow \beta)$  em  $\Gamma$ , como é fácil verificar.

<sup>22</sup> O sentido em que estamos utilizando as palavras ‘variáveis de valor não especificado’ está bem claro, mas, se quisermos evitar esse modo informal de falar, podemos introduzir o conceito de uma substituição completa  $\xi/\bar{c}$  como sendo uma função de um conjunto de variáveis individuais em um conjunto de constantes individuais, e o conceito de uma substituição incompleta  $\xi/\bar{c}$  como sendo uma relação entre um conjunto de variáveis individuais em um conjunto de constantes individuais, tal que cada variável está relacionada a, no máximo, uma constante. Nesse caso, podemos dizer que  $\mathbb{B}'$  foi modificado acima de modo a poder lidar com substituições incompletas aplicadas às componentes atômicas das universais em  $\Gamma$  (definindo à parte o conceito de aplicação de uma tal substituição a uma fórmula atômica, do modo óbvio).

<sup>23</sup> É possível portanto, obter uma prova completa de que  $\mathbb{A}$  e  $\mathbb{B}''$  são equivalentes, isto é, de que produzem as mesmas saídas para as mesmas entradas, e também de que  $\mathbb{A}$  e  $\mathbb{B}'$  são equivalentes nos casos em que cada universal na entrada  $\Gamma$  possui ao menos uma fórmula atômica no antecedente com ao menos uma ocorrência de cada variável quantificada na universal em questão. Essas provas são um exercício relativamente simples, e não vamos apresentá-las aqui.

# LÓGICAS DE SYLVAN E IMPLICAÇÃO ESTRITA

CEZAR A. MORTARI

Departamento de Filosofia, UFSC

cmortari@cfh.ufsc.br

## 1. Introdução

Em seu artigo “Relational semantics for all Lewis, Lemmon and Feys’s modal logics, most notably for the systems between **S0.3°** and **S1**” (Sylvan 1989), Richard Sylvan procurou apresentar semânticas relacionais para vários sistemas bem conhecidos de lógica modal (sistemas de Lewis, Lemmon e Feys). A estratégia geral foi a de considerar estruturas (*frames*) em que o universo de mundos possíveis está dividido em um conjunto de mundos normais e um ou mais conjuntos de mundos não normais, mundos em que fórmulas modalizadas são valoradas de acordo com diferentes condições. Uma condição bem conhecida, por exemplo, é a de que todas as necessidades são falsas e todas as possibilidades verdadeiras: temos aqui os mundos “estranhos” (*queer*) de Kripke 1965. Outros tipos de mundos não normais considerados por Sylvan são os seguintes:

- *opostos*:  $\Box\alpha$  é verdadeira sse  $\alpha$  é falsa em todos os mundos acessíveis;
- *contrários*:  $\Box\alpha$  é verdadeira sse  $\alpha$  é falsa em algum mundo acessível;
- *perversos*:  $\Box\alpha$  é verdadeira sse  $\alpha$  é verdadeira em algum mundo acessível;
- *arbitrários*:  $\Box\alpha$  é arbitrariamente verdadeira ou falsa.

Combinando esses tipos de mundos nas estruturas, obtemos semânticas para várias lógicas modais não normais. Em particular, Sylvan sustentou que estruturas consistindo em mundos normais e opostos caracterizariam os sistemas **S1°** e **S1**, caso tomemos o conjunto de mundos normais como distinguidos. Contudo, como M. J. Cresswell demonstrou (ver Cresswell 1995), a semântica de Sylvan não caracteriza **S1**, mas um sistema mais forte, **S1+** (a denominação é de Cresswell). Ou seja, a semântica relacional proposta por Sylvan para **S1** e **S1°** é inadequada.

No entanto, estruturas com vários tipos de mundos não normais, e as lógicas determinadas por elas, são certamente interessantes – várias delas são mais fracas que as lógicas normais usuais, o que as torna atrativas se pensarmos em interpretações epistêmicas ou deônticas dos operadores modais. Tanto Sylvan quanto Cresswell, contudo, ocuparam-se de apenas alguns poucos sistemas, e tão somente aqueles cuja semântica toma o conjunto dos mundos normais como distinguido.

Em trabalhos anteriores,<sup>1</sup> examinei várias lógicas – que denominei “lógicas de Sylvan” – obtidas se considerarmos estruturas com vários tipos de mundos não normais sem tomar nenhum desses tipos como distinguidos. No presente trabalho, pretendo investigar algumas dessas lógicas de Sylvan e sua relação com certos sistemas de implicação estrita. A motivação é a semântica relacional apresentada por Sylvan para os sistemas estritos mais fracos, começando com **S0.6**<sup>o</sup>. Mostrarei, contudo, que também nesses casos a semântica relacional proposta é inadequada, e que as lógicas caracterizadas acabam sendo outras, também sistemas mais fortes.

## 2. Lógicas e estruturas

Tomarei como ponto de partida uma linguagem modal proposicional básica consistindo em um conjunto  $\Phi$  de variáveis proposicionais, a constante proposicional  $\perp$ , e os operadores primitivos  $\rightarrow$  e  $\Box$ . Os demais operadores e a constante  $\top$  são definidos do modo usual.

Para os propósitos deste artigo, uma lógica é um conjunto de fórmulas que inclui o conjunto PL de todas as tautologias e é fechado sob modus ponens (MP) e substituição uniforme (SU). Iniciarei por caracterizar as lógicas semanticamente, como conjuntos de fórmulas válidas em certa classe de estruturas que incluirão um ou mais tipos de mundos não normais. De acordo com o tipo de mundos não normais admitidos poderemos ter diferentes tipos de estruturas: talvez apenas mundos normais, talvez opostos e perversos, e assim por diante. No caso geral, uma estrutura é definida assim:

**Definição 1.** Uma *estrutura*  $\mathfrak{F}$  é uma sequência ordenada  $\langle U, N, O, C, P, R \rangle$ , em que  $U$  é um conjunto não vazio, o universo de mundos (ou estados) da estrutura;  $N, C, O$  e  $P$  são subconjuntos de  $U$  dois a dois disjuntos, respectivamente, o conjunto dos mundos normais, opostos, contrários e perversos da estrutura, tais que  $U = N \cup O \cup C \cup P$ . Finalmente,  $R \subseteq U \times U$  é uma relação de acessibilidade.<sup>2</sup>

Permitir que diferentes subconjuntos de  $U$  sejam vazios dá origem a diferentes lógicas. Por exemplo, se todos os subconjuntos, exceto  $N$ , forem vazios, teremos lógicas modais normais. Por outro lado, se tivermos somente mundos opostos, digamos, uma simples tradução nos dirá que fórmulas válidas teremos, já que, nesses mundos,  $\Box$  comporta-se como  $\Box\top$  nas lógicas modais normais. Não considerarei tais casos aqui, pois os casos interessantes surgem com a combinação de mundos normais e diferentes tipos de mundos. Assim, estarei supondo, adicionalmente, que  $N$  não é vazio.

Finalmente, a definição de estrutura acima apresentada é mais geral do que a de Sylvan, uma vez que não exige, de início, sua *condição crítica*: de que todo elemento de  $U$  seja acessível a algum mundo normal. (Estruturas com tal característica serão consideradas logo em seguida.)

**Definição 2.** Se  $\mathfrak{F}$  é uma estrutura,  $\mathfrak{M} = \langle \mathfrak{F}, V \rangle$  é um *modelo*, em que  $V$  é uma valoração, ou seja, uma função de  $\Phi$  em  $\mathfrak{B}(U)$ . Dizemos, nesse caso, que o modelo  $\mathfrak{M}$  é *baseado em*  $\mathfrak{F}$ .

A definição usual de verdade em um mundo de um modelo é a usual, exceto pelas cláusulas para fórmulas modalizadas, em que a condição de verdade dependerá de que espécie de mundo se trata.

**Definição 3.** Sejam  $\mathfrak{M} = \langle \mathfrak{F}, V \rangle$  um modelo e  $x$  um mundo em  $\mathfrak{M}$ . Então:

- (a)  $\mathfrak{M}, x \Vdash \mathbf{p}$  sse  $x \in V(\mathbf{p})$ , para  $\mathbf{p} \in \Phi$ ;
- (b)  $\mathfrak{M}, x \not\Vdash \perp$ ;
- (c)  $\mathfrak{M}, x \Vdash \alpha \rightarrow \beta$  sse  $\mathfrak{M}, x \not\Vdash \alpha$  ou  $\mathfrak{M}, x \Vdash \beta$ ;
- (d)  $\mathfrak{M}, x \Vdash \Box \alpha$  sse
  - $x \in N$  e, para todo  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \Vdash \alpha$ , ou
  - $x \in O$  e, para todo  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \not\Vdash \alpha$ , ou
  - $x \in C$  e, para algum  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \not\Vdash \alpha$ , ou
  - $x \in P$  e, para algum  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \Vdash \alpha$ .

Obviamente, as condições de verdade para fórmulas com o operador de possibilidade são as seguintes:

- (e)  $\mathfrak{M}, x \Vdash \Diamond \alpha$  sse
  - $x \in N$  e, para algum  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \Vdash \alpha$ , ou
  - $x \in O$  e, para algum  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \not\Vdash \alpha$ , ou
  - $x \in C$  e, para todo  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \not\Vdash \alpha$ , ou
  - $x \in P$  e, para todo  $y$  tal que  $Rxy$ ,  $\mathfrak{M}, y \Vdash \alpha$ .

**Definição 4.** Uma fórmula  $\alpha$  é *válida* em uma estrutura  $\mathfrak{F}$  sse para todo mundo  $x$  de  $\mathfrak{F}$  e todo modelo  $\mathfrak{M}$  baseado em  $\mathfrak{F}$ ,  $\alpha$  é verdadeira em  $x$  (ou seja,  $\mathfrak{M}, x \Vdash \alpha$ ). Dizemos que  $\alpha$  é *válida* em uma classe  $C$  de estruturas ( $\Vdash_C \alpha$ ) se é válida em toda estrutura em  $C$ .

Finalmente, definimos como segue a noção de consequência lógica:

**Definição 5.** Seja  $C$  uma classe de estruturas,  $\Gamma$  um conjunto de fórmulas, e  $\alpha$  uma fórmula. Dizemos que  $\alpha$  é *consequência lógica de*  $\Gamma$  em  $C$  ( $\Gamma \Vdash_C \alpha$ ) sse para todos os modelos  $\mathfrak{M} = \langle \mathfrak{F}, V \rangle$  baseados em uma estrutura  $\mathfrak{F}$  em  $C$ , e todos os mundos  $x$  em  $\mathfrak{M}$ , se  $\mathfrak{M}, x \Vdash \gamma$  para toda  $\gamma \in \Gamma$ , então  $\mathfrak{M}, x \Vdash \alpha$ .

A noção de consequência lógica aqui apresentada, evidentemente, é a noção local.

Dadas essas definições, a proposição a seguir é facilmente demonstrada.

**Proposição 6.** Seja  $C$  a classe de todas as estruturas. Então:

- (i) se  $\alpha$  é uma consequência tautológica de  $\alpha_1, \dots, \alpha_m$  ( $m \geq 0$ ), e  $\Vdash_C \alpha_1, \dots, \Vdash_C \alpha_m$ , então  $\Vdash_C \alpha$ ;

(ii) se  $\Vdash_c \alpha \leftrightarrow \beta$ , então  $\Vdash_c \Box\alpha \leftrightarrow \Box\beta$ .

Em decorrência disso, vemos que a lógica da classe de todas as estruturas é uma lógica modal clássica, pois (i) inclui as tautologias e (ii) é fechada sob a regra de inferência RE ( $\alpha \leftrightarrow \beta / \Box\alpha \leftrightarrow \Box\beta$ ).

Como mencionado acima, contudo, Sylvan considerou estruturas satisfazendo a seguinte condição crítica:

$$\mathbf{cc}. \forall x \exists y (y \in N \wedge Ryx).$$

Ou seja, cada mundo na estrutura é acessível a algum mundo normal.

Indicando por  $\Vdash_{cc}$  validade e consequência lógica na classe das estruturas satisfazendo a condição crítica, podemos demonstrar sem muita dificuldade o seguinte:<sup>3</sup>

**Teorema 7.**

- (a)  $\Vdash$  a sse  $\Vdash_{cc} \alpha$ .  
 (b)  $\Gamma \Vdash \alpha$  sse  $\Gamma \Vdash_{cc} \alpha$ .

Em decorrência disso, nossa linguagem não é suficiente para distinguir a classe das estruturas com **cc** da classe de todas as estruturas. Evidentemente, se dispuséssemos de um operador  $\blacksquare$  cujas condições de verdade fossem

$$x \Vdash \blacksquare\alpha \text{ sse para todo } y \in N, \text{ se } Ryx \text{ então } y \Vdash \alpha,$$

teríamos  $\Vdash_{cc} \blacksquare\Box\alpha \rightarrow \alpha$  mas  $\not\vdash \blacksquare\Box\alpha \rightarrow \alpha$ , o que nos permitiria distinguir estruturas satisfazendo **cc** das demais.

A condição **cc** é o que se costuma denominar uma condição para uma *extensão segura* de uma estrutura. Considerando estruturas com **cc**, é fácil mostrar que valem as regras:

$$(\text{RN}^*) \quad \Box\alpha / \alpha,$$

$$(\text{RnN}^*) \quad \Box\top \rightarrow \Box\alpha / \alpha,$$

regras que serão necessárias posteriormente (por exemplo, na prova do Teorema 8).

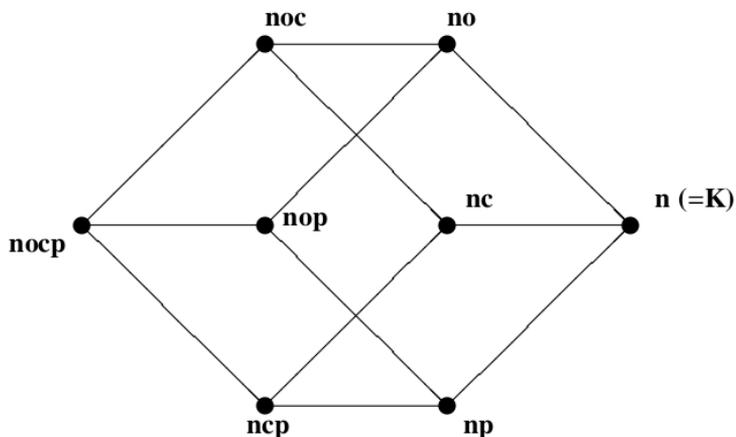
Em função do Teorema 7 acima, passaremos a considerar somente estruturas em que **cc** vale, também de acordo com a especificação de Sylvan.

### 3. A lógica de base nocp e suas extensões

Definindo validade em uma estrutura como verdade em *todos* os mundos de todos os modelos baseados na estrutura, e tomando estruturas em que, como na Definição 1, aparecem todos os tipos de mundos não normais aqui

considerados, obtemos uma primeira lógica de Sylan, que denominaremos **nocp**. Evidentemente, se eliminarmos um ou mais espécies de mundos não normais – embora mantendo sempre  $N$  – teremos um total de oito lógicas de Sylan distintas, dependendo de que tipos de mundos estão presentes em um estrutura. Para padronizar a notação, além de **nocp**, teremos **noc**, **nep**, etc., até ficarmos com **n** – a lógica da classe de estruturas contendo apenas mundos normais, que é a menor lógica normal, **K**.

O diagrama a seguir apresenta as relações entre as 8 lógicas, começando com **nocp**, a mais fraca, do lado esquerdo, até **n**, do lado direito do diagrama.



Podemos axiomatizar **nocp** acrescentando os seguintes axiomas à menor lógica modal clássica, **E** (uma lógica modal é dita *clássica* se fechada sob a regra de inferência RE):

- (nM)  $\Box \top \rightarrow (\Box(\alpha \wedge \beta) \rightarrow \Box\alpha \wedge \Box\beta)$ ,
- (qW)  $\Diamond \perp \rightarrow (\Box(\alpha \vee \beta) \rightarrow \Box\alpha \wedge \Box\beta)$ ,
- (NP)  $\Box \top \rightarrow (\Box\alpha \wedge \Box\beta \rightarrow \Box(\alpha \wedge \beta)) \vee (\Box(\gamma \vee \delta) \rightarrow \Box\gamma \vee \Box\delta)$ ,
- (OC)  $\Diamond \perp \rightarrow (\Box\alpha \wedge \Box\beta \rightarrow \Box(\alpha \vee \beta)) \vee (\Box(\gamma \wedge \delta) \rightarrow \Box\gamma \vee \Box\delta)$ .

Além dos axiomas acima, podemos mostrar que valem em **nocp** os esquemas:

- (nV)  $\Box \top \rightarrow (\Box\alpha \wedge \Box\beta \rightarrow \Box(\alpha \vee \beta))$ ,
- (qC)  $\Diamond \perp \rightarrow (\Box\alpha \wedge \Box\beta \rightarrow \Box(\alpha \wedge \beta))$ ,
- (qK)  $\Diamond \perp \rightarrow (\Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta))$ ,
- (qA)  $\Diamond \perp \rightarrow (\Box\alpha \vee \Box\beta \rightarrow \Box(\alpha \wedge \beta))$ .

Do mesmo modo, **nocp** tem também as regras:

- (RnM)  $\alpha \rightarrow \beta / \Box \top \rightarrow (\Box\alpha \rightarrow \Box\beta)$ ,
- (RqW)  $\alpha \rightarrow \beta / \Diamond \perp \rightarrow (\Box\beta \rightarrow \Box\alpha)$ ,

$$(RRE) \quad \alpha \leftrightarrow \beta / \gamma \leftrightarrow \gamma[\alpha/\beta].$$

Esta última, RRE, é a regra de substituição de equivalentes, em que  $\gamma[\alpha/\beta]$  é a fórmula obtida de  $\gamma$  substituindo-se ocorrências de  $\alpha$  por  $\beta$ .

Por outro lado, **nocp** não tem nenhum dos esquemas a seguir:

$$(nC) \quad \Box\top \rightarrow (\Box(\alpha \wedge \beta) \rightarrow \Box\alpha \wedge \Box\beta),$$

$$(nK) \quad \Box\top \rightarrow (\Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta)),$$

$$(M) \quad \Box(\alpha \wedge \beta) \rightarrow (\Box\alpha \wedge \Box\beta),$$

$$(C) \quad (\Box\alpha \wedge \Box\beta) \rightarrow \Box(\alpha \wedge \beta),$$

$$(K) \quad \Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta),$$

$$(X) \quad (\Box(\alpha \rightarrow \beta) \wedge (\Box(\beta \rightarrow \gamma))) \rightarrow \Box(\alpha \rightarrow \gamma),$$

$$(N) \quad \Box\top.$$

Vejam os que relações **nocp** tem com algumas outras lógicas conhecidas. Uma lógica modal é *clássica* se fechada sob RE; *monotônica*, se é clássica tem o esquema M, e *regular* se é clássica e tem os esquemas M e C. Lógicas *pré-normais* foram introduzidas por Chellas e Segerberg (em 1996); são as lógicas clássicas que tem o esquema nK, e a menor delas é o sistema **P**. Finalmente, lógicas de Cresswell, também introduzidas em Chellas e Segerberg (1996), são as extensões de **P** por meio do esquema

$$(X_n) \quad \Box(\alpha \rightarrow \beta_1 \wedge \dots \wedge \beta_n) \wedge \Box(\beta_1 \rightarrow \gamma_1) \wedge \dots \wedge \Box(\beta_n \rightarrow \gamma_n) \rightarrow \Box\top, \text{ para } n \geq 1.$$

É fácil ver que **nocp** é uma lógica modal *clássica*, mas não monotônica nem regular e nem pré-normal, já que M, C e nK não valem. Igualmente, não é uma lógica de Cresswell. Por outro lado, **nocp** tem nM e qW, nenhuma das quais vale em **E**. Assim, **nocp** é uma extensão própria de **E**.

Em função das limitações de espaço, não poderei apresentar aqui axiomatizações das demais lógicas de base apresentadas no diagrama anterior. Todas elas são lógicas modais clássicas, mas não são clássicas estritas, que é o que nos interessa neste artigo.

## 4. Lógicas de Lewis

Se **L** é uma lógica, **Lew(L)** é o menor conjunto de fórmulas que contém **L**,  $\Box\top$  e é fechado sob MP e SU. Semanticamente, em muitos casos isso corresponde a tomar algum subconjunto do universo  $U$  de uma estrutura como distinguidos – por exemplo, **S2°** é caracterizada pelas mesmas estruturas que **E2°**, tomando os mundos normais como distinguidos. Assim, uma fórmula é válida em uma estrutura se verdadeira nos mundos normais de todos os modelos baseados na estrutura. Como veremos depois, isso não funciona em todos os casos – precisaremos também considerar os mundos perversos das estruturas.

Em geral, lógicas de Lewis *não são* clássicas, já que não são fechadas sob RE. Mesmo que  $\alpha \leftrightarrow \beta$  seja válida (isto é, verdadeira em todo mundo normal),  $\Box\alpha \leftrightarrow \Box\beta$  pode falhar na medida em que  $\alpha$  e  $\beta$  recebam valores diferentes em algum mundo não normal. (RE não é uma regra válida em, por exemplo, **S2**°.)

Por outro lado, as versões de Lewis são lógicas *clássicas estritas* se a lógica de base é clássica, tendo:

$$\begin{aligned} (\Box\text{PL}) \quad & \{\Box\alpha: \alpha \text{ é uma tautologia}\}, \\ (\text{RSE}) \quad & \alpha \Leftrightarrow \beta / \Box\alpha \Leftrightarrow \Box\beta. \end{aligned}$$

RSE é a contraparte estrita da regra RE; a equivalência estrita de duas fórmulas  $\alpha$  e  $\beta$  está sendo aqui representada por  $\alpha \Leftrightarrow \beta$  e pode ser definida como  $\Box(\alpha \leftrightarrow \beta)$ , a necessidade da equivalência material. Analogamente, representarei por  $\Rightarrow$  a implicação estrita, que pode ser definida como a necessidade da implicação material,  $\Box(\alpha \rightarrow \beta)$ .<sup>4</sup>

Definimos uma lógica modal como *clássica estrita* se for fechada sob a substituição de equivalentes estritos, ou seja:

$$(\text{RRSE}) \quad \alpha \Leftrightarrow \beta, \gamma / \gamma[\alpha/\beta].$$

Em Chellas e Segerberg (1996) temos o teorema a seguir, que estabelece certas relações entre lógicas clássicas e suas versões de Lewis:

**Teorema 8.** (Chellas–Segerberg) Seja **L** uma lógica modal clássica.

- (i)  $\Vdash_{\mathbf{L}} \Box\top \rightarrow \alpha$  sse  $\Vdash_{\text{Lew}(\mathbf{L})} \alpha$ ;  
(ii) se **L** tem  $\text{RnN}^*$ ,  $\Vdash_{\mathbf{L}} \alpha$  sse  $\Vdash_{\text{Lew}(\mathbf{L})} \Box\alpha$ .

Em particular, temos então que:

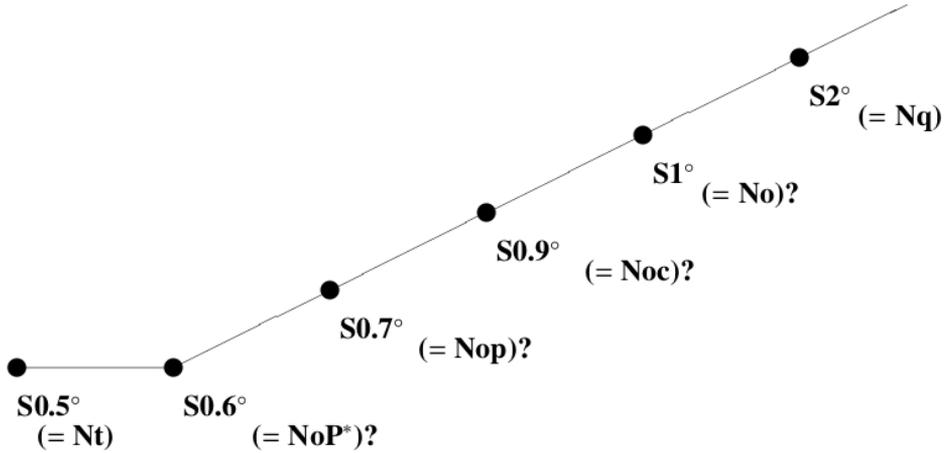
$$\Vdash_{\mathbf{L}} \alpha \rightarrow \beta \text{ sse } \Vdash_{\text{Lew}(\mathbf{L})} \alpha \Rightarrow \beta.$$

Em seu artigo, Sylvan começa apresentando uma semântica relacional para uma lógica clássica estrita básica – a menor de tais lógicas, denominada **S0.6**°, uma extensão do cálculo proposicional clássico que pode ser axiomatizada, por exemplo, acrescentando-se as regras a seguir:

$$\begin{aligned} (\text{RP1}) \quad & \alpha \Rightarrow \beta, \Box\alpha / \Box\beta, \\ (\text{RN+}) \quad & \alpha / \Box\alpha, \text{ se } \alpha \text{ é uma tautologia}, \\ (\text{RN}^*) \quad & \Box\alpha / \alpha, \\ (\text{RSE}) \quad & \alpha \Leftrightarrow \beta / \Box\alpha \Leftrightarrow \Box\beta. \end{aligned}$$

O diagrama a seguir mostra as relações entre algumas das lógicas estritas conhecidas. Também foi incluído, para fins de comparação, o sistema **S0.5**°, que não é uma lógica estrita. Tal sistema poderia ser denominado **Nt**, pois é caracterizados por estruturas em que temos mundos normais como distinguidos (N) e também mundos arbitrários (t). Ao lado do nome padrão, temos a sugestão de Sylvan de qual semântica determinaria a lógica. Por exemplo, **S0.6**° seria **NoP**\*,

ou seja, a lógica das estruturas que tem mundos normais, opostos e perversos, mas em que apenas os mundos normais e perversos são distinguidos (daí estarem  $N$  e  $P$  em maiúsculas na denominação da lógica). Além disso,  $P^*$  representa a exigência da condição de serialidade para os mundos em  $P$ : para todo  $x$  em  $P$ , existe algum  $y$  tal que  $Rxy$ . (Isso é necessário para validar  $\Box\top$  nos mundos perversos.)



É sabido que  $S2^\circ$  é determinada por estruturas com mundos normais distinguidos ( $N$ ) e os mundos “estranhos” ( $q$ ) de Kripke. Contudo, como mencionei no início deste artigo, M. J. Cresswell demonstrou que a conjectura de Sylvan de que  $S1^\circ$  seria  $No$  estava equivocada (ver Cresswell 1995). Na verdade,  $No = S1^{\circ+}$ , ou seja,  $S1^\circ$  acrescido dos seguintes esquemas:

$$(\Box\alpha \wedge \Box\beta) \Rightarrow \Box(\alpha \vee \beta),$$

$$(\Diamond\perp \wedge \Box\alpha \wedge \beta) \Rightarrow \Box(\alpha \wedge \beta).$$

## 5. Lógicas de Sylvan estritas

Consideremos então a semântica apresentada por Sylvan para os demais sistemas. A tese de Sylvan é que a semântica relacional por ele apresentada para os sistemas  $S0.6^\circ$ ,  $S0.7^\circ$  e  $S0.9^\circ$  corresponderia a  $NoP^*$ ,  $Nop$  e  $Noc$  – mas isso não é o caso. Vejamos então como ficam as coisas.

Em primeiro lugar, uma vez que o axioma  $N$  vale em lógicas estritas, precisamos requerer nas estruturas que os mundos normais  $N$  e perversos  $P$  com serialidade sejam distinguidos, pois  $N$  falha em mundos opostos e contrários. A serialidade para os mundos em  $P$  é necessária pois, em um mundo perverso para o qual não há mundos acessíveis,  $\Box\top$  é automaticamente falsa (o operador de necessidade comporta-se em mundos perversos como o operador de

possibilidade em mundos normais). Pareceria assim que a lógica estrita mais básica definida sobre a classe de todas as estruturas seria **NocP\*** (e não **NoP\***, a sugestão de Sylvan, que desconsidera mundos contrários). É fácil mostrar por um argumento semântico que, primeiro, em **NoP\*** – *mas não em NocP\** – valem os esquemas a seguir:

$$\begin{aligned} (\Box V) \quad & (\Box\alpha \wedge \Box\beta) \Rightarrow \Box(\alpha \vee \beta), \\ (\Box X_n) \quad & (\alpha \Rightarrow \beta_1 \wedge \dots \wedge \beta_n) \wedge (\beta_1 \Rightarrow \gamma_1) \wedge \dots \wedge (\beta_n \Rightarrow \gamma_n) \Rightarrow \Box\top, \text{ para } n \geq 1. \end{aligned}$$

Adicionalmente, o que também podemos mostrar construindo contra-exemplos, **NocP\*** é uma lógica diferente de **S0.6°**, já que em **NocP\***, mas não em **S0.6°**, valem:

$$\begin{aligned} (M) \quad & \Box(\alpha \wedge \beta) \rightarrow (\Box\alpha \wedge \Box\beta) \\ (V) \quad & (\Box\alpha \wedge \Box\beta) \rightarrow \Box(\alpha \vee \beta) \end{aligned}$$

Assim, as relações de inclusão entre essas lógicas são:

$$\mathbf{S0.6^\circ} \subset \mathbf{NocP^*} \subset \mathbf{NoP^*}$$

Talvez se pudesse pensar agora que a lógica de Sylvan correspondente a **S0.6°** seria dada ou tomando-se a versão de Lewis de **nocp**, ou tomando-se a lógica gerada se considerarmos o conjunto dos mundos normais como distinguido – **Nocp**. Que gerar a versão de Lewis de uma lógica **L** não é equivalente a tomar o conjunto de mundos normais como distinguido é mostrado pelo caso a seguir. **Lew(nocp)** não é a mesma lógica que **Nocp**. Neste segundo sistema, mas não em **Lew(nocp)**, são válidos:

$$\begin{aligned} (C) \quad & (\Box\alpha \wedge \Box\beta) \rightarrow (\Box\alpha \wedge \beta), \\ (K) \quad & \Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta), \\ (X) \quad & \Box(\alpha \rightarrow \beta) \wedge \Box(\beta \rightarrow \gamma) \rightarrow \Box(\alpha \rightarrow \gamma). \end{aligned}$$

Isso é fácil ver, pois todos os esquemas acima valem em mundos normais, não importando que tipo de mundo sejam a eles acessíveis. Podemos mostrar também que:

**Teorema 9.** **Lew(nocp)** e **NocP\*** são a mesma lógica.

O que acontece, então, com **S0.6°**?

Recordemos as lógicas pré-normais de Chellas e Segerberg (1996). Os autores apresentaram seis sistemas distintos, a saber, **P**, a menor lógica pré-normal e suas extensões **PK** e **PX** (bem como as extensões de cada um desses sistemas acrescentando-se o axioma T,  $\Box\alpha \rightarrow \alpha$ ). Os autores demonstraram então que

$$\begin{aligned} \mathbf{Lew(PK)} &= \mathbf{S0.9^\circ}, \quad \mathbf{Lew(PKT)} = \mathbf{S0.9}, \\ \mathbf{Lew(PX)} &= \mathbf{S1^\circ}, \quad \mathbf{Lew(PXT)} = \mathbf{S1}. \end{aligned}$$

E embora isso não tenha ficado explícito no texto, temos também que **Lew(P)** = **S0.7°**, e **Lew(PT)** = **S0.7**.

Evidentemente, parece então ser natural a conjectura de que **S0.6°** seja a versão de Lewis da menor lógica modal clássica **E**.

**Lew(E)**, claro, é o fecho por MP de  $\mathbf{E} \cup \{\Box\top\}$ . Note-se, porém, que tal lógica não deve ser confundida com a lógica modal clássica **EN**. Esta lógica é uma extensão de **E** através do acréscimo de  $\mathbf{N}$  (ou seja,  $\Box\top$ ) como um novo axioma – mas em **EN** preservamos as regras de inferência de **E**, em particular, RE. Note-se ainda que **EN** é uma versão de Lewis de si mesma, ou seja, **EN** = **Lew(EN)**; assim, **EN** não é só *clássica* como também *clássica estrita*. Contudo, **EN** e **Lew(E)** não são a mesma lógica. **EN** é sabidamente fechada sob a regra de necessitação  $\mathbf{RN}^*$  ( $\alpha / \Box\alpha$ ), o que não acontece com **Lew(E)**.

No entanto, se a menor lógica estrita determinada pela semântica proposta por Sylvan é **NocP\***, que é ainda mais forte que **Lew(E)**, quais seriam as estruturas para determinar esta última lógica?

A resposta é que precisamos abandonar a semântica relacional para tanto, passando a uma semântica de vizinhanças com mundos distinguidos.

**Definição 10.** Uma estrutura para **Lew(E)** é uma tripla  $\mathfrak{F} = \langle U, N, S \rangle$ , tal que

- (i)  $U$  é um conjunto não vazio;
- (ii)  $S$  é uma função que associa a cada  $x \in U$  um conjunto de subconjuntos de  $U$ ;
- (iii) para ao menos um  $x \in U$ ,  $U \in S(x)$ ;
- (iv)  $N = \{x \in U : U \in S(x)\}$ .

Os elementos de  $N$  são distinguidos na estrutura, ou seja, a validade em uma estrutura é definida somente a partir de elementos de  $N$ . Não há espaço para mostrar isso, mas pode-se demonstrar que:

**Teorema 11.** **Lew(E)** é determinada pela classe das estruturas acima definidas.

Finalmente, podemos demonstrar que **Lew(E)** de fato é a menor lógica estrita tal como apresentada por Sylvan, o sistema **S0.6°**.

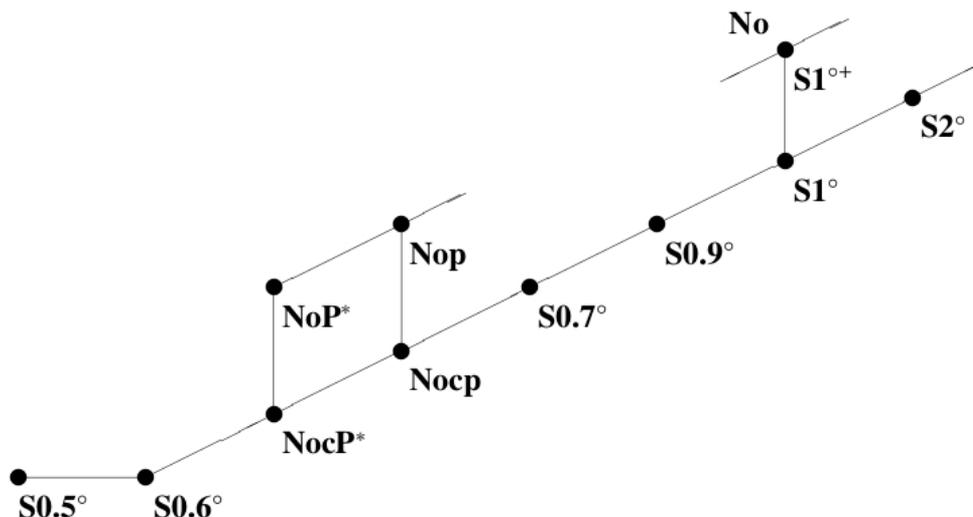
Quanto a **NocP\***, também é equivalente a uma versão de Lewis de uma lógica básica, a saber, **Lew(EnM)**, em que **EnM** é obtida acrescentando-se a **E** o esquema

$$(nM) \quad \Box\top \rightarrow (\Box(\alpha \wedge \beta) \rightarrow (\Box\alpha \wedge \Box\beta)).$$

As estruturas para **Lew(EnM)** são definidas como para **Lew(E)**, exigindo-se adicionalmente que mundos distinguidos (elementos de  $N$ ) satisfaçam a condição (m), isto é, onde  $X$  e  $Y$  são subconjuntos quaisquer de  $U$ ,

$$(m) \quad \text{se } X \cap Y \in S(x) \text{ então } X \in S(x) \text{ e } Y \in S(x).$$

As relações entre alguns dos sistemas aqui investigados aparecem na figura a seguir:



Em conclusão, embora a semântica apresentada por Sylan para sistemas como **S0.6°** e **S0.7°** não fosse adequada, mesmo assim obtivemos outros sistemas de implicação estrita que estendem **S0.6°**. O quadro, contudo, não está completo, e há ainda muitos outros sistemas a investigar.

## Referências

- Chellas, B. F. & Segerberg, K. 1996. Modal Logics in the Vicinity of S1. *Notre Dame Journal of Formal Logic* 37(1): 1-25.
- Cresswell, M. J. 1995. S1 is not so simple. Em W. Sinnott-Armstrong, D. Raffman, e N. Asher (eds.) *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus*. Cambridge: Cambridge University Press, pp.29-40.
- Kripke, S. 1965. Semantical analysis of modal logic II. Non-normal modal propositional calculi. Em J. Addison, L. Henkin, e A. Tarski (eds.) *The Theory of Models: Proceedings of the 1963 International Symposium at Berkeley*. Amsterdam: North-Holland, pp.206-20.
- Mortari, C. A. 2008. Relational Semantics for Some Non-normal Logics: opposite logics. *Annals of the CLE 30 Years XV Brazilian Logic Conference XIV Latin-American Symposium on Mathematical Logic*. Campinas: Center for Logic, Epistemology and History of Science/State University of Campinas, p.149.
- \_\_\_\_\_. 2010. Lógicas de Sylan. *ANPOF Atas do XIV Encontro Nacional de Filosofia*. Campinas: Associação Nacional de Pós-Graduação em Filosofia, p.99-100.
- Sylvan, R. 1989. Relational semantics for all Lewis, Lemmon and Feys' modal logics, most notably for the systems between **S0.3°** and **S1**. *The Journal of Non-Classical Logic* 6(2): 19-40.

## Notas

<sup>1</sup> Uma monografia sobre “Lógicas de Sylvan” encontra-se ainda em preparação. Resultados parciais foram apresentados sob a forma de comunicações em eventos, cujos resumos estão publicados em Mortari 2008, 2010.

<sup>2</sup> Não considerarei neste trabalho estruturas envolvendo mundos estranhos nem arbitrários.

<sup>3</sup> Dadas as limitações de espaço, não é possível incluir neste texto as demonstrações dos teoremas apresentados.

<sup>4</sup> Alternativamente, pode-se definir  $\alpha \Leftrightarrow \beta$  como  $(\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)$ . Ver Chellas e Segerberg 1996.

# ANÁLISE CONCEITUAL DO QUANTIFICADOR “POUCOS” E UMA LÓGICA GERADA POR UM DESSES CONCEITOS.

KLEIDSON ÊGLICIO CARVALHO DA SILVA OLIVEIRA

UNESP, Campus Marília

kecso10@yahoo.com.br

**Resumo:** Em sua Tese de doutorado, Grácio (1999), inspirada pelos trabalhos de Carnielle, Sette e Veloso (1999) e Carnielle e Veloso (1997), apresentou uma família de lógicas – lógicas moduladas – cuja função é formalizar sentenças que expressam quantificações da linguagem natural, que não podem ser definidas em função dos quantificadores da lógica clássica de primeira ordem. Dentre as lógicas moduladas, destaca-se a lógica do muito, que formaliza expressões do tipo “muitos  $x$  satisfazem  $\varphi(x)$ ” através de  $Gx \varphi(x)$ , na qual  $G$  é um novo quantificador inserido na lógica clássica de primeira ordem. Neste trabalho, apresentamos algumas considerações acerca de um quantificador com características duais ao “muitos”, o quantificador “poucos”. Comparamos como diferentes noções de “poucos” geraria diferentes tipos de lógica e desenvolvemos uma lógica com uma noção escolhida.

## 1. Introdução

A questão da quantificação é um tema de interesse de várias áreas do conhecimento, tais como a Lógica, a Linguística, a Matemática e a Computação. Na lógica clássica, tanto de primeira ordem como de ordens superiores, a quantificação é formalizada por meio de dois quantificadores denominados *quantificadores lógicos*: o quantificador existencial ( $\exists$ ) e o quantificador universal ( $\forall$ ). Apesar de possuir apenas estes dois, outros quantificadores podem ser definidos por meio deles, aumentando o escopo das formalizações possíveis nesta lógica.

Um dos primeiros trabalhos acerca de quantificadores diferentes dos usuais foi de Mostowski (1957), onde temos uma das primeiras definições de *quantificadores generalizados*, a partir deste trabalho, muitos outros autores se dedicaram a questão da quantificação. Barwise e Cooper (1981) deram grande contribuição a esta área ao estudar a relação entre os quantificadores das linguagens naturais e os quantificadores lógicos. A autora Sanz, em 2007, fez uma síntese de alguns dos mais importantes trabalhos já publicados sobre a área da quantificação.

Embora a quantificação seja um tema de interesse de vários autores, não há muitas referências na literatura acerca do significado do quantificador “poucos”, Barwise e Cooper (1981), por exemplo, quase não abordam ou comentam sobre o

significado deste quantificador, apenas afirmam que a expressão “poucos”, por si só, não representa um quantificador, mas um determinante. Isso porque não se pode diferenciar expressões do tipo “poucas flechas de John” de “pouco de todas as coisas”. Porém, neste trabalho, a expressão “poucos”, por si só, será considerada um quantificador, posto que se tem sempre um universo de discurso definido.

Os estudos de Peterson (1979) esclarecem como a oposição entre os termos “muitos” e “poucos” ocorre, porém, não define nem um nem outro. Para o autor, uma sentença do tipo “poucos S são P” seria oposta a outra do tipo “muitos S são P”, sendo necessariamente sentenças contraditórias. Este fato vem da hipótese conceitual de que “muitos S são P” possui o mesmo significado que “poucos S são não-P”. Deste modo, se “muitos x têm a propriedade y”, então se pode afirmar que “não poucos x têm a propriedade y”.

Tendo como base a Tese de doutorado de Grácio (1999), mais propriamente a lógica do muito apresentada, estudamos como o quantificador “poucos” pode se opor ao “muitos” de algumas formas diferentes, escolhendo uma dentre elas e gerando uma lógica com aspectos duais a lógica do muito, denominada, *lógica do poucos*. Para isso, consideramos a relação entre os quantificadores “muitos” e “poucos” de maneira um pouco diferente da adotada por Peterson (1979). Um conjunto “não tem poucos” elementos se, e somente se, tiver “muitos” elementos ou for vazio. Do mesmo modo, um conjunto tem “poucos” elementos se, e somente se, não tem “muitos” elementos e não for vazio.

## 2 Noções preliminares

Neste trabalho, apresentamos a formalização de um raciocínio indutivo usado quando houver proposições do tipo “poucos”. Mais especificamente, trata-se de algum “comportamento incomum” entre os indivíduos de um universo, ou um “comportamento não frequente”, posto que não seja usual para os indivíduos daquele universo. Apesar de não podermos atribuir um cardinal a noção de “poucos”, podemos atribuir algumas características básicas que queremos que esta noção possua.

Na sua tese de doutorado, Grácio (1999) formalizou argumentos indutivos, como “muitos” para a “lógica do muito”, introduzida como um caso da família das lógicas moduladas, também definida na tese.

Neste trabalho, propomos a formalização de um novo quantificador com aspectos duais ao quantificador “muitos” de Grácio (1999), denominado quantificador “poucos”.

De maneira análoga à usada por Grácio (1999) para a noção de “muitos”, a noção de “poucos”, é uma noção abstrata, vaga e mais flexível do que seria a noção de “minoría”. O conceito de “minoría” está sempre associado ao conceito de “menor parte”, ao passo que a noção de “poucos” não se atém à noção de

cardinal, pois está associada à noção de um *conjunto pequeno de evidências*, é uma noção relativa, não absoluta, de modo dual ao usado por Grácio (1999). Por motivo semelhante, na linguagem natural, difere-se “poucos” de “quase nenhum”.

Apesar de quantificadores como “muitos”, “maioria” e “quase todos” terem sido estudados em vários trabalhos, os quantificadores com sentido dual a eles, como “poucos”, “minoridade” e “quase nenhum”, ainda não possuem formalização e pouco foram abordados na literatura.

Assim como Grácio (1999) adotou a noção de “muitos”, adota-se uma noção mais subjetiva para o quantificador “poucos”, em que, para o estudo deste quantificador, trata-se da noção de poucos como significando um *“conjunto pequeno de evidências”*, que não possui uma noção de cardinalidade usual, ou menor parte, mas que tenha algumas regras específicas. Por exemplo: ao afirmar que “poucos brasileiros possuem doutorado”, tem-se associada a ideia de um conjunto pequeno de evidências favoráveis à propriedade de ter doutorado entre os brasileiros. Do mesmo modo, afirmar que “poucos brasileiros pilotam aviões” refere-se a outro conjunto de brasileiros, também pequeno, ou seja, os brasileiros que pilotam aviões. É importante salientar que os conjuntos de brasileiros que possuem doutorado e de brasileiros que pilotam aviões não são, necessariamente, do mesmo tamanho.

Assim, explicitamos algumas propriedades que caracterizam nossa noção de “poucos”.

Dada uma sentença  $\varphi$ , indica-se  $[\varphi]$  o conjunto dos indivíduos do universo de discurso que satisfazem  $\varphi$ , observa-se que:

- (a) Se o conjunto de indivíduos que satisfaz  $\varphi$  também satisfaz  $\psi$ , e existem indivíduos que satisfazem  $\varphi$  (isto é,  $[\varphi]$  está contido em  $[\psi]$  e  $[\varphi]$  é não vazio), então se poucos indivíduos satisfazem  $\psi$ , então são poucos os indivíduos que satisfazem  $\varphi$ .
- (b) Se poucos indivíduos satisfazem a sentença  $\varphi$ , então existe alguém que satisfaz  $\varphi$ .
- (c) O conjunto universo não tem poucos indivíduos.

Considerando estas três propriedades para sentenças do tipo “poucos”, define-se uma estrutura matemática baseada nessas propriedades, para captar a noção inicial de poucos, que é denominada *família fechada inferiormente*. Entretanto, esta estrutura não está totalmente adequada à noção que se pretende associar ao quantificador “poucos” e, desse modo, será apresentada uma nova estrutura % *família quase fechada inferiormente* %, que mais se aproxima desta proposta.

Para a criação da lógica do poucos, introduz-se na linguagem usual da lógica clássica de primeira ordem um novo quantificador generalizado denotado por  $K$ , de maneira que:

$$Kx \varphi(x)$$

tenha o significado de “para poucos  $x$ ,  $\varphi(x)$ ”.

Deste modo, a noção de “poucos” aqui proposta é uma noção intermediária entre a noção de “existe” e “todos”. É uma noção diferente de “minoría” e “quase nenhum”, pois o conjunto que não satisfaz uma sentença sobre o escopo dos quantificadores “minoría” ou “quase nenhum” é, necessariamente, um conjunto grande, ao passo que para o quantificador “poucos”, o conjunto que não satisfaz uma proposição sobre seu escopo não é, necessariamente, um conjunto grande.

O motivo para nossa noção de “poucos” não ser dual à utilizada por Grácio (1999) é o fato de termos que, se todos os indivíduos do universo satisfazem uma sentença, então temos, pela lógica do muito, que muitos satisfazem esta mesma sentença. O dual disto seria: se nenhum indivíduo do universo satisfaz determinada sentença, então poucos satisfazem essa sentença. Consideramos isto anti-intuitivo, pois, por exemplo, imaginemos um diálogo entre dois alunos sobre as aulas do dia anterior em que o primeiro faltou. O primeiro aluno pergunta: – Quantos alunos compareceram ontem? O segundo responde: – Muitos alunos estiveram presentes. Claramente, não se conclui que estariam poucos alunos presentes naquelas aulas.

No próximo capítulo, são apresentadas algumas definições e propriedades para a formalização do conceito de poucos no ambiente lógico e que permitem a criação de uma lógica do poucos.

### 3 Uma lógica quantificacional para “poucos”

Inicialmente, apresentamos a definição de uma família fechada inferiormente sobre um universo  $A$ .

**Definição 2.1** Uma *família fechada inferiormente*  $I$  sobre um conjunto  $A$  é uma coleção de subconjuntos de  $A$  que satisfaz as seguintes condições:

- (i) Se  $B \subseteq C$  e  $C \in I$ , então  $B \in I$ ;
- (ii)  $\emptyset \in I$ .

O Conjunto das partes do Universo, denotado por  $\mathcal{P}(A)$ , isto é, o conjunto constituído por todos os subconjuntos do Universo, inclusive o próprio Universo e o vazio, é uma família fechada inferiormente sobre  $A$ , como podemos facilmente observar pela proposição seguinte.

**Proposição 2.2** O conjunto  $\mathcal{P}(A)$  é uma família fechada inferiormente sobre  $A$ .

*Demonstração:* Se um conjunto  $C \in \mathcal{P}(A)$  e  $B \subseteq C$ , então  $B \in \mathcal{P}(A)$ . Além disso,  $\emptyset \in \mathcal{P}(A)$ .

Intuitivamente, pode-se dizer que o conjunto das partes do universo é a maior família fechada inferiormente sobre  $A$ , mas ela não captura a noção de poucos pretendida neste trabalho, pois se  $\mathcal{P}(A)$  tiver poucos elementos, qualquer

conjunto, menor que este, também teria poucos elementos, o que acarretaria que a noção de “poucos” seria trivial.

Como a família fechada inferiormente não captura a noção de “poucos” esperada, apresentamos algumas propriedades para se estabelecer uma maior proximidade à intuição desejada.

**Definição 2.3** Uma família fechada inferiormente  $I$  é chamada de *imprópria* se  $I = \mathcal{P}(A)$ . Todas as outras famílias fechadas inferiormente sobre  $A$  são chamadas de *próprias*.

Abaixo, tem-se uma propriedade que todas famílias fechadas inferiormente próprias possuem.

**Proposição 2.4** Uma família fechada inferiormente  $I$  sobre  $A$  é *própria* se, e somente se,  $A \notin I$ .

*Demonstração:* Se  $A \notin I$ , então  $I \neq \mathcal{P}(A)$  e, portanto,  $I$  é uma família fechada inferiormente própria sobre  $A$ . Reciprocamente, se  $A \in I$ , então, pela definição de família fechada inferiormente,  $I = \mathcal{P}(A)$ . ■

Intuitivamente, diz-se que uma família fechada inferiormente é própria se o Universo não pertence à família.

Para se determinar a lógica do poucos, é preciso, porém, introduzir uma nova condição na família fechada inferiormente própria, pois esta ainda não se adéqua à intuição de que “se poucos indivíduos satisfazem uma sentença, então existe alguém que a satisfaz”, portanto, define-se a seguir a *Família quase fechada inferiormente*.

**Definição 2.5** Uma *Família quase fechada inferiormente* sobre um conjunto universo  $A$  é uma coleção não-vazia  $I$  de subconjuntos de  $A$  que satisfaz as seguintes condições:

- (i) Se  $B \neq \emptyset$ ,  $B \subseteq C$  e  $C \in I$ , então  $B \in I$ ;
- (ii)  $\emptyset \notin I$ ;
- (iii)  $A \notin I$ .

Para a formalização do conceito de “poucos”, utiliza-se na semântica da lógica do poucos o conceito de *família quase fechada inferiormente*.

Seja  $L$  a linguagem da lógica de primeira ordem clássica  $\mathcal{L}$  com símbolos para predicados, funções e constantes usuais e que seja fechada para os conectivos  $\wedge$ ,  $\vee$ ,  $\rightarrow$ ,  $\neg$  e para os quantificadores  $\exists$  e  $\forall$ .

Para gerarmos a lógica do poucos, que será denotada por  $\mathcal{L}(K)$ , estendemos a linguagem  $L$  da lógica clássica de primeira ordem para a linguagem  $L(K)$ , pela inclusão de um novo quantificador generalizado para “poucos”, denotado por  $K$ .

As fórmulas de  $\mathcal{L}(K)$  são as mesmas fórmulas de  $\mathcal{L}$ , mais aquelas geradas pela seguinte cláusula:

· se  $\varphi(x)$  é uma fórmula de  $\mathcal{L}(K)$ , então  $Kx \varphi(x)$  é uma fórmula de  $\mathcal{L}(K)$ .

As definições de variável livre e ligada numa fórmula são idênticas para o quantificador  $K$ , isto é, toda ocorrência de  $x$  em  $Kx \varphi(x)$  é ligada.

O resultado da substituição de todas as ocorrências de uma variável livre  $x$  em  $\varphi$  pelo termo  $t$  é denotado por  $\varphi(t/x)$ . Quando não houver problema em identificar a substituição, denota-se apenas por  $\varphi(t)$ .

Os axiomas de  $\mathcal{L}(K)$  são todos os axiomas de  $\mathcal{L}$ , incluindo os axiomas da identidade, acrescidos dos seguintes axiomas para o quantificador  $K$ :

$$(Ax1) \forall x (\varphi(x) \leftrightarrow \psi(x)) \rightarrow (Kx \varphi(x) \leftrightarrow Kx \psi(x))$$

$$(Ax2) Kx \varphi(x) \rightarrow Ky \varphi(y), \text{ quando } y \text{ é livre para } x \text{ em } \varphi(x)$$

$$(Ax3) Kx \varphi(x) \rightarrow \exists x \varphi(x)$$

$$(Ax4) Kx \varphi(x) \rightarrow \exists x \neg \varphi(x)$$

$$(Ax5) (\forall x (\varphi(x) \rightarrow \psi(x)) \wedge \exists x \varphi(x)) \rightarrow (Kx \psi(x) \rightarrow Kx \varphi(x)).$$

Os dois primeiros axiomas são necessários para a proposta de dualização da lógica do muito. Os axiomas  $Ax_3$ ,  $Ax_4$  e  $Ax_5$  são específicos para a lógica do poucos, e possuem as seguintes caracterizações intuitivas:

$Ax3$  – Se poucos indivíduos satisfazem uma sentença  $\varphi$ , então existem indivíduos que satisfazem  $\varphi$ .

$Ax4$  – Se poucos indivíduos satisfazem a sentença  $\varphi$ , então existe alguém que não satisfaz a sentença  $\varphi$ .

$Ax5$  – Se todos os indivíduos do universo que satisfazem  $\varphi$  também satisfazem  $\psi$ , e se o conjunto de indivíduos que satisfazem  $\varphi$  é não-vazio, então se poucos indivíduos satisfazem  $\psi$ , também poucos indivíduos satisfazem  $\varphi$ .

As regras de inferência de  $\mathcal{L}(K)$  são *Modus Ponens* e Generalização.

As noções sintáticas usuais, como sentença, demonstração, teorema, consequência lógica, consistência, etc., para  $\mathcal{L}(K)$ , são definidas de modo análogo às definidas na lógica clássica.

Abaixo, temos alguns teoremas de  $\mathcal{L}(K)$

**Teorema 2.6** As fórmulas abaixo são teoremas de  $\mathcal{L}(K)$ :

$$(1) \neg Kx (\varphi(x) \vee \neg \varphi(x))$$

$$(2) Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x) \rightarrow Kx \varphi(x)$$

$$(3) Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \psi(x) \rightarrow Kx \psi(x)$$

$$(4) (Kx (\varphi(x) \vee \psi(x)) \rightarrow (Kx \varphi(x) \vee Kx \psi(x)))$$

$$(5) \neg Kx (\varphi(x) \wedge \neg \varphi(x))$$

$$(6) ((Kx (\varphi(x) \vee \psi(x))) \wedge \exists x (\varphi(x) \wedge \psi(x))) \rightarrow (Kx \varphi(x) \wedge Kx \psi(x))).$$

*Demonstração:*

(1)

- |  |                          |
|--|--------------------------|
| 1. $\varphi \vee \varphi$  | Teorema de $\mathcal{L}$ |
| 2. $\forall x (\varphi(x) \vee \neg\varphi(x))$  | Gen 1                    |
| 3. $Kx (\varphi(x) \vee \neg\varphi(x)) \rightarrow \exists x \neg(\varphi(x) \vee \varphi(x))$  | Ax4                      |
| 4. $\forall x (\varphi(x) \vee \neg\varphi(x)) \rightarrow \neg Kx (\varphi(x) \vee \varphi(x))$ | CPC em 3                 |
| 5. $\neg Kx (\varphi(x) \vee \neg\varphi(x))$  | MP 2, 4                  |

(2)

- |  |                     |
|--|---------------------|
| 1. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x)$  | pp <sup>1</sup>     |
| 2. $\exists x \varphi(x)$  | CPC 1               |
| 3. $\varphi(x) \rightarrow \varphi(x) \vee \psi(x)$  | CPC                 |
| 4. $\forall x (\varphi(x) \rightarrow \varphi(x) \vee \psi(x))$  | Gen 3               |
| 5. $\forall x (\varphi(x) \rightarrow \varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x)$  | CPC 2, 4            |
| 6. $\forall x (\varphi(x) \rightarrow \varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x) \rightarrow (Kx (\varphi(x) \vee \psi(x)) \rightarrow Kx \varphi(x))$ | Ax 5                |
| 7. $Kx (\varphi(x) \vee \psi(x)) \rightarrow Kx \varphi(x)$  | MP 5, 6             |
| 8. $Kx (\varphi(x) \vee \psi(x))$  | CPC 1               |
| 9. $Kx \varphi(x)$   | MP 7, 8             |
| 10. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x) \rightarrow Kx \varphi(x)$   | TD <sup>2</sup> 1-9 |

(3)

- |   |          |
|---|----------|
| 1. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \psi(x)$  | pp       |
| 2. $\exists x \psi(x)$  | CPC 1    |
| 3. $\psi(x) \rightarrow \varphi(x) \vee \psi(x)$  | CPC      |
| 4. $\forall x (\psi(x) \rightarrow \varphi(x) \vee \psi(x))$  | Gen 3    |
| 5. $\forall x (\psi(x) \rightarrow \varphi(x) \vee \psi(x)) \wedge \exists x \psi(x)$   | CPC 2, 4 |
| 6. $\forall x (\psi(x) \rightarrow \varphi(x) \vee \psi(x)) \wedge \exists x \psi(x) \rightarrow (Kx (\varphi(x) \vee \psi(x)) \rightarrow Kx \psi(x))$ | Ax 5     |
| 7. $Kx (\varphi(x) \vee \psi(x)) \rightarrow Kx \psi(x)$  | MP 5, 6  |
| 8. $Kx (\varphi(x) \vee \psi(x))$   | CPC 1    |
| 9. $Kx \psi(x)$   | MP 7, 8  |
| 10. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \psi(x) \rightarrow Kx \psi(x)$  | TD 1-9   |

(4)

- |   |                |
|---|----------------|
| 1. $Kx (\varphi(x) \vee \psi(x))$   | pp             |
| 2. $Kx (\varphi(x) \vee \psi(x)) \rightarrow \exists x (\varphi(x) \vee \psi(x))$ | Ax 3           |
| 3. $\exists x (\varphi(x) \vee \psi(x))$  | MP 1, 2        |
| 4. $\exists x \varphi(x) \vee \exists x \psi(x)$                                  | Teorema de L 3 |
| 5. $\exists x \varphi(x)$   |                |
| 6. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x)$                     | CPC 1, 5       |

7. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x) \rightarrow Kx \varphi(x)$	Teorema L(K)
8. $Kx (\varphi(x))$	MP 6, 7
9. $Kx \varphi(x) \vee Kx \psi(x)$	CPC 8
10. $(Kx (\varphi(x) \vee \psi(x)) \rightarrow (Kx \varphi(x) \vee Kx \psi(x)))$	TD 1-9
5'. $\exists x \psi(x)$	pp
6'. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \psi(x)$	CPC 1, 5
7'. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \psi(x) \rightarrow Kx \psi(x)$	Teorema L(K)
8'. $Kx \psi(x)$	MP 6', 7'
9'. $Kx \varphi(x) \vee Kx \psi(x)$	CPC 8'
10'. $(Kx (\varphi(x) \vee \psi(x)) \rightarrow (Kx \varphi(x) \vee Kx \psi(x)))$	TD 1-4, 5'-9'

(5)

1. $\neg \exists x (\varphi(x) \wedge \neg \varphi(x))$	Teorema de L
2. $Kx (\varphi(x) \wedge \neg \varphi(x)) \rightarrow \exists x (\varphi(x) \wedge \neg \varphi(x))$	Ax3
3. $\neg \exists x (\varphi(x) \wedge \neg \varphi(x)) \rightarrow \neg Kx (\varphi(x) \wedge \neg \varphi(x))$	CPC 2
4. $\neg Kx (\varphi(x) \wedge \neg \varphi(x))$	MP 1, 3

(6)

1. $Kx (\varphi(x) \dot{\cup} \psi(x)) \wedge \exists x (\varphi(x) \wedge \psi(x))$	pp
2. $\exists x (\varphi(x) \dot{\cup} \psi(x))$	CPC 1
3. $\exists x \varphi(x) \dot{\cup} \exists x \psi(x)$	Teorema de L 2
4. $\exists x \varphi(x)$	CPC 3
5. $Kx (\varphi(x) \vee \psi(x))$	CPC 1
6. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x)$	CPC 4, 5
7. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \varphi(x) \rightarrow Kx \varphi(x)$	Teorema L(K)
8. $Kx \varphi(x)$	MP 6, 7
9. $\exists x \psi(x)$	CPC 3
10. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \psi(x)$	CPC 5, 9
11. $Kx (\varphi(x) \vee \psi(x)) \wedge \exists x \psi(x) \rightarrow Kx \psi(x)$	Teorema L(K)
12. $Kx \psi(x)$	MP 10, 11
13. $Kx \varphi(x) \wedge Kx \psi(x)$	CPC 8, 12
14. $((Kx (\varphi(x) \vee \psi(x))) \wedge \exists x (\varphi(x) \wedge \psi(x))) \rightarrow (Kx \varphi(x) \wedge Kx \psi(x)))$	TD 1-13.

**Teorema 2.7 (Teorema da Dedução):** Seja  $\Gamma \cup \{\varphi, \psi\}$  um conjunto de fórmulas de  $\mathcal{L}(K)$  e  $\Gamma \cup \{\varphi\} \vdash \psi$  uma dedução em que  $x$  ocorre livre em  $\varphi$ . Se nesta dedução de  $\psi$  a partir de  $\Gamma \cup \{\varphi\}$  a regra Gen não é aplicada a nenhuma fórmula  $\varphi_i$  que depende de  $\varphi$  para a obtenção de  $\forall x_i \varphi_i$ , então  $\Gamma \vdash \varphi \rightarrow \psi$ .

A demonstração do Teorema da dedução para a lógica do poucos é idêntica a utilizada nas lógicas moduladas.

## 2.8 Semântica para $\mathcal{L}(K)$

Do mesmo modo como em Grácio (1999), definimos a semântica para as fórmulas de  $\mathcal{L}(K)$  da seguinte maneira. Seja  $\tau = \langle I, J, K, T_0, T_1 \rangle$  um tipo de similaridade e  $\mathcal{A} = \langle A, \{R^{\mathcal{A}}_{i \in I'}\}, \{f^{\mathcal{A}}_{j \in J'}\}, \{c^{\mathcal{A}}_{k \in K'}\} \rangle$  uma estrutura clássica de primeira ordem de tipo de similaridade  $\tau$ . Uma *Estrutura de família quase fechada inferiormente* será uma estrutura de tipo  $\tau$  para  $\mathcal{L}(K)$  construída sobre  $\mathcal{A}$  com o acréscimo de uma *família quase fechada inferiormente*  $I^{\mathcal{A}}$  sobre  $A$ . Em termos formais, tem-se:

$$\mathcal{A}^I = \langle A, \{R^{\mathcal{A}}_d\}_{d \in D}, \{f^{\mathcal{A}}_j\}_{j \in J'}, \{c^{\mathcal{A}}_q\}_{q \in Q}, I^{\mathcal{A}} \rangle = \langle A, I^{\mathcal{A}} \rangle,$$

em que  $A$  é o conjunto universo, cada  $R^{\mathcal{A}}_d$ , com  $d \in D$  e aridade  $T_0(d) = m$ , é uma relação de aridade  $m$  sobre  $A$ ; cada  $f^{\mathcal{A}}_j$ ,  $j \in J$  e  $T_1(j) = n$ , é uma função de  $A^n$  em  $A$ ; cada  $c^{\mathcal{A}}_q$ , para  $q \in Q$ , é um membro de  $A$ ; e  $I^{\mathcal{A}}$  é uma *família quase fechada inferiormente* definida sobre  $A$ .

Todos os símbolos funcionais, relacionais e constantes individuais têm a mesma *interpretação* de  $\mathcal{L}(K)$  em  $\mathcal{A}$ . Define-se a relação de satisfação das fórmulas de  $\mathcal{L}(K)$ , na estrutura  $\mathcal{A}^I$ , recursivamente, da maneira usual, com o acréscimo da seguinte cláusula:

- sejam  $\varphi$  uma fórmula cujo conjunto de variáveis livres esteja contido em  $\{x\} \cup \{y_1, \dots, y_n\}$  e  $\bar{a} = (a_1, \dots, a_n)$  uma sequência de elementos de  $A$ . Então

$$\mathcal{A}^I \models Kx \varphi[\bar{a}] \text{ se e somente se } \{b \in A : \mathcal{A}^I \models \varphi[b, \bar{a}]\} \in I^{\mathcal{A}}.$$

Em particular,  $\mathcal{A}^I \models \varphi[\bar{b}]$  denota que  $\mathcal{A}^I \models_s \varphi$ , quando as variáveis livres da fórmula  $\varphi$  ocorrem no conjunto  $\{z_1, \dots, z_n\}$ ,  $s(z_i) = b_i$  e  $\bar{b} = (b_1, \dots, b_n)$ . Dado que  $A \neq \emptyset$ , então  $\mathcal{A}^I \models Kx \varphi[\bar{a}]$  se e somente se  $\mathcal{A}^I \models \varphi[\bar{a}]$  quando  $x$  não ocorre livre em  $\varphi$ . Em particular, para uma sentença  $Kx \psi(x)$ :

$$\mathcal{A}^I \models Kx \psi(x) \text{ se e somente se } \{a \in A : \mathcal{A}^I \models \psi(a)\} \in I^{\mathcal{A}}.$$

Com esta axiomática e esta semântica, conseguimos demonstrar que o cálculo de predicados é consistente e que a lógica do poucos, assim proposta, é correta e completa.

Apresentamos, agora, outras noções possíveis para o quantificador “poucos”.

## 4 Outras noções para o quantificador “poucos”

Analizamos, neste capítulo, três outras noções possíveis para o quantificador “poucos”. Embora muitas outras noções possam ser atribuídas a este quantificador, apresentamos as três que consideramos mais interessantes.

No primeiro caso, considere uma noção de poucos como estritamente oposta à noção de “muitos” utilizada por Grácio (1999), ou seja, ao contrário do quantificador proposto para “poucos” no Capítulo 2, vale que sentenças vazias satisfazem “poucos”. Assim, nesta situação, temos que se nenhum indivíduo satisfaz uma sentença, esta sentença é satisfeita por “poucos” indivíduos.

Nesta situação, para o quantificador “poucos”, não valeria o axioma (Ax3)  $Kx \varphi(x) \rightarrow \exists x \varphi(x)$ , proposto no Capítulo 2. Deste modo, nessa intuição de “poucos”, não estaria garantida a existência de um indivíduo para satisfazer qualquer sentença nos axiomas e nos teoremas. Destacamos que outras propriedades apresentadas para o quantificador “poucos” seriam mantidas, por exemplo, uma sentença que satisfaz “poucos” ainda estaria associada a conjuntos de indivíduos que apresentem um “comportamento incomum”, no sentido de que não é o comportamento usual do universo de discurso em questão.

Não se optou por essa noção de “poucos” neste trabalho por se julgar anti-intuitiva a propriedade de que se nenhum indivíduo satisfaz uma proposição, então poucos a satisfazem. Como exemplo, apresentamos a seguinte situação: se alguém disser “poucos indivíduos fizeram a avaliação”, com certeza alguém fez a avaliação.

Considere agora outra possibilidade para a noção de “poucos”, que esteja mais associada ao ambiente matemático de teoria dos conjuntos, ao atender à questão da inclusão, ou subalternação, de conjuntos: se uma sentença satisfaz muitos indivíduos, então ela também satisfaz poucos. Com base nesta noção para “poucos”, se todos os indivíduos do universo satisfazem determinada sentença, então “muitos” a satisfazem e, por conseguinte, poucos indivíduos a satisfazem.

Consideramos que essa noção para “poucos”, associada à noção de subalternação de conjuntos, deve ser entendida como “poucos ou mais”, comparando com “todos” e “alguns” do quadrado das oposições de Aristóteles. Assim, nesta situação, o quantificador “poucos” teria os seguintes aspectos:

- (a) Se todos os indivíduos do universo satisfazem determinada proposição, então poucos também a satisfazem.
- (b) Se poucos indivíduos satisfazem determinada proposição, então existe alguém que satisfaz essa proposição.

Denominando por  $K^*$  um quantificador que atenda essas condições, temos que o sistema gerado para este quantificador seria um caso particular das lógicas moduladas:

- (Ax1)  $\forall x(\varphi(x) \leftrightarrow \psi(x)) \rightarrow (K^*x(\varphi(x)) \leftrightarrow K^*x(\psi(x)))$ ;
- (Ax2)  $K^*x(\varphi(x)) \rightarrow K^*y(\varphi(y))$ , se  $y$  é livre para  $x$  em  $\varphi(x)$ ;
- (Ax3)  $K^*x(\varphi(x)) \rightarrow \exists x(\varphi(x))$ ;
- (Ax4)  $\forall x \varphi(x) \rightarrow K^*x \varphi(x)$ .

Temos, como mencionado, que se muitos indivíduos satisfazem determinada sentença, então poucos também o satisfazem. Observamos que esta noção de “poucos” gera um sistema com o mesmo conjunto de axiomas construído para o quantificador “muitos”, proposto por Grácio (1999).

O problema com esta noção de “poucos” é que não conseguimos dualizar o quantificador “muitos”, nem preservar alguns aspectos duais com o mesmo. Assim, consideramos que esta noção de “poucos” seja de uso mais complicado, caso queiramos tratar da oposição de sentenças que satisfazem “muitos” ou “poucos” indivíduos.

Para  $G$  o quantificador “muitos” de Grácio (1999) e  $K^*$  o quantificador “poucos”, e assumindo que a recíproca da contrária seja válida para um sistema lógico que contenha estes dois quantificadores, temos que:

1.  $Gx \varphi(x) \rightarrow K^*x \varphi(x)$
2.  $\neg K^*x \varphi(x) \rightarrow \neg Gx \varphi(x)$

Tem-se que (1) é verdadeira se, e somente se (2) é verdadeira.

Como um exemplo para explicitar os problemas de um sistema construído para esta segunda noção de “poucos”, considere-se as seguintes sentenças: “Muitos brasileiros vivem na região sudeste” e “Não são poucos brasileiros que vivem na região sudeste”. Estas duas sentenças parecem razoavelmente intuitivas, uma vez que ambas parecem verdadeiras, porém considerando as duas, tem-se que, se “muitos brasileiros vivem na região sudeste” então “poucos brasileiros vivem na região sudeste”, pois  $Gx \varphi(x) \rightarrow K^*x \varphi(x)$  e, por outro lado, tem-se que “Não são poucos brasileiros que vivem na região sudeste” implica que “Não são muitos brasileiros que vivem na região sudeste”, pois  $\neg K^*x \varphi(x) \rightarrow \neg Gx \varphi(x)$ . Desse modo, tem-se que “muitos brasileiros vivem na região sudeste” e “não são muitos brasileiros que vivem na região sudeste”, são sentenças verdadeiras, ao se assumir que as duas iniciais são verdadeiras, o que seria uma contradição, uma vez que “muitos” indivíduos satisfizerem uma proposição, tem-se que “muitos” também não a satisfazem.

Tentar encontrar uma situação onde “poucos” indivíduos satisfazem determinada sentença implica que “muitos” não satisfazem essa sentença, temos que se “muitos” satisfazem, então “poucos” não a satisfazem, contrariando o conceito inicial de que sempre que “muitos” indivíduos satisfazem uma determinada sentença, “poucos” também a satisfazem. Por este motivo, por não se encontrar uma oposição entre “poucos” e “muitos”, esta noção de “poucos” não foi a adotada neste trabalho.

Uma terceira noção de “poucos”, diferente das duas anteriores apresentadas neste capítulo, é uma tentativa de construir uma axiomática para o quantificador “poucos” fazendo com que a lógica para este quantificador seja uma lógica modulada, porém, quando consideramos que aconteça a oposição entre “muitos” e “poucos” teremos problemas.

Nesta situação, considere que a noção de “poucos” presente as seguintes propriedades:

- (a) Se o conjunto de indivíduos que satisfaz uma proposição  $\varphi$  satisfaz  $\psi$ , i.e.,  $[\varphi]$  está contido em  $[\psi]$ , e existem indivíduos que satisfazem  $\varphi$ , então se poucos indivíduos satisfazem  $\psi$ , poucos indivíduos satisfazem  $\varphi$ .
- (b) Se poucos indivíduos satisfazem uma proposição  $\varphi$ , então existe alguém que satisfaz  $\varphi$ .
- (c) Se todos os indivíduos satisfazem uma proposição  $\varphi$ , então poucos satisfazem a proposição  $\varphi$ .
- (d) Se muitos indivíduos satisfazem uma proposição  $\varphi$ , então não são poucos que a satisfazem.

Para diferenciar esta noção da anterior, concluímos que “muitos” não implica “poucos”, ou seja, apesar de, se todos indivíduos satisfazem uma proposição, temos que “muitos” satisfazem e “poucos” satisfazem, não se tem que se “muitos” satisfazem, então “poucos” satisfazem.

Para essa intuição de “poucos”, denominando por  $K'$  o seu quantificador e  $G$  o quantificador da lógica do muito de Grácio (1999), os axiomas seriam:

- (Ax1)  $\forall x (\varphi(x) \leftrightarrow \psi(x)) \rightarrow (K'x \varphi(x) \leftrightarrow K'x \psi(x))$ ;
- (Ax2)  $K'x \varphi(x) \rightarrow K'y \varphi(y)$ , quando  $y$  é livre para  $x$  em  $\varphi(x)$ ;
- (Ax3)  $K'x \varphi(x) \rightarrow \exists x \varphi(x)$ ;
- (Ax4)  $\forall x \varphi(x) \rightarrow K'x \varphi(x)$ ;
- (Ax5)  $(\forall x (\varphi(x) \rightarrow \psi(x)) \wedge \exists x \varphi(x)) \rightarrow (K'x \psi(x) \rightarrow K'x \varphi(x))$ .
- (Ax6)  $Gx \varphi(x) \rightarrow \neg K'x \varphi(x)$

Assim, por conta do Ax6 que afirma que “se muitos indivíduos satisfazem uma sentença, então, não são poucos que a satisfazem”, este sistema lógico também não seria uma lógica modulada.

Para esta terceira noção e formalização para “poucos”, ao reunir em um único sistema lógico, a lógica do muito e os axiomas Ax1 a Ax6 acima apresentados, por conta do axioma Ax3 “ $\forall x \varphi(x) \rightarrow Gx \varphi(x)$ ” da lógica do muito, tem-se que se todos os indivíduos satisfazem uma proposição, então “poucos” satisfazem e “poucos” não satisfazem, assim como, “muitos” satisfazem e “muitos” não satisfazem, pois a partir do axioma Ax3 da lógica do muito e do axioma (Ax6) acima apresentado, temos que  $\forall x \varphi(x) \rightarrow \neg K'x \varphi(x)$ . Mas, pelo (Ax4) acima, temos que se todos indivíduos satisfazem uma proposição, então poucos a satisfazem, e portanto, “poucos” a satisfazem e “poucos” não a satisfazem. Para obter que “muitos” satisfazem e “muitos” não satisfazem, basta considerar a recíproca da contrária do axioma (Ax6) apresentada para esse conceito de “poucos”.

Portanto, se for considerada a segunda ou a terceira noção e consequente formalização de “poucos”, haveria problemas, como “muitos” e “muitos’ não”, e

“poucos” e “poucos’ não”. Ainda, além de não se ter o dual do quantificador “muitos”, se obteria um sistema considerado menos intuitivo que o proposto no Capítulo 2. Quanto à primeira noção, em que o vazio possui “poucos” elementos, consideramos que ela pode ser algo a ser estudado posteriormente, quando a preocupação com a intuição desse quantificador não estiver no contexto.

## Considerações Finais

Neste trabalho, apresentamos um sistema lógico para o quantificador “poucos” que possui aspectos duais ao quantificador “muitos” de Grácio (1999), além disto, apresentamos algumas outras noções possíveis para o quantificador “poucos” e mostramos como algumas delas seriam problemáticas se o intuito do estudo for dualizar o quantificador “muitos”.

Acreditamos que o estudo de quantificadores da linguagem natural, assim como uma formalização lógica para os mesmos seja de grande interesse tanto para a Linguística, para a Lógica, quanto para a Computação e Inteligência Artificial (I.A.), nesta última, pois, ao implementarmos uma linguagem em um sistema dotado de I.A. temos de dar o significado mais aproximado possível daquilo que queremos que o sistema entenda.

As análises realizadas neste trabalho acerca do quantificador “poucos” são importantes a várias áreas e, do mesmo modo como a Tese de Grácio (1999) e os trabalhos de Sette, Carnielle e Veloso (1999), traz algumas elucidações sobre noções que temos na linguagem natural através de uma linguagem artificial.

## Referências

- BARWISE, J.; COOPER, R. Generalized quantifiers and natural language. *Linguistics and Philosophy*, Providence, v. 4, p. 159-219, 1981.
- CARNIELLI, W.A.; VELOSO, P.A.S. Ultrafilter logic and generic reasoning. In: KURT GÖDEL COLLOQUIUM, 5, 1997, Berlin. *Proceedings...* Berlin: Springer-Verlag, 1997. p.34-53.
- GRÁCIO, M. C. C. *Lógicas moduladas e raciocínio sob incerteza*. 1999. 194 f. Tese de doutorado (Doutorado em Lógica e Filosofia da Ciência)-Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 1999.
- MOSTOWSKI, A. On a generalization of quantifiers. *Fund. Mathematical*, Warszawa, v.44, p.12-36, 1957.
- PETERSON, P. L. On the logic of “few”, “many”, and “most”. *Notre Dame Journal Logic*, Notre Dame, v.20, p.155-79, 1979.
- SANZ, M. J. F. (Org.). *Filosofía de La Lógica*. Madri: Editora Tecnos, 2007.

SETTE, A. M.; CARNIELLI, W. A.; VELOSO, P. An alternative view of default reasoning and its logic. In: HAUESLER, E. H.; PEREIRA, L. C. (Ed.) *Pratica: Proofs, types and categories*. Rio de Janeiro: PUC, 1999. p. 127-58.

## Notas

<sup>1</sup> Premissa Provisória

<sup>2</sup> Teorema da Dedução (Teorema 2.7)

# **SOBRE A NOÇÃO DE “ESTADOS DE COISAS CONTRADITÓRIOS”: UMA APROXIMAÇÃO DE NOÇÕES LÓGICAS A NOÇÕES MÉTRICO-TOPOLÓGICAS.**

WALTER GOMIDE

(Universidade Federal de Mato Grosso – UFMT)

waltergomide@yahoo.com

Resumo: Neste trabalho, procuro mostrar como estados de coisas relacionados com a atribuição de pares de propriedades contraditórias a um objeto pode ser associado ao conceito de “ponto fronteiro” em um espaço métrico. Para tanto, utiliza-se a noção formulada por Graham Priest de “dialetheia” (contradições que efetivamente acontecem no mundo real) a fim de mostrar como as contradições associadas ao movimento podem ser vistas como situações fronteiriças em um espaço métrico. Também neste artigo é apresentada uma análise da “hipótese do espalhamento de Priest”, a partir de noções métrico-topológicas. O artigo se encerra com uma apresentação intuitiva da tese de que qualquer propriedade monádica pode ser compreendida a partir de correspondências com vizinhanças em espaços métricos; nestes contextos, as contradições engendram situações claramente relacionadas com “pontos de fronteira”.

Palavras-chave: espaço métrico – contradição – dialetéia – vizinhança – propriedade.

## **1. O Princípio de Não-Contradição e a Paraconsistência.**

Talvez um dos princípios mais diretivos do pensamento intelectual ocidental seja o de *não-contradição*, elaborado inicialmente por Aristóteles da forma seguinte:

É impossível para uma mesma coisa, ao mesmo tempo, pertencer e não pertencer à mesma coisa e sob o mesmo aspecto (*Metafísica*, G3, 1005b 18-23).

Contemporaneamente, muitas podem ser as traduções formais deste princípio<sup>1</sup>. Entretanto, quaisquer que sejam estas, a síntese do que é por ele enunciado é o seguinte: uma propriedade qualquer não pode, simultaneamente e sob o mesmo aspecto, ser afirmada e negada de um objeto. Neste sentido, a lei ou princípio da não-contradição (PNC) tem um caráter regulador daquilo que se pode esperar no campo ontológico: é interdita ao mundo das coisas a existência de um objeto que, simultaneamente e sob o mesmo aspecto, possua propriedades contraditórias. Mas também se depreende do PNC um elemento lógico, isto é, de PNC surgem interdições semântico-sintáticas: sob o ponto de vista semântico, o PNC interdita que uma proposição possa ser afirmada e negada ao mesmo tempo,

isto é, a conjunção ( $A \wedge \neg A$ ) é sempre falsa na lógica clássica ou aristotélica; em sua dimensão sintática, PNC proíbe que um cálculo lógico qualquer deduza  $A$  e  $\neg A$ . Segundo Francesco Berto, o PNC é um princípio do pensamento (dimensão lógica) e do ser (dimensão ontológica) que, na perspectiva aristotélica, é visto como “o mais certo de todos”:

Aristóteles o chama de [...] “o princípio mais certo de todos” – *firmissimum omnium principiorum*, diziam os medievais. É o Princípio de não-Contradição [...] A qualidade de *firmissimum* exprime o fato de que [o princípio de não-contradição] é considerado a *lei* mais certa e sem controvérsias do pensamento e do ser, e portanto é posto como fundamento supremo do conhecimento e da ciência (BERTO, p.21, [2006])<sup>2</sup>.

Sob a perspectiva semântica, o PNC não permite que a conjunção ( $A \wedge \neg A$ ) seja verdadeira. Mais especificamente, segundo a concepção de lógica clássico-aristotélica, a fórmula molecular ( $A \wedge \neg A$ ) é falsa em qualquer interpretação, ou, dito de outra maneira, é *insatisfatível*: não existe um modelo para a supracitada conjunção. Sintaticamente considerado, a violação de PNC acarreta, classicamente falando, a trivialização de teorias.

O conceito de um cálculo (ou teoria) trivial é explicitamente formulado por Decio Krause no trecho seguinte, retirado da apresentação da obra *Sistemas formais Inconsistentes*, inicialmente publicada em 1964, considerada o marco zero da lógica paraconsistente – o *antídoto* contra os efeitos daninhos que a contradição gera no âmbito sintático:

A teoria  $[\Gamma]$  é dita *trivial* se o conjunto de suas fórmulas coincide com os de seus teoremas, ou seja, dito informalmente, se todos os enunciados sintaticamente corretos do ponto de vista da linguagem de  $[\Gamma]$  puderem ser provados em  $[\Gamma]$ ; se este for o caso, a teoria não permite que se distinga o “demonstrável” do “não-demonstrável”, não apresentando, aparentemente, interesse algum, uma vez que não se poderá, por assim dizer, separar o verdadeiro do falso (KRAUSE.D, ix, in: DA COSTA, 1993).

Dentro da lógica clássica, a presença de enunciados contraditórios e demonstráveis acarreta, por força de regras ou esquemas inferenciais tipicamente inspirados em um uso “normal” da negação lógica, que qualquer enunciado sintaticamente correto da linguagem é teorema da teoria em questão; dentro da lógica clássica, a contradição elimina que o verdadeiro seja identificado com o que é sintaticamente provável. De fato, esta situação lógica indesejável é expurgada dos sistemas lógicos paraconsistentes, iniciados pelo lógico brasileiro Newton C. A. Da Costa: nos sistemas criados por Da Costa<sup>3</sup>, os procedimentos inferenciais são escolhidos de tal forma que a possibilidade de explosão demonstrativa, a trivialização, seja interdita na presença de enunciados que se contradizem. A íntima relação entre a presença de enunciados contraditórios como teoremas (o que qualifica uma teoria como “inconsistente – ver KRAUSE, in: DA COSTA, ix, *Op. Cit*) e a trivialização é aludida por KRAUSE, na mesma apresentação dos *Sistemas Formais Inconsistentes*:

Inconsistências e trivialidade, à luz das lógicas de cunho clássico (e da grande maioria dos sistemas lógicos em geral), são na realidade conceitos inseparáveis, e a normalidade com que isso tem sido aceito se deve, talvez, ao fato de que um requisito mínimo sobre um sistema de crenças, para que ele possa ser definido como “racional”, é que seja consistente [...] Com efeito, o *princípio da não-contradição*, excluindo a possibilidade de que uma proposição e sua negação sejam ambas verdadeiras, é um dos pressupostos básicos da lógica tradicional, tendo em vista que a presença de uma contradição nas teorias  $[\Gamma]$  em geral tornam-nas triviais (KRAUSE, in: DA COSTA, x, 1993).

Não há dúvida de que, sob o ponto de vista filosófico, as questões levantadas pela paraconsistência são de extremo interesse. Em uma primeira análise, o primeiro respeito à natureza daquilo que é a “negação”. Quais as propriedades lógicas que um operador deve satisfazer para ser considerado uma autêntica “negação”? Seria a negação paraconsistente uma verdadeira negação, ou outra coisa qualquer próxima da autêntica “negação”, esta perfeitamente definida nos moldes da lógica clássica? Questões desta natureza, por si sós, já são suficientemente ricas para legitimar as lógicas paraconsistentes como objetos de instigantes discussões filosóficas<sup>4</sup>.

Mas há uma outra questão que de imediato se coloca na origem, por assim dizer, de um dito “pensamento paraconsistente”: há de fato no mundo contradições? O princípio de não-contradição, tal como formulado por Aristóteles, com toda a sua força reguladora do pensamento, encontraria objetos ou propriedades que não o satisfizessem? Sem dúvida alguma, o problema sintático que a paraconsistência resolve (o da trivialização) se expande e ganha uma dimensão semântica, isto é, surge o problema de conceber o mundo como de fato sendo um portador de autênticas contradições, ou então como o lugar de contradições que não seriam reais, mas apenas aparentes. Dito de outra maneira, o que se levanta aqui é o postulado de que o enunciado  $(A \wedge \neg A)$  possa realmente encontrar uma exemplificação no mundo, postulado este que se antagoniza com a visão de que as contradições se dariam apenas de forma aparente (obviamente, o pensamento clássico postularia que a contradição supracitada não se dá nem de forma real e nem de forma aparente).

## **2. O “Dialeteísmo” de Graham Priest e as Mudanças do Tipo $\Delta$ .**

Segundo o filósofo inglês Graham Priest, há de fato no mundo situações em que a contradição se faz presente. Tal concepção filosófica é conhecida por “dialeteísmo”. Em linhas gerais, os dialeteístas afirmam que, em determinados contextos, certas propriedades, por sua própria natureza, não podem ser atribuídas aos objetos de uma maneira totalmente definida e, desta forma, a própria indefinição inerente a estas propriedades, em tais contextos, gera casos em que não se pode contar com o princípio de não contradição, este sendo

formulado como a interdição de que uma proposição e sua contraditória possam simultaneamente ser verdadeiras. Segundo Priest:

Dialeteísmo é simplesmente a visão de que algumas contradições são verdadeiras. Isto é, há algumas sentenças (enunciados, proposições, ou qualquer coisa que tenha valores de verdade),  $\alpha$ , tais que ambas  $\alpha$  e  $\neg\alpha$  são verdadeiras. [...] Pode-se perguntar se “segue disto que há contradições reais no mundo?”. Em um sentido nada problemático, [pode-se dizer que sim] (PRIEST, [2006], p.299)<sup>5</sup>.

Como exemplo de tais proposições, tomemos aquelas que estão relacionadas ao movimento e à localização no tempo. Como ilustração de contradições que surgem a partir do movimento, Priest apresenta o exemplo de um sistema  $s$  que se movimenta, através do tempo, percorrendo vários estados físicos:

Antes de um tempo  $t_0$ , um sistema  $s$  está no estado  $s_0$ , descrito por  $\alpha$ . Depois de  $t_0$  ele está no estado  $s_1$ , descrito por  $\neg\alpha$ . Em qual estado ele está em  $t_0$ ? *A priori*, há quatro possibilidades:

- (A)  $s$  está em  $s_0$  somente.
- (B)  $s$  está em  $s_1$  somente.
- (G)  $s$  não está nem em  $s_0$ , e nem em  $s_1$
- (D)  $s$  está em  $s_0$  e em  $s_1$

Naturalmente, não deve haver uma resposta uniforme. Mudanças diferentes podem ser mudanças de vários tipos. A questão crucial que eu gostaria de propor é se há mudanças do tipo  $\Delta$ , isto é, mudanças dialéticas [...] Argumentarei que há algumas mudanças que são do tipo  $\Delta$  (PRIEST, [2006], p.160)<sup>6</sup>

Como exemplo de uma *real* mudança dialética, Priest apresenta o caso seguinte:

Eu estou em uma sala. Enquanto eu caminho através da porta, eu estou dentro ou fora da sala? Para enfatizar que isto não é um problema de vagueza, suponha que identifique minha posição com o meu centro de gravidade, e a porta com o plano vertical que passa pelo centro. Enquanto eu deixo a sala, deve haver um instante para o qual o ponto se situa no plano. Em relação a este instante, estou dentro ou fora [da sala]? Claramente, não há razão para afirmar um caso em detrimento do outro (PRIEST, *op. cit.*, p.161)<sup>7</sup>.

Uma sala quarto, na qualidade de uma região espacialmente determinada, tem uma parte *interna* e outra *externa*. Entretanto, isto não significa que tais partes sejam perfeitamente determinadas em toda a extensão do quarto; há os *pontos fronteiros* para os quais não se pode afirmar se um corpo está dentro ou fora do quarto; antes de ser um caso de vagueza, tal situação reflete uma propriedade métrica dos pontos que constituem a região espacial em questão.

Para esclarecer tal situação, introduzamos o conceito de espaço métrico<sup>8</sup>. Seja  $X$  qualquer conjunto não vazio. Seja  $d$  uma função de  $X^2$  em  $\mathbf{R}^+$  (o conjunto

dos números reais maiores ou iguais a zero).  $d$  é chamada uma métrica sobre  $X$  se satisfizer as seguintes condições:

- i)  $d(x,y) \geq 0$ , para todo  $x, y \in X$ ;
- ii)  $d(x,y) = d(y,x)$ , para todo  $x, y \in X$ ;
- iii)  $d(x,y) = 0$  se, e somente se,  $x = y$ ;
- iv)  $d(x,y) + d(y,z) \geq d(x,z)$ , para todo  $x, y$  e  $z \in X$ .

O conjunto  $X$  em que se define a função métrica  $d$ , satisfazendo as condições acima dadas, é chamado de *espaço métrico* e pode ser denotado como  $\langle X, d \rangle$ .

Seja agora  $\langle X, d \rangle$  um espaço métrico. Consideremos todos os  $y$  de  $X$  que satisfazem a seguinte inequação:

$$v) \quad d(p, y) < \varepsilon,$$

sendo  $p$  um elemento de  $X$  e  $\varepsilon \in \mathbf{R}^+$ . Assim sendo, denominamos de  $d_\varepsilon$  - vizinhança de  $p$  (em símbolos,  $d_\varepsilon$  - Viz  $p$ ) o conjunto:

$$vi) \quad \{y \in X / d(p, y) < \varepsilon\}.$$

Este conjunto contém os elementos de  $X$  que estão dentro de  $d_\varepsilon$  Viz  $p$  - o conjunto dos elementos de  $X$  que distam de  $p$  algo *menor* que  $\varepsilon$ ; o conjunto dos  $y$  definido pela relação:

$$vii) \quad d(p, y) > \varepsilon$$

claramente definem os  $y \in X$  que estão fora de  $d_\varepsilon$  - Viz  $p$ . Mas o que dizer dos  $y \in X$  que satisfazem a relação:

$$viii) \quad d(p, y) = \varepsilon ?$$

Por *vi*), tais  $y$  estão fora da vizinhança de  $p$  determinada por  $\varepsilon$  e, por conseguinte, pertencem ao complementar de tal vizinhança. Neste caso, *vii*) tem de ser substituído por

$$ix) \quad d(p, y) \geq \varepsilon.$$

Mas não há nenhuma razão intuitiva para que *ix*) seja aceito em detrimento de, por exemplo, *viii*), uma nova definição de  $d_\varepsilon$  - Viz  $p$ :

$$vi^*) \quad d_\varepsilon - Viz p \equiv \{y \in X / d(p, y) \leq \varepsilon\}$$

Por *ix*) e *vi*\*)), os elementos de  $X$  que satisfazem exatamente a equação *viii*) são aqueles que, como o centro de gravidade de uma pessoa que atravessa a porta de uma sala, encontram-se no tanto dentro como fora de uma vizinhança definida em um espaço métrico; tais pontos que satisfazem a identidade expressa por *vii*) são chamados de *pontos de fronteira de X, em relação a  $d_\varepsilon$  Viz  $p$* .

O exemplo de uma pessoa se movimentando dentro de uma sala é expressamente *modelado* por Graham Priest nos seguintes termos (ver PRIEST, *op. cit.*, p. 163-164): tomemos a tripla

$$x) \langle W, <, v \rangle,$$

em que  $W$  é o conjunto de instantes temporais (tomados com a mesma topologia da reta real), “ $<$ ” é a relação de ordem, definida de forma usual, e  $v$  é uma função de parâmetros proposicionais, indexados com instantes  $x$ , no conjunto  $\pi$ , definido como:

$$xi) \pi = \{\{0\}, \{1\}, \{0,1\}\}.$$

A partir destas considerações semânticas, Priest caracteriza que toda mudança do já citado tipo  $\Delta$  (aquelas que, como a transição de dentro para fora de uma sala – ou *vice-versa* – pressupõem a passagem ou a localização em *pontos fronteiros* de conjuntos) satisfazem as seguintes condições:

$$\begin{aligned} v_x(p) &= \{1\}, \text{ se } x < t_0 \\ v_x(p) &= \{0\}, \text{ se } x > t_0 \\ v_x(p) &= \{0,1\}, \text{ se } x = t_0. \end{aligned}$$

Nas equações acima,  $p$  é um enunciado do tipo: “ $a$  está em  $R$ ”. Pelas valorações definidas para os conectivos proposicionais<sup>10</sup>, chega-se à conclusão de que, em  $t_0$ ,

$$1 \in v_{t_0}(p \wedge \neg p).$$

Assim, a contradição “ $a$  está em  $R$  em  $t_0$  e  $a$  não está em  $R$  em  $t_0$ ” é verdadeira – lembremos que Priest define a relação semântica “ $1 \in v(\alpha)$ ” como “ $\alpha$  é verdadeira sob  $v$ ”.

A contradição acima citada, típica dos casos em que um objeto  $a$  se encontra na passagem de uma região a outra (ou se encontra nos pontos de fronteira de tal região), também poderia ser a tradução semântica da seguinte equação:

$$x) d(p, {}^*a_0) = r.$$

$x)$  nos diz que, no instante  $t_0$ , a posição do objeto  $a$  (indicada por “ ${}^*a_0$ ”), dista da posição  $p$  uma distância  $r$ – o raio que determina o escopo da vizinhança  $d_r$ -Viz  $p$ . Neste caso, as contradições que Priest verifica nas mudanças do tipo  $\Delta$  podem ser compreendidas como representativas de situações métricas que envolvem vizinhanças.

Mais não só isto: para Priest, a força motriz do movimento é a contradição. Estar em movimento é justamente estar na vivência plena de estados contraditórios. Segundo Priest, seguindo a tradição hegeliana, o movimento é causado pela contradição, tese esta explicitamente posta pela seguinte citação de Hegel que Priest insere em seu texto:

[A] contradição é a raiz de todo movimento e vitalidade; e é somente na medida em que algo contém uma contradição dentro de si que ela se move [...](HEGEL in: PRIEST, *op.cit.*, p.170)<sup>11</sup>.

Mas a emergência do movimento a partir da contradição não impede uma análise inversa: a contradição pode ser compreendida como causada pelo movimento; e não qualquer movimento: um movimento que se dê de forma contínua, dentro de intervalos temporais (estes compostos de pontos geométricos – os instantes – dispostos em espaços métricos) que se comportam como vizinhanças bem determinadas. De fato, a relação entre contradição e vizinhanças em espaços métricos é indicada por Priest na sua análise do *princípio de continuidade* (ou, como Priest denomina, “*The Leibniz Condition of Continuity*”) e na formulação de sua *Hipótese do espalhamento* (“*The Spread Hypothesis*”), como veremos a seguir.

### **3. A Condição de Leibniz de Continuidade e a Hipótese do Espalhamento.**

Para Priest, uma das razões que fazem da contradição um elemento intrínseco ao movimento é o fato de que os movimentos (ou as transformações, de quaisquer espécies) se dão em conformidade com o princípio de continuidade. A fim de formular claramente tal princípio, Priest toma como referência a concepção de Leibniz sobre continuidade, a saber:

Quando a diferença entre duas grandezas em uma dada série [...] pode ser diminuída até se tornar menor do que qualquer quantidade dada, a correspondente diferença é [...] necessariamente diminuída e menor do que qualquer grandeza dada (LEIBNIZ in: PRIEST, *op. cit.*, p.165)<sup>12</sup>

A partir da formulação de Leibniz, Priest estabelece uma definição, por assim dizer, *contemporânea* de continuidade – denominada por Priest de “*The Leibniz Condition of Continuity*” (LCC). Para tanto, é explícito o uso de noções métricas e do conceito de limite<sup>13</sup>:

[...] Com cerca de trezentos anos de retrospecto matemático, é fácil pensar que o que Leibniz está dizendo é que, para duas seqüências matemáticas,  $(s_n)$  e  $(t_n)$ , se  $\lim_{n \rightarrow \infty} s_n - t_n = 0$ , então  $\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} t_n$  (PRIEST, *op. cit.*, p.165)<sup>14</sup>

Uma vez sendo enunciada “a condição de Leibniz” (daqui por diante, como o faz Priest, abreviada por LCC), torna-se uma tarefa interessante a apresentação de uma interpretação semântica para LCC, isto é, uma versão de LCC que envolva valores de verdade, parâmetros proposicionais e valorações – funções de tais parâmetros sobre valores de verdade. Para tanto, Priest faz uso da já citada

estrutura  $\langle W, <, v \rangle$ , em que  $W$  é uma coleção de instantes, com a mesma topologia da reta real (e, portanto, toma-se  $W$  como um *espaço métrico euclidiano*), “<” a relação de ordem usual, e  $v$ , uma função de parâmetros proposicionais (indexados com os instantes de tempo) sobre o já citado conjunto  $\pi = \{\{0\}, \{1\}, \{0,1\}\}$ . Surge então a versão semântica de LCC:

Para todo parâmetro proposicional  $p$ , e para todo  $x, y \in W$ , se  $1[0] \in v_z(p)$  para todo  $z$ , tal que  $x < z < y$ , então  $1[0] \in v_x(p)$  e  $1[0] \in v_y(p)$ <sup>15</sup> (PRIEST, *op. cit.*, p. 168).

Através de LCC, como enunciado acima, Priest apresenta a tese de que o movimento, compreendido como uma função contínua de instantes de tempo sobre posições no espaço, é um fator gerador de contradição. Tomemos o caso em que um corpo se movimenta conforme a função  $x = kt$ , em que  $x$  é a sua posição e  $t$  é o tempo. Em um  $t \neq t_0$ , temos que  $x \neq kt_0$ . Em qualquer intervalo de tempo  $\{t/ t_1 < t < t_0\}$ , o corpo em questão, em todos os termos intermediários do intervalo supracitado, terá a propriedade posicional  $x \neq kt_0$ . Entretanto, por LCC, esta propriedade se encontra nos extremos do intervalo. Logo, em  $t_0$ , o móvel tem simultaneamente as propriedades posicionais  $x = kt_0$  e  $x \neq kt_0$ , uma nítida contradição instantânea. Mais ainda: além de ser uma contradição, tal situação gera um claro caráter indeterminado quanto à posição exata que o móvel possui em  $t_0$ .

Tais indeterminações posicionais (inerentes a qualquer deslocamento que se faça de forma contínua) leva Priest a postular a *Hipótese do Espalhamento* (“*Spread Hypothesis*”):

Um corpo não pode ser localizado em um ponto que ele está ocupando em um instante de tempo, mas somente naqueles pontos que ele ocupa em uma pequena vizinhança daquele [instante de tempo] (PRIEST, *op. cit.*, p. 177).<sup>16</sup>

Para Priest, a hipótese do espalhamento “foi proposta inicialmente no contexto da análise do movimento: sua aplicação soluciona [...] aspectos intrigantes sobre tal tema”. (ver PRIEST, *op. cit.*, p. 214).

No contexto da análise que Priest faz do movimento, a hipótese do espalhamento é um elemento que não tem caráter epistêmico, mas pretende ser uma apresentação de como as coisas de fato ocorrem na natureza: não é possível na natureza isolar as coisas a ponto de precisar com rigor onde elas estão (PRIEST, *op. cit.*, p. 214).

Propõe-se agora aqui uma análise da hipótese do espalhamento, a partir de LCC e de considerações métricas. Primeiramente, consideremos uma estrutura formada pelos seguintes termos:

$$\langle W, <, v, d, \Theta \rangle,$$

em que  $W$  é um conjunto de instantes, “<” uma relação de ordem usual,  $v$  são valorações indexadas por instantes de  $W$  sobre parâmetros proposicionais,  $d$  uma

função métrica definida em  $W^2$  sobre  $\mathbf{R}^+$  (os números reais positivos), e  $\Theta$  um conjunto de vizinhanças definidas para números reais positivos e, tais que e  $\varepsilon \geq 0$  e existe um  $(x,y) \in W^2$ , tal que  $d(x,y) = \varepsilon$ .

Tomemos agora outra estrutura

$$\langle W', <, v', d, \Theta' \rangle,$$

tal que  $W'$  é um conjunto de posições, “<” uma relação de ordem usualmente definida,  $v'$  são valorações indexadas por posições de  $W'$ ,  $d$  uma função métrica definida em  $W'^2$  sobre  $\mathbf{R}^+$ , e  $\Theta'$  um conjunto de vizinhanças definidas para números reais  $\sigma \geq 0$ , e existe um  $(x',y') \in W'^2$ , tal que  $d(x',y') = \sigma$ . Postulemos também uma função contínua  $f'$ , tal que:

a)  $f(W) = W'$ ,

Desta forma, para todo  $x' \in W'$ , temos  $x' = f(x)$ , para algum  $x \in W$ , e, por conseguinte:

b)  $d(x', y') = d(f(x), f(y))$ .

Consideremos agora o intervalo de  $W$  igual a:

$$T = \{t / a < t < b\}.$$

A este intervalo de tempo corresponderá, em  $W'$ , o intervalo:

$$X = \{x / f(a) = a' < f(t) = x < f(b) = b'\}.$$

Consideremos agora um corpo em movimento contínuo durante o intervalo  $T$ . Neste intervalo, para cada instante  $t$ , associa-se ao corpo uma posição  $f(t)$ . Em  $T$ , determinemos o conjunto  $\Theta$  de todas as vizinhanças definidas em  $T^2$ . Por sua vez, este conjunto definido em  $T^2$ , determina, através da função  $f$ , o conjunto  $\Theta'$  de todas as vizinhanças definidas em  $X^2$ .

Em seu movimento contínuo, o referido corpo vai da posição  $f(a)$  até a posição  $f(c)$ , sendo  $c > a$  e  $c < b$ . Consideremos agora um posição  $x_0$ , de tal forma que

$$f(a) = x_0 - \sigma < x_0 < f(b) = x_0 + \sigma$$

Tomemos agora um conjunto  $Y = \{y / 0 < y < \sigma\}$ . Tomemos agora todas as vizinhanças  $\Theta'$  de  $x_0$  cujos raios percorrem os valores de  $Y$ . Consideremos o valor  $\beta$ , tal que  $f(c) = x_0 - \beta$ . Seja agora o conjunto  $Y' = \{y' / 0 < y' < \beta\}$ . Para qualquer valor de  $y' \in Y'$ , temos que um corpo que se encontra na vizinhança  $d_{y'}$ . - Viz  $x_0$ , também se encontra na vizinhança  $d_{\beta}$  - Viz  $x_0$ : a propriedade de estar na vizinhança  $d_{\beta}$  - Viz  $x_0$  vale para todos os  $y'$  de  $Y'$  e, portanto, por LCC, vale também em  $y' = 0$ , isto é, vale para  $x_0$ . Isto implica que um corpo em movimento contínuo<sup>18</sup>, ao passar por uma posição  $x_0$  qualquer, com certeza se encontra em qualquer

vizinhança infinitesimal deste ponto; e isto nos assegurado por LCC. Mas, e quanto à posição específica  $x_o$ , o que dizer quanto à localização de tal corpo?

Consideremos que o móvel em questão se movimentava dentro das vizinhanças que satisfazem os valores pertencentes a  $Y'$ . Quando o móvel estiver em  $x_o - \beta'$ , com  $\beta' < \beta$ , então ele vai estar situado em um ponto fronteiro de  $d_{-\beta}$ . *Viz*  $x_o$ , a saber o ponto  $x_{\beta'}$ , para o qual vale a relação  $d(x_o - \beta', x_o) = \beta'$ . Na qualidade de fronteiro, para tal ponto vale, para o corpo em deslocamento – denominado de  $\mathbf{b}$  –, a propriedade:

$$(\mathbf{b} \text{ está em } d_{\beta'} \text{ - Viz } x_o \wedge \mathbf{b} \text{ não está em } d_{\beta'} \text{ - Viz } x_o).$$

A propriedade acima afirma que, quando  $\mathbf{b}$  estiver a uma distância  $\beta'$  de  $x_o$ , ele se encontrará simultaneamente *dentro* e *fora* desta vizinhança de raio  $\beta'$  – *imageticamente*, podemos conceber a situação em estrita analogia com o caso já abordado de um corpo que se encontra em movimento, passando pela porta de uma sala; no caso aqui tratado, uma sala de *dimensões infinitesimais*.

Para qualquer  $y' \in Y'$ , vale a propriedade que existe uma vizinhança  $V^o$  de  $x_o$ , de raio  $y'$ , tal que o corpo está em  $V$  e não está em  $V$ , quando dista um valor  $y'$  de  $x_o$ , isto é, a propriedade

$$(\exists V^o_{\epsilon \Theta'}) (\mathbf{b} \text{ está em } V^o \wedge \mathbf{b} \text{ não está em } V^o)$$

é verdadeira para todos os valores de  $Y'$  e, por força de LCC, também é verdadeira no extremo inferior  $y' = 0$ . Como  $y' = 0$  equivale ao ponto em que o corpo está em  $x_o$  e, por sua vez,  $x_o = f(t_o)$ , para algum  $t_o \in T$ , em  $t_o$  não podemos determinar exatamente onde se encontra o móvel, pois, por força da expressão imediatamente acima, este móvel tanto está na vizinhança  $\{x_o\}$ , quanto não está em  $\{x_o\}$  – lembremo-nos de que a fórmula existencial acima, no caso de  $y = 0$ , só pode ser satisfeita pela vizinhança determinada por

$$d(x_o, y) < 0 \text{ ou } d(x, y) = 0.$$

Excluindo o caso em que os valores da função distância são negativos, resta o valor  $d(x, y) = 0$ . Tal valor para a função só é satisfeito pelo par ordenado  $(x_o, x_o)$ . Logo, a *única* vizinhança que instancia a supracitada fórmula existencial é a seguinte:

$$V^o_{\epsilon \Theta'} = \{y / d(x_o, y) = 0\} = \{x_o\}.$$

Desta maneira, mediante uma análise de vizinhanças em espaços métricos e com o auxílio de LCC, embora não se possa *efetivamente demonstrar a hipótese do espalhamento*, pode-se mostrar a sua plausibilidade. De fato, um corpo que esteja em movimento, em um instante  $t_o$ , encontra-se em uma vizinhança infinitesimal da posição  $x_o$ , sendo que  $x_o = f(t_o)$ , e  $f$  é uma função contínua. Neste caso, a localização do corpo, em  $x_o$ , se faz a contento dentro desta vizinhança, se por “localizar” entende-se uma indeterminação  $\Delta x$  tão pequena quanto queiramos. Mas se o ensejo é determinar onde está precisamente o móvel em  $t_o$  – o que,

topologicamente, equivale a determinar se o corpo está ou não está, de forma exclusiva, na vizinhança  $\{x_0 = f(t_0)\}$  -, então tem-se o malogro da determinação: a continuidade do movimento faz com que  $x_0$  *herde*, por intermédio da condição leibniziana de continuidade (LCC), a propriedade de ser um *ponto fronteiro* da vizinhança  $\{x_0\}$ .

#### 4. Conclusão: os estados de coisas contraditórios e os “Pontos Fronteiros” de uma propriedade

De fato, percebe-se que os exemplos aludidos pelos dialeteístas como contraditórios e intuitivos - e associados ao movimento - são aqueles que, de alguma forma, estão relacionados com propriedades ou relações que admitem *pontos fronteiros*: as situações intuitivamente contraditórias são aquelas em que objetos, ao mesmo tempo, estão *dentro* e *fora* de uma *região conceitual*; como se tais objetos estivessem na *fronteira* entre uma propriedade, entendida como análoga a uma região geométrica, e sua contraditória. No caso do corpo em movimento que passa por um ponto, devido à própria natureza não-extensa do ponto, tal analogia com as fronteiras de regiões geométricas não se dá de forma tão óbvia, mas, no exemplo do homem que se adentra em um recinto, a analogia é evidente: há uma situação, aquela em que ele está passando pela porta, que indica o estado de coisas de tal homem estar dentro e fora do recinto ao mesmo tempo, justamente porque ele se situa em uma região de fronteira.

De maneira geral, podemos postular que os estados de coisas contraditórios, em grandes linhas, são redutíveis às situações nas quais se pode falar, em estrita analogia com situações geométricas, de pontos fronteiros de uma propriedade. As propriedades ou relações que admitiriam ser tratadas como possuindo um *dentro* e um *fora* e, por conseguinte, um limite entre tais regiões interna e externa (os ditos pontos de fronteira) seriam aquelas para as quais o princípio de não contradição, sob a forma do enunciado molecular

$$R^n(a_1, a_2, \dots) \wedge \neg R^n(a_1, a_2, \dots),$$

não valeria de forma absoluta; na realidade, afirmar a conjunção acima para uma certa seqüência de objetos  $(a_1, a_2, \dots)$  seria o mesmo que afirmar que tal seqüência está no ponto de fronteira de  $R^n(x_1, x_2, \dots)$ .

Mas como se pode falar em “pontos de fronteira” de propriedades, entidades intensionais por excelência e, como tais, desprovidas de qualquer natureza extensiva ou espacial? Primeiramente, é preciso entender que as propriedades com pontos de fronteira são aquelas que podem ser analisadas, em grandes linhas, como se fossem *regiões geométricas*. Mais precisamente, as propriedades com fronteiras (ou os estados de coisas a elas associados) admitem ser interpretadas como se estivessem embebidas em um espaço métrico. Uma vez situadas (ou

relacionadas a) em um espaço métrico, tais propriedades se comportariam como *vizinhanças* em tal espaço.

Para exemplificar como esta relação entre propriedades e espaço métrico pode ser feita, consideremos um ponto  $z$ , situado em um espaço métrico  $\langle \mathbf{Z}, \mathbf{d} \rangle$ . Consideremos as vizinhanças  $d_w$ -Viz  $z$ , definidas para números reais cujos valores estão no intervalo  $0 \leq w \leq \gamma$ , sendo  $\gamma$  um número real tão próximo de zero quanto quisermos. Desta maneira, define-se no conjunto  $\wp \mathbf{Z}$  o subconjunto:

$$V_{z;\gamma} \equiv_{df} \{d_w\text{-Viz } z / 0 \leq w \leq \gamma\}.$$

Consideremos agora o conjunto de todas as propriedades. Obviamente, tal conjunto é absurdamente grande e, por conseguinte, é fonte de paradoxos. Entretanto, tal fato será por ora ignorado e postulamos que se possa tomar uma parte própria do conjunto das propriedades como um autêntico contra-domínio de uma função  $h$ , cujo domínio é  $V_{z;\gamma}$ ; tal parte própria é uma parte das propriedades de aridade 1; chamemos tal conjunto de propriedades monádicas de  $\Pi$ , de tal forma que:

$$(\exists h) [h(V_{z;\gamma}) = \Pi]$$

A função  $h$  relaciona cada vizinhança  $d_w$ -Viz  $z$  que pertence a  $V_{z;\gamma}$  a uma propriedade monádica  $\pi_w$  que pertence a  $\Pi$ .

Postulemos agora uma função  $g$  que associa para cada objeto  $a$ , de um domínio de objetos  $\mathbf{A}$ , uma posição  $g(a)$  no espaço métrico  $\langle \mathbf{Z}, \mathbf{d} \rangle$ . Para cada vizinhança  $d_w$ -Viz  $z$ , relacionamos um número ordinal  $i$ , tal que  $0 \leq i \leq \mu$ , sendo  $\mu$  um ordinal transfinito (o valor adequado de  $\mu$  dependerá da cardinalidade do intervalo  $[0; \gamma]$ ; se a cardinalidade de tal intervalo for  $c$ , o cardinal do *continuum*, então  $\mu$  será um ordinal transfinito da terceira classe de números, desde que admitamos a hipótese do contínuo como verdadeira). Sendo assim, o contra-domínio de  $g$  pode ser circunscrito ao conjunto  $\chi = \cup_{i=1,2,\dots,\mu} (d_w\text{-Viz } z)_i$ . Desta forma, temos que  $g(\mathbf{A}) = \chi$ , tal que  $\chi \subset \mathbf{Z}$  e, por conseguinte,  $\langle \chi, \mathbf{d} \rangle$  é um subespaço de  $\mathbf{Z}$ .

Consideremos agora um número real  $\varepsilon$ , tal que  $0 \leq \varepsilon < \gamma$ . Como  $d_\varepsilon$ -Viz  $z \in V_{z;\gamma}$ , então podemos admitir que há um objeto  $c$ ,  $c \in \mathbf{A}$ , tal que  $g(c) \in d_\varepsilon$ -Viz  $z$  ou  $g(c) \notin d_\varepsilon$ -Viz  $z$ .

A partir de uma análise inteiramente intuitiva da noção de vizinhança, chegamos aos seguintes condicionais:

$$C1) d(z, g(c)) < \varepsilon \rightarrow g(c) \in d_\varepsilon\text{-Viz } z.$$

$$C2) d(z, g(c)) > \varepsilon \rightarrow g(c) \notin d_\varepsilon\text{-Viz } z.$$

Também de forma intuitiva, podemos estabelecer a relação que  $c$  tem com a propriedade  $\pi_\varepsilon$ , a partir da pertinência ou não de  $g(c)$  à vizinhança  $d_\varepsilon$ -Viz  $z$ :

$$C3) g(c) \in d_\varepsilon\text{-Viz } z \rightarrow \pi_\varepsilon(c).$$

$$C4) g(c) \notin d_\varepsilon\text{-Viz } z \rightarrow \neg \pi_\varepsilon(c).$$

Partamos agora da hipótese de que  $\neg\pi_\epsilon(c)$ . Disto resulta que  $g(c) \notin d_\epsilon$  - Viz  $z$  e que, por conseguinte,  $\neg(d(z, g(c)) < \epsilon)$ . Introduzamos a hipótese contraditória à primeira, isto é, partamos da tese de que  $\pi_\epsilon(c)$ . Segue-se então que  $g(c) \in d_\epsilon$  - Viz  $z$  e que, por *modus tollens*,  $\neg(d(z, g(c)) > \epsilon)$ . Portanto, da introdução das hipóteses contraditórias  $\pi_\epsilon(c)$  e  $\neg\pi_\epsilon(c)$ , chegamos à conjunção dos seguintes *estados de coisas*:

$$D1) \neg(d(z, g(c)) < \epsilon) \wedge \neg(d(z, g(c)) > \epsilon).$$

Como  $d(z, g(c))$  é um número real, vale então a tricotomia da relação de ordem total, o que nos dá que, a partir de e), chegamos à terceira possibilidade da aludida tricotomia:

$$D2) d(z, g(c)) = \epsilon$$

Reparemos que a condição D1) nos foi dada a partir da aceitação da contradição expressa pela conjunção de  $\pi_\epsilon(c)$  e  $\neg\pi_\epsilon(c)$ ; D2) é a resultante *métrica* da admissão que se pode predicar de um objeto tanto  $\pi_\epsilon$ , quanto não- $\pi_\epsilon$ . Mais do que isto: D2) afirma que o estado de coisas associado a uma contradição simples, relacionada às fórmulas atômicas de *aridade* 1, pode ser identificado com a situação métrica que diz que uma vizinhança em um espaço métrico (vizinhança esta postulada como a contrapartida topológica da referida propriedade) tem um ponto fronteiroço, justamente aquele em que as posições distam da *origem*  $z$  de um valor real e positivo igual a  $\epsilon$ .

O que se quer postular aqui é o fato de podermos introduzir um estado de coisas intuitivo que possa ser correspondente ao fato um tanto quanto extravagante de atribuirmos a um mesmo objeto, ao mesmo tempo, um par de propriedades contraditórias.

Não há dúvida de que o apelo à tradução de um conceito ou propriedade, entes intensionais por excelência, para o universo das vizinhanças, entidades que se associam aos espaços métricos (pressupostamente, entes eminentemente extensionais), pareça artificial. De fato, várias críticas muito pertinentes poderiam ser feitas a esta tentativa de “metrificar” os conceitos. Entretanto, a noção de espaço métrico, apesar de seu lastro geométrico e, por assim dizer, euclidiano (o que nos faz associar tal termo a aspectos do mundo real, físico, em estrito antagonismo com o caráter abstrato dos entes lógicos) de há muito perdeu sua ligação com os dados imediatos da sensibilidade: atualmente, um espaço métrico é uma estrutura algébrica, definida por postulados ou axiomas que não visam, de antemão, ser satisfeitos por estes ou aqueles objetos; os espaços métricos não privilegiam o espaço físico ou qualquer outro conjunto de entidades geométricas como alvo de seus postulados, apesar de seus elementos serem chamados de “pontos”, pontos estes que mantêm entre si a relação métrica de “distância”.

Mas é claro que a tradução de propriedades para espaços métricos requer uma série de quesitos formais que devem ser explicitamente dados. Por enquanto, o postulado de tal tradução é apenas uma idéia seminal que, em outras

ocasiões, deverá ser claramente desenvolvida. Neste artigo, apenas aspectos intuitivos (e formalmente problemáticos, diga-se de passagem) desta possível tradução são apresentados.

## Referências

- BERTO, Francesco. *Theorie dell'Assurdo. I rivali dell'Prinzipio della non-Contraddizione*. [2006];
- DA COSTA, Newton C.A. "The Philosophical Import of Paraconsistent Logic". In: *Journal of Non-Classical Logic* 1, 1-19, [1982];
- DA COSTA, Newton C. A. *Sistemas Formais Inconsistentes*. Editora UFPR, Curitiba, [1964];
- GEMIGNANI, Michael. *Elementary Topology*. Addison-Wesley Publishing Company, Reading-Palo Alto-London-Don Mills, [1967];
- HEGEL, Gottfried W. H. *Wissenschaft der Logik*, published in English. Translation by A. V. Miller as *Hegel's Science of Logic*, Allen & Unwin, [1969];
- LEIBNIZ, G. W. "Letter of Mr. Leibniz on a General Principle Useful in Explaining the Laws of Nature through a Consideration of the Divine Wisdom; to Serve as a Reply to the Response of the Rev. Father Malebranche". Published in English in translation in *Leibniz's Philosophical Papers and Letters*, ed., L. E. Loemker, 31-4, [1969];
- PRIEST. *In Contradiction. A Study of Transconsistent*. Clarendon Press, Oxford, [2006].

## Notas

<sup>1</sup> Sobre as várias formulações do princípio de não-contradição, confira Francesco BERTO, [2006]: *Theorie dell'assurdo. I rivali del Principio di Non-Contraddizione*, capítulo I, pg 29-31.

<sup>2</sup> No original: "Aristotele lo chiama [...] "il principio più saldo di tutti"-*firmissimum omnium principiorum*, dicevano i medioevali. É il Principio di Non-Contraddizione[...] La qualifica di *firmissimum* esprime il fatto che il (PNC) è stato la legge più certa e incontrovertibile del pensiero del pensiero e dell'essere, e quindi è stato posto come fondamento supremo della conoscenza e della scienza"

<sup>3</sup> Em seu livro *Sistemas Formais Inconsistentes*, Da Costa introduz uma hierarquia de sistemas lógicos. Esta hierarquia, que se estende de 0 até  $\omega$  (o primeiro ordinal transfinito), se verifica tanto no cálculo proposicional, quanto no cálculo dos predicados (com ou sem identidade). Sobre os sistemas de Da Costa, ver DA COSTA, [1964], capítulo 1, pg 6-17.

<sup>4</sup>Sobre a importância filosófica da lógica paraconsistente, ver DA COSTA, "The Philosophical Import of Paraconsistent Logic", in: *The Journal of Non-Classical Logic* 1: 1-19, [1982].

<sup>5</sup>No texto original: "Dialetheism is simply the view that some contradictions are true. That is, there are some sentences (statements, propositions, or whatever one takes truth bearers to be),  $\alpha$ , such that both  $\alpha$  and  $\neg\alpha$  are true. [...] One may therefore ask 'does it follow that are contradictions in the world'? In one quite unproblematic sense it does"

<sup>6</sup> “Before a time  $t_0$ , a system  $s$  is in a state  $s_0$ , described by  $\alpha$ . After  $t_0$  it is in a state  $s_1$ , described by  $\neg\alpha$ . What state is it in at  $t_0$ ? *A priori*, there are four possibilities:

(A)  $s$  is in  $s_0$  and  $s_0$  only

(B)  $s$  is in  $s_1$  and  $s_1$  only

( $\Gamma$ )  $s$  is in neither  $s_0$  nor  $s_1$

( $\Delta$ )  $s$  is in  $s_0$  and  $s_1$ .

Of course, there may be no uniform question. Different changes may be changes of different kinds. The crucial question I wish to ask is whether there are any changes in class  $\Delta$ , that is, dialethic changes. [...] I shall argue that there are some changes of type  $\Delta$ ”

<sup>7</sup> “I am in a room. As I walk through the door, am I in the room or out of (not in) it? To emphasize that this is not a problem of vagueness, suppose we identify my position with that of my centre of gravity, and the door with the vertical plane passing through its centre of gravity. As I leave the room there must be an instant at which the point lies on the plane. At that instant am I in or out? Clearly, there is no reason for saying one rather than the other”

<sup>8</sup> Sobre a noção de espaço métrico ver GEMIGNANI, M. C., [1967], p.16.

<sup>9</sup> O termo “ponto de fronteira” é usado aqui com alguma liberalidade. Mais precisamente, o conceito de *fronteira* surge em contextos métricos da forma seguinte. Seja  $\langle X, d \rangle$  um espaço métrico. Tomemos um conjunto  $A$ , tal que  $A \subset X$ . Definamos a distância entre um elemento  $x$  de  $X$ , e o conjunto  $A$ , como sendo dada pela relação seguinte:

$$d(x, A) = \max \lim \inf \{d(x, a) / a \in A\}.$$

Isto é, a distância entre  $x$  e o conjunto  $A$  é o valor máximo que esta distância tem, para todo  $a \in A$ , dentre os valores que são limites inferiores para esta mesma distância.

O conceito de *Fronteira de A em X*, com  $A \subset X$ , é definido da forma seguinte:

$$Fr A_x = \{x \in X / d(x, A) = 0 \text{ e } d(x, X - A) = 0\}.$$

Assim sendo, o que há são conjuntos de fronteira, e não propriamente “*pontos fronteirios*” (ver GEMIGNANI, [1967], pgs 38 e 54).

<sup>10</sup> As condições semânticas de Priest para os conectivos lógicos são as seguintes (ver PRIEST, p.75, *op. cit*):

(1a)  $1 \in v(\neg\alpha)$  sse  $0 \in v(\alpha)$

(1b)  $0 \in v(\neg\alpha)$  sse  $1 \in v(\alpha)$

(2a)  $1 \in v(\alpha \wedge \beta)$  sse  $1 \in v(\alpha)$  e  $1 \in v(\beta)$

(2b)  $0 \in v(\alpha \wedge \beta)$  sse  $0 \in v(\alpha)$  ou  $0 \in v(\beta)$

(3a)  $1 \in v(\alpha \vee \beta)$  sse  $1 \in v(\alpha)$  ou  $1 \in v(\beta)$

(3b)  $0 \in v(\alpha \vee \beta)$  sse  $0 \in v(\alpha)$  e  $0 \in v(\beta)$

<sup>11</sup> “...contradiction is the root of all movement and vitality; and it is only in so far as something contains a contradiction within it that it moves [...]”. A versão inglesa adotada por Priest da *Wissenschaft der Logik*, de 1812, é a seguinte: *Hegel’s Science of Logic*. Published in English translation by A. V. Miller, Allen & Unwin, [1969].

<sup>12</sup> “When the difference between two instances in a given series [...] can be diminished until it becomes smaller than any given quantity whatever, the corresponding difference [...] must of necessity also be diminished or become less than any given quantity whatever”. A citação de Leibniz que Priest usa é retirada da obra seguinte: “Letter of Mr. Leibniz on a General Principle Useful in Explaining the Laws of Nature through a Consideration of the Divine Wisdom; to Serve as a Reply to the Response of the Rev. Father Malebranche”. Published in English translation in Leibniz’. *Philosophical Papers and Letters*, ed, L. E. Loemker, Reidel, [1969].

<sup>13</sup> Embora a tradução matemática que Priest fará a seguir da citação de Leibniz seja adequada, é bom salientar que Leibniz também salienta que, quando uma variável independente caminha continuamente até um valor determinado, então a correspondente variável dependente também o faz da mesma maneira *por necessidade*. Neste sentido, a lei de continuidade estipula um princípio

regulador dos processos naturais, o que mais tarde será explicitamente formulado por Kant, em sua *Crítica da Razão Pura*, de 1781.

<sup>14</sup> “[With] some three hundred years of mathematical hindsight, it is easy enough to think that Leibniz is just saying that, for two mathematical sequences,  $(s_n)$  and  $(t_n)$ , if

$$\lim_{n \rightarrow \infty} s_n - t_n = 0 \text{ then } \lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} t_n$$

<sup>15</sup> “For every propositional parameter,  $p$ , and every  $x, y \in W$ , if  $1[0] \in v_z(p)$  for every  $z$  such that  $x < y < z$ , then  $1[0] \in v_x(p)$  and  $1[0] \in v_y(p)$ ”

<sup>16</sup> “A body cannot be localized to a point it is occupying at an instant of time, but only to those points it occupies in a small neighbourhood of that time”

<sup>17</sup> Por “função contínua”, entende-se o seguinte: Sejam  $\langle X, d \rangle$  e  $\langle Y, d' \rangle$  dois espaços métricos. Uma função  $f$  de  $X$  em  $Y$  é dita *contínua* se, dado qualquer ponto  $f(a) \in Y$  e qualquer número positivo  $k$ , existe um número positivo  $q$  tal que, se  $x \in d_q$  - *Viz*  $a$ , então  $f(x) \in d_k$  - *Viz*  $f(a)$ . (ver GEMIGNANI, *op. cit.*, p.29).

<sup>18</sup> As considerações feitas aqui sobre o movimento contínuo não dizem respeito ao movimento contínuo em geral. Na realidade, o caso analisado no presente texto se limitou a um deslocamento posicional contínuo que é *similar* ao fluxo do tempo, isto é, sendo  $t$  e  $t'$  dois instantes de tempo, tais que  $t < t'$ , então a função contínua  $f$  preserva a relação de ordem entre  $t$  e  $t'$  em suas respectivas imagens, ou seja,  $f(t) < f(t')$ .

II

---

ONTOLOGIA DA  
MECÂNICA QUÂNTICA



# O REALISMO CIENTÍFICO SCHRÖDINGUERIANO

CAROLINE ELISA MURR

*Universidade Federal de Santa Catarina*

e-mail: caromurr@gmail.com

## 1. Introdução

Erwin Schrödinger (1887-1961), nascido em Viena, Áustria, foi um dos mais importantes cientistas do seu tempo. Além disso, ele foi um dos fundadores da Mecânica Quântica não relativística, para a qual a Equação de Schrödinger é sua a mais conhecida contribuição. No entanto, muito menos conhecida do que seu trabalho como cientista é a sua obra como filósofo.

Schrödinger esboçou suas ideias filosóficas em muitos textos e livros ao longo da sua vida. Os primeiros foram publicados por volta de 1932. Nesses textos, ele refletiu sobre questões como o objetivo e o papel da Ciência na sociedade humana. Mais tarde, ele começou a explorar mais profundamente algumas questões epistemológicas. Por exemplo, Schrödinger analisou o que ele chamou de “construção das coisas da vida cotidiana”, bem como a “construção dos objetos da Ciência”. Neste texto, vamos expor suas reflexões com respeito a ambos os assuntos, enfatizando o segundo anteriormente citado. Seguindo não somente os textos de Schrödinger mas também alguns comentadores, nosso objetivo é chegar a um esboço de sua posição com respeito à realidade dos objetos da Ciência, especialmente da Física. Resumidamente, defendemos que sua abordagem pode ser considerada como um tipo de realismo científico, embora mais complexo do que a “forma pura” dessa posição filosófica. Por essa razão, o denominamos “realismo científico Schrödingueriano”.

## 2. A construção das coisas da vida cotidiana

Em um de seus primeiros textos filosóficos, Schrödinger propõe uma questão: em que sentido se diz que alguma coisa, tal como uma cesta de frutas, é real? Com o objetivo de responder a essa pergunta, ele começa dizendo que se reconhece as coisas tangíveis da vida cotidiana não somente com base em sensações reais, mas também algumas expectativas. De acordo com ele, tais objetos são “nada mais do que uma *configuração*, a qual serve para unir certas percepções sensoriais, algumas das quais são reais, enquanto que a maioria é

apenas virtual” (SCHRÖDINGER, [1928], p. 120, grifo do autor). Consequentemente, para Schrödinger, os objetos da vida cotidiana são configurações montadas, ou construções, e tais configurações são constituídas por sensações reais (percebidas através dos sentidos) e percepções que ele chama de “virtuais”. Esse último tipo de percepção envolve nossas expectativas para com o objeto. Por exemplo, o que imaginamos fazer com ele. Também pode envolver nossas expectativas com respeito à percepção dos objetos por parte de outras pessoas (Schrödinger, [1954a], p. 146). Então, a realidade dos objetos não dependeria apenas de sensações causadas por eles. Respondendo à pergunta de Schrödinger, chama-se uma cesta de frutas de “real” em dois sentidos complementares. Primeiro, é possível ter sensações dela. Segundo, é possível imaginar como agir com relação a ela.

Segundo Schrödinger, as expectativas são instintivas no início do processo de construção das coisas da vida cotidiana. A razão para tal é que esse processo começa a acontecer por experimentação durante a mais tenra infância. Portanto, esse seria um processo em grande parte inconsciente desde o início. Schrödinger chama esse processo de formação de “invariantes”. O termo “invariantes” é emprestado da Matemática e significa, nesse contexto de construção das coisas da vida cotidiana, “o que não muda, independentemente do ponto de vista”. Tal processo começa com o próprio indivíduo, estendendo-se logo para invariantes mútuos, comuns a indivíduos em contato social (Schrödinger, [1954a], p. 146). Por meio desse processo é que todos os seres humanos acabam fazendo a mesma construção, percebendo um e o mesmo mundo invariante.

Embora invariante, o mundo não resultaria em algo fixo e acabado. Em primeiro lugar, não se trata de um método, que poderia ter um resultado fixo e acabado. Schrödinger diz que “a formação de invariantes é um comportamento que adotamos desde a mais tenra infância, desenvolvemos à mais alta perfeição e usamos a todo minuto desperto de nossa vida, para nos orientar com relação a tudo o que nos rodeia diariamente” (SCHRÖDINGER, [1954a], p. 146). Em segundo lugar, não haveria fim para as percepções e sensações potenciais que cada um poderia ter com relação a um objeto real (Schrödinger, [1954b], p. 95). Portanto, a realidade não poderia ser considerada acabada ou fixa.

Além disso, tal realidade não-fixa e mutável não seria independente dos indivíduos que fazem as construções. No entanto, seria conveniente para esses indivíduos considerar sua existência objetiva, independentemente deles ou de suas mentes. A conveniência dessa realidade é, no entanto, uma estratégia, válida para a vida prática. É o que Schrödinger chama de “hipótese do mundo real ao nosso redor” (SCHRÖDINGER, [1956], p.132). Michel Bitbol, um físico e filósofo francês que tem se dedicado ao estudo dos escritos de Schrödinger, afirma que, para Schrödinger, nós precisaríamos “tomar essas construções muito a sério, uma vez que elas são pré-condição para a nossa vida” (BITBOL, 1996, p. 14). Por essa razão, Bitbol sustenta que Schrödinger teria uma “atitude realista” fundada em uma “doutrina antirrealista” subjacente. Consequentemente, o realismo de

Schrödinger com respeito às coisas da vida cotidiana seria apenas metodológico (Bitbol, 1996, pp. 14 e 39).

Em suma, para Schrödinger a realidade seria, considerando as coisas da vida cotidiana, um construto ilimitado e não-fixo, não independente do sujeito que o monta. Entretanto, ele diz que a ideia de uma realidade independente é conveniente, conforme já indicamos. Ele explica esse ponto de vista aparentemente contraditório através da introdução do que ele chama de “princípio da objetivação” (SCHRÖDINGER, [1956], p. 131). Segundo Schrödinger, esse princípio consiste em inconscientemente remover “o Sujeito Conhecedor do domínio da natureza, o qual nos esforçamos por entender” (SCHRÖDINGER, [1956], p. 132). Consequentemente, esse processo torna o mundo objetivo, sendo que os sujeitos passam a vê-lo como se não fizessem parte dele.

É a partir do estabelecimento tácito desse princípio que surge a distinção sujeito/objeto (Schrödinger, [1956], p. 140). Além disso, para Schrödinger a ideia de uma realidade independente do sujeito também tem como consequência a conhecida ideia Kantiana da “coisa em si”. Schrödinger não concorda com essa ideia, segundo a qual há uma realidade em si mesma, independente do sujeito, a ponto de que não se pode ter conhecimento dela. Schrödinger afirma que “o mundo me é dado apenas uma vez, não uma vez como existente e uma vez como percebido” (SCHRÖDINGER, [1956], p. 140). Logo, a construção das coisas da vida cotidiana não resultaria em outra manifestação do mundo, distinta da sua forma real. Para ele, o mundo é o que todos os seres humanos constroem através da formação de invariantes, sem uma realidade em si mesma escondida do conhecimento humano.

### **3. A construção dos objetos da Ciência**

Os objetos da Ciência também são considerados construtos, para Schrödinger. No entanto, na interpretação de Bitbol, seu significado não é diminuído diante dessa concepção. Não deveríamos chamá-los de meros produtos de nossas mentes. Segundo Bitbol, Schrödinger aconselha que, com respeito a construtos teóricos adequados, dever-se-ia “pensar que eles são exatamente o mesmo tipo de estrutura que aquela que estamos acostumados a chamar de ‘objeto real’” (BITBOL, 1996, p. 14, ênfase do autor). Portanto, construtos científicos teriam o mesmo status ontológico que as coisas da vida cotidiana, para Schrödinger. No entanto, há uma diferença no modo pelo qual nós moldamos cada um deles.

A diferença concerne a percepção dos objetos. As reflexões de Schrödinger foram motivadas pela Física Quântica, cujos objetos são coisas diminutas, dos quais não se pode ter as chamadas “percepções reais”. O único tipo de percepção que se pode ter desses objetos é o que Schrödinger chamou de “virtual”. Essas percepções virtuais seriam baseadas em experimentos que envolvem objetos macroscópicos (Schrödinger, [1928], p. 123).

Além dessa diferença, há um ponto comum entre ambos os construtos, da vida cotidiana e científicos. Schrödinger conecta a construção dos objetos da Ciência ao princípio da objetivação em *Mente e Matéria*, publicado em 1956. Esse seria um dos princípios básicos do método científico, segundo ele (Schrödinger, [1956], p. 131), sendo que seu uso seria um acordo tácito entre os cientistas. Para Schrödinger, a humanidade concluiu que “um quadro moderadamente satisfatório do mundo somente seria alcançado pelo elevado preço de nos retirarmos dele, descendo ao papel de um observador desinteressado” (SCHRÖDINGER, [1956], p. 133). Por essa razão, o princípio da objetivação tem sido a base do método científico. Portanto, a Ciência teria se desenvolvido a partir do princípio da objetivação, ao qual permanece profundamente ligada.

Pode-se dizer que, para Schrödinger, a Ciência do seu tempo constrói os objetos baseando-se na observação de experimentos. No entanto, a Ciência não descreve tais experimentos ou os objetos utilizados nos experimentos, que fazem parte do mundo de coisas cotidianas. Deste modo, para Schrödinger, a Ciência faz previsões sobre objetos que são construídos pela própria Ciência. As construções da Física, por exemplo, seriam simplesmente abstrações. Ben-Menahem, que escreveu o artigo “Lutando contra o realismo: o caso de Schrödinger”, interpreta que tais construções somente seriam possíveis, para ele, através da abstração da experiência. Além disso, o abismo entre o modelo construído pela Ciência e a experiência na qual esse modelo é baseado seria “intransponível” (Ben-Menahem, 1992, pp. 27-8). Em suma, a Ciência tentaria explicar a experiência humana, que envolve coisas da vida cotidiana. No entanto, a Ciência o faz abstraindo da experiência. Essa abstração envolve os construtos dos quais falamos, isto é, os objetos da Ciência. Logo, tais abstrações não se referem mais à experiência humana, mas a esses objetos científicos.

Tais abstrações utilizadas pela Ciência seriam as teorias científicas. Do parágrafo anterior podemos concluir que as teorias científicas não teriam uma perfeita relação de correspondência com os fatos observados. As teorias seriam relacionadas aos fatos por meio dos objetos da Ciência. Também concluímos que os objetos da Ciência são “puras configurações”, para Schrödinger. Entretanto, ele afirma que nós só podemos nos referir aos objetos da Ciência como sendo reais. Para ele, se nós chamamos reais as construções das coisas tangíveis ao nosso redor, nós teríamos que fazer o mesmo para os objetos da Ciência. Todavia, eles são metodologicamente distintos, embora com o mesmo status ontológico.

#### **4. Dois tipos de objetos, uma realidade**

De acordo com Bitbol, mais do que o mesmo status como objetos reais, Schrödinger tem a intenção de conferir o status de coisa inacabada, em construção, tanto aos objetos da Ciência quanto aos objetos da vida cotidiana. Bitbol defende que “é nessa atmosfera de ‘abertura’ que Schrödinger concebe o

status dos modelos e teorias científicas” (BITBOL, 1992, p. 56, ênfase do autor). Consequentemente, não apenas os objetos da vida cotidiana estão sujeitos a mudanças, mas também os construtos científicos. Em ambos os casos, é a presença das percepções virtuais que permite tal abertura. Conforme já destacamos, as percepções virtuais incluem expectativas, ou saber como agir com relação aos objetos. Consequentemente, objetos que são reais nesse sentido guiam nossas ações. Bitbol conclui que as representações construídas pela Ciência teriam um papel regulatório na pesquisa científica (Bitbol, 1996, p. 41). De fato, pode-se incluir as previsões científicas entre essas expectativas. Essas previsões guiam o trabalho científico em grande medida. Logo, o realismo de Schrödinger implica que os objetos reais nos constroem a agir. Ben-Menahem propõe uma interpretação semelhante, dizendo que Schrödinger “vê o conceito de realidade como um construto inevitável, que guia nossos pensamentos e comportamentos” (BEN-MENAHM, 1992, p. 26).

O próprio Schrödinger reitera a realidade comum das coisas da vida cotidiana e dos construtos da Ciência com o objetivo de destacar seu caráter revisável. Referindo-se ao início do processo de formação de invariantes, ele afirma que não se pode dizer que essa “é tão somente a maneira pela qual a criança aprende como o mundo realmente é” (SCHRÖDINGER, [1954a], p. 149). Isso seria trivialmente verdadeiro para ele, pois nós consideraríamos que o mundo é realmente como tal porque todos nós o construímos da mesma maneira. No entanto, a unicidade dessa construção não implica a sua inevitabilidade: “eu os equiparo a construtos científicos. Isso os torna passíveis e capazes de serem sujeitos a revisão e mudados e melhorados, como todas as teorias científicas são” (SCHRÖDINGER, [1954a], p. 149).

Embora revisáveis, os construtos científicos e as coisas da vida cotidiana não são completamente abertos. Eles dependem da formação de invariantes, que unifica a realidade como sendo a mesma para todos. Na Ciência, a formação de invariantes seria uma continuação do comportamento em grande parte inconsciente descrito para a vida cotidiana. No entanto, o processo não é mais inconsciente na Ciência, embora seja tacitamente aceito, como o princípio da objetivação. Para Schrödinger, a busca de invariantes se torna completamente consciente na Ciência (SCHRÖDINGER, [1954b], p. 146). É desse comportamento que surge a busca por teorias válidas ou verdadeiras.

## **5. A realidade dos construtos científicos**

O realismo científico tradicional envolve a correspondência entre teorias científicas e coisas no mundo. Por exemplo, seria necessário, de acordo com essa posição filosófica, que um elétron na Teoria Quântica tivesse um correspondente no mundo. Além disso, seria necessário que ele se comportasse como a teoria prevê. Logo, ser um realista nesse sentido implica acreditar que a matéria como

a percebemos nas coisas da vida cotidiana é realmente constituída dos objetos que a teoria afirma.

Segundo French (2009, p. 105), o Realismo Científico inclui a ideia de que as teorias científicas podem ser falsas ou verdadeiras. Tal posição cria um problema, uma vez que teorias científicas mudam. Tal problema traz o argumento contra o realismo científico conhecido como Meta-indução Pessimista, abreviado como MIP. Segundo o MIP, a História da Ciência tem mostrado que algumas teorias, antes consideradas verdadeiras, são mais tarde descartadas. Logo tem-se boas razões para acreditar que algumas teorias consideradas verdadeiras hoje serão consideradas falsas no futuro. MIP é uma das principais correntes de crítica contra o realismo científico. E Schrödinger concordaria com tal crítica. Ele afirma que “não temos o direito hoje de dizer que a teoria corpuscular de Newton era a errada, embora fosse costume por um bom tempo declarar isso” (SCHRÖDINGER, [1932], p. 74).

Não apenas as opiniões de Schrödinger, mas também suas ideias, confirmam as diferenças entre esse tipo de realismo científico e aquele encontrado em seus escritos. Em primeiro lugar, no Realismo Científico de Schrödinger não há correspondência entre as teorias científicas e as coisas da vida cotidiana. Elas podem ser ligadas por meio de uma interpretação. Mesmo assim, essa ligação é completamente diferente daquela afirmada pelo realismo científico que aqui estamos chamando de tradicional. Por um lado, o realismo científico afirma a correspondência entre construtos científicos e coisas no mundo, no sentido que já apontamos acima. Os objetos científicos são tomados como reais uma vez que são parte de coisas reais, por exemplo, o elétron é supostamente parte da constituição da matéria. Por sua vez, a matéria para o cientista é correspondente à matéria da vida cotidiana. Por outro lado, no Realismo Schrödingueriano não há essa correspondência. Por meio de uma interpretação, no entanto, há uma correspondência não-perfeita entre as teorias científicas e os experimentos, isto é, coisas da vida cotidiana. Embora conecte a experiência da vida cotidiana e construtos científicos, tal interpretação apenas explica o estabelecimento de uma com base nos outros. A razão para se considerar elétrons, por exemplo, como reais, não é mais a sua existência como constitutivo da matéria da vida cotidiana. Eles são considerados reais uma vez que guiam o trabalho científico, permitem previsões e fornecem resultados que têm correlação com as coisas da vida cotidiana.

## 6. Conclusão

Finalmente, podemos resumir o Realismo Científico Schrödingueriano apontando algumas características. A primeira é a presença das percepções virtuais no processo de construção da realidade. O que coloca os construtos científicos e da vida cotidiana juntos como objetos reais, para Schrödinger, são

as percepções virtuais. Elas o fazem através da sua característica de guiar a ação. A segunda característica vem da primeira. A realidade é aberta, não-fixa e não-acabada devido ao caráter das percepções virtuais. Em terceiro lugar, não há correspondência da matéria da Ciência e da matéria da vida cotidiana. Apenas uma correlação é aplicável para explicar a construção dos objetos científicos com base nos experimentos que envolvem construtos da vida cotidiana.

A quarta característica do Realismo Schrödingeriano é que a realidade vem do princípio de objetivação. De fato, os objetos são resultado da necessidade do estabelecimento de sujeito e objeto, para Schrödinger. Finalmente, a quinta e última característica que gostaríamos de ressaltar é a formação de invariantes, que permite a unicidade de realidade. Na Ciência, a busca por invariantes também permite o acordo sobre o que deve ser considerado verdadeiro dentre tantas teorias científicas.

## Referências

- BEN-MENAHM, Yemima. 1992. "Struggling with realism: Schrödinger's case". In: BITBOL, Michel and DARRIGOL, Olivier. (eds.). *Erwin Schrödinger. Philosophy and the Birth of Quantum Mechanics*; pp. 25-40. Paris: Editions Frontières.
- BITBOL, Michel. 1992. "Esquisses, Forme et Totalité: Schrödinger et Le concept d'objet". In: BITBOL, Michel and DARRIGOL, Olivier. (eds.). *Erwin Schrödinger. Philosophy and the Birth of Quantum Mechanics*; pp. 41-80. Paris: Editions Frontières.
- \_\_\_\_\_. 1995. "Introduction". In: BITBOL, Michel (org.). *The Interpretation of Quantum Mechanics: Dublin Seminars (1949-1955) and Other Unpublished Essays*. Woodbridge: Ox Bow Press.
- \_\_\_\_\_. 1996. *Schrödinger's Philosophy of Quantum Mechanics*. Dordrecht: Kluwer Academic Publishers.
- FRENCH, Steven. [2007] 2009. *Ciência. Conceitos Chave em Filosofia*. Porto Alegre: Artmed.
- SCHRÖDINGER, Erwin. [1928] 1935. "Conceptual Models in Physics and their Philosophical Value". In: *Science and the Human Temperament*, pp. 119-138. Translated from German by: JOHNSTON, W. H. London: George Allen & Unwin Ltd.
- \_\_\_\_\_. [1932] 1935. "Is Science a Fashion of the Times?". In: *Science and the Human Temperament*, pp. 66-85. Tradução do alemão por: MURPHY, J. London: George Allen & Unwin Ltd.
- \_\_\_\_\_. [1954a] 1995. "The Part of the Human Mind". In: BITBOL, Michel (ed.). *The Interpretation of Quantum Mechanics: Dublin Seminars (1949-1955) and Other Unpublished Essays*, pp. 141-149. Woodbridge: Ox Bow Press.
- \_\_\_\_\_. [1954b] 1996. "Nature and the Greeks". In: *Nature and the Greeks and Science and Humanism*, pp. 3-99. Cambridge: Cambridge University Press.
- \_\_\_\_\_. [1956] 1992. "Mente e Matéria". In: *O que é Vida? O Aspecto Físico da Célula Viva seguido de Mente e Matéria e Fragmentos Autobiográficos*, pp. 107-174. Translated from English by: ASSIS, J. P. and ASSIS, V. Y. K. P. Cambridge and São Paulo: Cambridge University Press and Editora Unesp.

# A CALCULUS OF *NON-INDIVIDUALS* (IDEAS FOR A QUANTUM MEREOLGY)

DÉCIO KRAUSE

*Department of Philosophy, Federal University of Santa Catarina*

dkrause@cfh.ufsc.br

**Abstract:** In this paper we discuss some difficulties we see in the attempts of developing a mereology suitable to cope with typical facts involved in a possible interpretation of quantum mechanics. Our aim is to present a calculus of individuals (really, a calculus of *non*-individuals) that (hopefully) cope with the strange behaviour of quantum objects. We sketch a basic axiomatic assuming the concept of quasi-set (French & Krause [2010]), more or less in the directions indicated by Tarski [1983], who assumed a set theory in his approach to the subject. We do not develop the corresponding theory here in full, for our aim is just to introduce some elements in the discussion; for instance, in this first approach we do not consider time (Simons [1987], Cha.5). But we touch on what we believe may be the main difficulties in the development of a quantum mereology, namely, quantum holism, according to which a whole cannot simply be reduced to the (mereological) sum of its parts, and quantum indistinguishability, which entails that it would be indifferent whether a certain object has other objects as its parts or some indiscernible from them instead. We also consider that a physical system is not only a mere aggregate of parts in whatever sense of this expression, but its parts must be aggregated (composed, structured) in some way. Thus we suggest that some notion of *form*, or *shape*, must be considered too and we discuss a possible relationship of this concept with that of a *structure*. In our proposal, our quantum mereology is non extensional in a sense (but in a different sense than standard intensional mereologies), obeying a weak form of extensionality ascribed by quasi-set theory.

**Keywords:** Quantum mereology, calculus of non-individuals, quasi-sets, indiscernibility, non-individuality, mereology.

## 1. Introduction

Mereology, as commonly characterized, is the theory of the wholes and their parts, and of associated concepts. Stanislaw Lesniewski pioneered its formal development in 1916, although the philosophical roots can be traced back to antiquity (Varzi [2009]). In the remaining XXth century, the subject was treated formally and improved in several different ways by a cluster of authors. Simons [1987] surveys most of the known mereologies; see also Varzi op.cit..

In the most common approaches, mereologies are dealt with according to two basic lines: either by presupposing a set theory (Tarski and others),<sup>1</sup> or

without assuming a set theory, as preferred for instance by Nelson Goodman (in his *Structure and Appearance* [1966], due to his nominalistic credos).<sup>2</sup> Here we shall use the first method, but instead of a set theory, we shall be dealing with *quasi-set theory* for reasons to be explained below.

There are some questions to be considered in proposing a mereology directed to deal with quantum systems and their ‘parts’. The first question is related precisely with the sense we can ascribe to the notion of ‘part’. We may call *quantum holism* an idea that emerges in the quantum realm in different forms; for instance, Healey distinguishes between ‘explanatory holism’, the view that a satisfactory explanation of the behavior of a system cannot be given by explaining the behavior of its parts, and ‘property holism’, namely, the view that the properties of a whole are not wholly determined by those of its parts. Both views, he says, appear in quantum phenomema (Healey [2009]). As also says Healey, “[m]any phenomena in condensed matter physics are explained by applying quantum mechanics directly to systems composed of very large numbers of atomic and subatomic particles: only in special cases can the theory be applied at the level of these components”, and continues: “[i]n quantum mechanics, the state of a compound system is not always determined by the states of its components: each such failure of determination in quantum mechanics is an example of *state holism*. Schrödinger called the subsystems in such a compound state ‘entangled’. Assuming a system’s state specifies its properties, state holism implies property holism.” (ibid.). Thus, we may say that a form of quantum holism arises from entanglement and from the nonseparability of quantum systems (say, in an EPR typical experiment). Another question is, in our view, the fact that the ‘parts’ may be so that the exchange of one of them by ‘another’ of the similar species would not cause any difference to the whole (due to the existence of permutational indiscernibility).<sup>3</sup> This is different in saying that once cat Tibbles loses her tail in an accident and some miraculous transplant provides her another tail, what we shall have after the surgery is *the same Tibbles with another tail* (when the notion of ‘the same’ makes sense), that is, her ‘parts’ have of course be changed, although she is the very same individual. But in regarding quantum systems, this is apparently not so (at least believing in quantum mechanics), as we shall discuss in more details soon. In addition to these questions, we should also consider that a physical system is not only a mere aggregate of parts in whatever sense of this expression (that is, an individual is not a *set* in the sense of standard set theories), but its parts must be aggregated (composed, structured) in some way. Concerning this point, we suggest that some notion of *form*, or *shape*, must be considered too. But, firstly, let us comment on our basic logic.

Quasi-set theory (see French & Krause [2010] for an updated discussion) was created to cope with collections of indiscernible (or indistinguishable) objects, which of course cannot form *sets* of standard set theories for, among other things, they would not obey the usual Extensionality Axiom. The lack of identity of the so called m-atoms (see the next paragraph) has its main motivation in the (assumed) non-individuality and indistinguishability of quantum objects,

and assumes that the very notion of identity, ascribed by the standard theory of identity of classical logic and mathematics (set theory), poses a cluster of problems in dealing with indiscernible objects (cf. French & Krause [2006]). The theory enables us to deal in a ‘set theoretical sense’ with the second mereological question posed above, namely, that a ‘part’ may be exchanged by a similar one keeping the whole indistinguishable from the original one. Suitable mereological axioms added to this theory maybe enable us to treat also the first and the third questions.

Quasi-set theory encompasses ur-elements of two sorts, termed M-atoms and m-atoms. The first ones behave as the standard ur-elements of ZFU, while for the later the standard notion of identity is assumed to fail (expressions such as  $x=y$  are not considered as formulas if either  $x$  or  $y$  stand for m-atoms), but there is a weakly relation of indiscernibility (or indistinguishability), represented by ‘ $\equiv$ ’ holding among them. Furthermore, there is a defined notion of identity holding among M-atoms and ‘sets’ (quasi-sets formed by M-atoms or another sets), which has the properties of standard identity –but this notion does not hold for m-atoms. But in the theory there is no relationship between the two kinds of atoms. If we suppose that an intended interpretation is given, according to which the M-objects (which interpret the M-atoms) represent the usual objects or our surroundings, while the m-objects stand for quantum entities, it would be interesting to analyse in a formal way how M-atoms can be ‘composed’ by m-atoms. Of course this relationship would be described by some kind of mereology.

## **2. The non-individuality of quantum objects and quantum holism**

The non-individuality of quantum objects was fully discussed by French and Krause [2006], so that we shall take here just a few points. As pioneered by Schrödinger and other forerunners of quantum theory (Bohmian mechanics is left out from this first discussion), in the beginnings of quantum mechanics quantum objects (termed ‘elementary particles’) were regarded as not having individuality in the *usual sense*. For instance, obeying Bose-Einstein statistics, they could not be taken as individuals in the standard sense, for a permutation of two bosons does not conduce to another state (the case of fermions is a little bit different, for they obey Fermi-Dirac statistics, but the permutations do not exchange the square of the wave function). This ‘usual sense’ might be roughly described as follows. By an *individual* we understand something “having an identity”, so that it can be recognized as such in different opportunities. An individual is something that at least in principle can be distinguished from any other individual, even of the similar species. Of course, as most defenders of the standard interpretation of quantum mechanics seem to accept, this is not so with quantum objects, for they may be totally indiscernible (mainly in the case of

bosons). Furthermore, if we have a collection of individuals and suppose that one of its elements is exchanged with another individual not belonging to the collection, the new collection turns to be distinct from the original one. We may say that collections of individuals are *extensional*, for they seem to obey the axiom of extensionality of standard set theories. Of course such a characterization presupposes a theory of identity, and we usually assume this theory as given by classical logic and mathematics (say, that mathematics that can be build in Zermelo-Fraenkel set theory, ZF).<sup>4</sup> An individual, leaving aside the very origin of the term, may be simple (with no proper parts) or complex, that is, formed by other individuals. But the relationship between an individual and its parts cannot be seen as similar to that there exists between a set and its elements; an individual is not a set, although sets of standard set theories may be taken as individuals in the above sense. Furthermore, as the case of the Tibbles cat mentioned above suggests, an individual may lose some of its parts without turning to be a different individual, contrary to an extensional set, which without an element turn to be *another* set. This shows that a theory of identity for individuals, given in terms of its parts, is something to be pursued (contrary to what happens, say, with extensional sets and their elements, whose theory of identity is given as described in the previous footnote). But we shall circumvent such a definition for we shall consider our objects differently, namely, as *non-individuals*.

There are some basic traits that induce us to consider the non-individuality of quantum objects. Firstly, in general quantum objects cannot be physically re-identified through time. If two quanta of the same kind (say, two electrons) have their wave functions overlapped, they *lose* their individualities. Secondly, they are invariant under permutations, that is, if a quantum object is substituted by another one of similar species, nothing physically relevant is perceived at all (we shall be back to this point below).<sup>5</sup> Notwithstanding, non-individuals may be collected into amounts, and their collections may have a cardinal; for instance, we can say that there are six electrons in the level 2p of a sodium atom ( $1s^2 2s^2 2p^6 3s^1$ ), despite there is no way of identifying them individually; these collections may have a cardinal, but not an associated ordinal, as described by quasi-set theory. In other words, it makes sense to say that there are six electrons at the level 2p, but not which is which.

Thus, by considering situations such as these ones, we think it is interesting to pursue the idea of non-individuality of quantum objects also from the point of view of a mereology.<sup>6</sup> In this first paper on the subject, we shall not consider neither a mereology of continuants, which takes temporal considerations into account, nor modalities, which turn mereology non extensional (our non extensionality will be reached by other means), although we report them as relevant for any discussion involving quantum theory and physical objects in general.

Important to remark that, in our account, not all non-individuals are indistinguishable. Really, the hitherto mentioned six electrons of the level 2p can be distinguished by their quantum numbers (for instance, by their magnetic

numbers or by their spins)<sup>7</sup> for, being fermions they obey Pauli's exclusion principle and so cannot be all in the same quantum state (but this is possible for bosons). Non-individuals are characterized by their discrepancy from the standard informal notion of individual suggested above and, we could say, also from the characterization of the objects given by the theory of identity of classical logic or by the individuals in the sense of Whitehead & Russell's *Principia Mathematica* (also cf. Tarski [1983a], p.25n). This concept can be extended also to relativistic quantum mechanics (quantum field theories). There, 'particles' can be seen as field excitations, but even quantum fields manifest a form of holism; as R. Healey says, "[t]heir indiscernibility [that is, of field quanta], superposability and failure of localization makes field quanta like photons poor candidates for distinct parts of the field, suggesting ontological holism" (Healey [2009]). *Ontological holism*, according to him, denies that some supposedly composite systems have proper parts.

### 3. Looking for a calculus of non-individuals: enters the form

Let us call L a language that extends the language of quasi-set theory by adding the further primitive binary predicates P (part of), and F (form, or shape). For facility, we shall use a three sorted language encompassing individual variables  $x, y, z, w, \dots$  for m-atoms,  $X, Y, Z, W, \dots$  for M-atoms, and small Greek letters for quasi-sets. The atoms, when spoken without specification of the kind they belong to, shall be represented by neutral variables  $s, t, w, u, v, \dots$ . The following definitions will be useful, and in most cases we shall omit the universal quantifiers. The definition of formula is standard; thus,

Definition 1.

- (a) [*Disjointness*]  $Dst := \sim \exists w(Pws \ \& \ Pwt)$ , where  $\sim$  is the symbol for negation,  $\&$  is the logical conjunction, and the variables  $s, t$  and  $w$  are of the same kind.
- (b) [*Overlapping*]  $Ost := \exists w(Pws \ \& \ Pwt)$ , with the same remarks above, which shall be not mentioned from now on.
- (c) [*s sums  $\alpha$* ] Let  $\alpha$  be a collection (quasi-set) of atoms of the same species. Then,
 
$$s\Sigma\alpha := Dts \leftrightarrow (u \in \alpha \rightarrow Dtu).$$
<sup>8</sup>
- (d) [*[x] sums s*] Let  $[x]$  be a pure quasi-set, that is a quasi-set whose elements are m-atoms, and let  $s$  be either a M-atom or a m-atom. Then we say that  $[x]$  sums  $s$  iff any object disjoint from the elements of  $[x]$  is disjoint from the parts of  $s$ . In symbols,

$$[x]\Sigma s := Dyx \leftrightarrow (Pvs \rightarrow Dyv).$$

(here,  $x$  is representing an arbitrary element of  $[x]$ ). Thus, we may say that the quasi-set whose elements are six Hydrogen atoms, two Carbon atoms and one

Oxygen atom may sum both an ethylic alcohol molecule and a methylic ether molecule (see the discussion that follows). The difference between these two molecules is given by their form, as we shall discuss later.

The first two definitions speak by themselves. The third one tells us that an object  $s$  sums a collection  $\alpha$ , or that  $\alpha$  has  $s$  as its sum, when everything disjoined from  $s$  is disjoined from the parts of  $\alpha$  and conversely. Thus,  $m$ -objects may sum  $m$ - or  $M$ -objects, but  $M$ -objects sum  $M$ -objects only. Here an interesting problem appears, which is characteristic of quantum mechanics, and captured by quasi-set theory, a problem we have already touched upon before. This is the problem of the invariance by permutations, which in quantum physics is described by the so-called Indistinguishability Postulate (French & Krause [2006], *passim*); roughly speaking, it says that the expectation value in the measure of any observable is the same before and after the permutation of indistinguishable particles. That is, no permutation of indistinguishable particles interferes in the physical results got from measurements. In quasi-set theory, we have the following theorem, which captures (at least part of) this idea. The theorem says that if in a quasi-set we ‘exchange’ (this can be done by the quasi-set theoretical operations –see French & Krause [2010], theorem 3.1, p.110) an element by an indistinguishable one, the resulting quasi-set is indistinguishable from the original one. Indistinguishability is a primitive notion in the theory and acts as an equivalence relation. Thus, similar ideas can be applied for the objects that sum a certain collection of objects (this should be proved, but we shall postpone the proof for a while). As a consequence, the objects which sum other objects cannot be taken as formed by *specific* parts, for indistinguishable parts may act in their places as well. So, we have here a form of holism: the whole cannot simply be reduced to the (mereological) sum of its parts, for these parts may change without altering the whole (truly, the exchanges turn the whole indiscernible from the original one).

This last remark may chock some philosophers and logicians, yet perhaps do not chock scientists. To see its reasonableness, just think in an atom in the neutral state, say a Helium atom, He. It comprises two electrons, and by ionization, we can ‘eliminate’ one of the electrons in order to get a positive ion, He<sup>+</sup>. Now, by another chemical process, we can capture an electron again, making the ion neutral once more. Two questions naturally arise: (a) is the ‘new’ atom *the same* as the original one? (b) once the two electrons (after capturing again one electron) become entangled, can we say which is the electron which was captured? It’s impossible to say. This problem is exactly similar to the old philosophical problem of persistence through time (Haslanger [2003]). I tend to agree with Hume’s answer to the problem of re-identification of objects (Hume [1986]) as being a fiction of our imagination, something we do by force of the habit. In physics, these possibilities (namely, the permutation by indiscernibles) is generally not taken into account, for it does not matter for the measurement results, and when some explanation is required, we report to some *ad hoc* assumption, such as the Indistinguishability Postulate. I think that the best (although perhaps very complicated from the technical point of view) way of talking about these things is

to speak in terms of the indistinguishability (or indiscernibility) of the considered entities after discontinuous observations. That is, despite a little bit pedantic, we should regard the objects (persons inclusive) we meet in different times as merely indiscernible from other appearances, and not as being *the same* object or person. Strange it would be for our day-to-day way of life, philosophically it might be interesting (can you imagine that your wife could be said today to be not identical from her as she was yesterday? But, is she?). We shall leave these intricate questions for another forum and discuss the problem from another point of view, yet we will touch in this point again below.

Standard extensional mereologies, as Leonard & Goodman's calculus and Lesniewski's systems, presuppose that two objects having the same parts are the same object: they are identical. This assumption is termed *mereological extensionality* by Peter Simons (op.cit., p.1). Should we postulate that? (The axioms of our calculus shall be presented in the sequence). I think not. And this is due not only to quantum facts, but to the standard interpretation too. Really, think of two tiles as in the figure below. We could say that they have the same parts (drawn in different ways), but they should not be regarded as identical.<sup>9</sup>

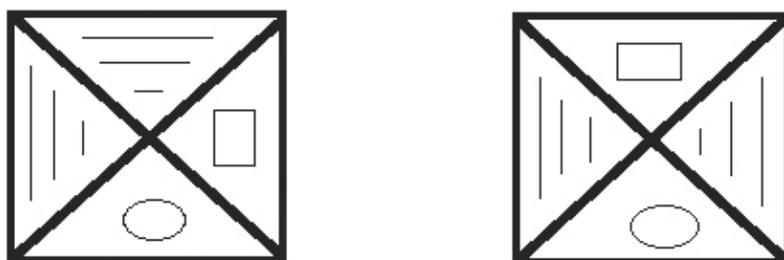


Figure 1. The two tiles have the same parts, but are not identical. Identity presupposes a certain disposition of the parts.

In chemistry, we know very well the concept of isomers. These are chemical compounds that have the same molecular formulas (thus they are composed by 'the same' atoms -more correctly, we should say that they are formed by *indistinguishable atoms*), but differ in their structural formulas. For instance, the molecular formula  $C_2H_6O$  may stand for both a molecule of ethylic alcohol and for a molecule of the methylic ether - see the figure below.

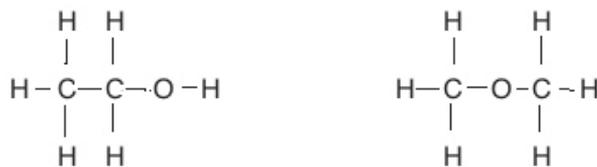


Figure 2. Structural molecules of the ethanol and of the dimethyl ether, both  $C_2H_6O$ .

So, what kind of axioms should we use? Let us suppose that L is as above, that is, the language of quasi-set theory, encompassing further primitive predicates P and F as before. The postulates are the following ones (others will appear soon):

1. The postulates of quasi-set theory (French & Krause [2010]).
2.  $Pst \ \& \ Pt_u \rightarrow Ps_u$ , being s and t variables of the same kind (this remark will be implicit from now on, and we shall comment on the kinds of variables only when absolutely necessary).
3.  $\exists s \in \alpha \rightarrow \exists s (s \Sigma \alpha \ \& \ (t \Sigma \alpha \rightarrow s = t))$ , where  $\equiv$  is the indistinguishability relation of quasi-set theory.

Hence, by the third axiom, which we call the Principle of Formation, if a given non empty collection of objects has a sum, and all sums are indistinguishable from one each other (notice that we are not saying that any collection of objects has a sum –see the last section). This is in agreement with quasi-set theory, where collections (quasi-sets) of indiscernible elements having the same cardinalities are indistinguishable (the relation of indistinguishability is an equivalence relation). When the considered entities are M-objects, the postulate (unicity of the sum) coincides with the defined notion of identity.

Our above discussion proposed that mereological extensionality should not be valid in general. Hence, we need to modify this assumption due to the possibility that objects may have their parts exchanged by indiscernible ones. In addition, this departure from extensional mereology should be motivated also by the fact that for those objects denoted by the first kind variables (m-objects), the concept of identity proper should be substituted by that of indiscernibility.

Since an object cannot be seen as formed by its parts, how can we identify it? Our answer follows Schrödinger in regarding quantum entities: we identify them by their *form*. His passage, in *Science and Humanism*, is quite beautiful and deserves reading:

“Finally we may observe that any palpable object in our environment is composed of atoms, which are composed of ultimate particles... and if the latter lack individuality, how does, say, my wrist-watch come by individuality? Where is the limit? How does individuality arise at all in objects composed of non-individuals?

“It is useful to consider this question in some detail, for it will give us the clue to what a particle or an atom really is –what there is permanent in it in spite of its lack of individuality. On my writing-table at home I have an iron letter-weight in the shape of a Great Dane, lying with his paws crossed in front of him. I have known it form many years. (...) It accompanied me to many places, until it stayed in Graz in 1938, when I had to leave in something of a hurry. But a friend of mine knew that I liked it so she took it and kept it for me. (...)

“I am quite sure it is the same dog (...). But *why* am I sure of it? That is quite obvious. It is clearly the peculiar *form* or *shape* (German: *Gestalt*) that rises the identity beyond doubt, not the material content. Had the material be melted and cast into the shape of a man, the identity would be much more difficult to establish. And what is more: even if the material identity were established beyond doubt, it would be of restricted interest. I should probably not care very much about the identity or not of the mass of iron, and should declare that my souvenir had been destroyed.” (Schrödinger [1996], pp.122-3).

This quotation motivates a further discussion on the relationship (if any) between the notions of form and structure. Really, perhaps we could say that we can represent the form of an object by means of a mathematical structure. For instance a  $H_2O$  molecule is composed by two H atoms and one O atom linked in a certain angle ( $104,45^\circ$ ); that is, it is not simple a collection (set) of two Hydrogen atoms and one Oxygen atom (really, there is more, but the details do not import here), but they must be *disposed* in a certain way. In a set theoretical context, by its turn, a structure is a n-tuple composed by one or more sets (the domains of the structure —we can suppose that there is just one domain if we take the union of the existing ones and enlarge the set of relations of the structure with adequate predicates telling us to which particular domain the objects belong) and a collection of relations whose relata are the objects of the domain (we reduce distinguished objects and operation symbols to relations in the standard way). Thus, in the case of the  $H_2O$  molecule, the domain comprises the involved atoms and the relations must be so that some of them give the necessary tools for its chemical and structural properties. If we can generalize this situation for more sophisticated objects, such as human beings, then perhaps we can identify our primitive predicate F with mathematical structures. But this would not entail that the involved objects are ‘pure structures’ as some form of Platonism could require or as the defenders of the ontological structural realism guess (Ladyman [2009]). In our opinion, a form is a form of something and, similarly, a structure structures something. Even in quantum mechanics, we feel the needs of associating the states of the systems, which are considered by the formalisms, with the *things* there are in those states. So, although we think that Schrödinger is right in posing that the material content is not enough, the characterization of the involved objects cannot be ruled out from the philosophical considerations; as Peter Simons says, “it is impossible to isolate mereological from more general ontological issues” (Simons [1987], p.210).

Given that F is a primitive predicates symbol in our approach, we need to find axioms that involve some notion of *form*. In considering our primitive binary predicate F, we interpret  $Fst$  as saying that s and t have the same form. Thus, our axioms might be (remember that we are dropping the universal quantifiers)

4.  $(Pus \rightarrow \exists w(Pwt \ \& \ w \equiv u)) \ \& \ (Put \rightarrow \exists w(Pws \ \& \ w \equiv u)) \ \& \ Fst \rightarrow s \equiv t$
5.  $(PZX \leftrightarrow PZY) \ \& \ (PXZ \leftrightarrow PYZ) \ \& \ FXY \rightarrow X=Y.$

Other axioms of F are

6. Fss
7. Fst  $\rightarrow$  Fts
8. Fst & Ftu  $\rightarrow$  Fsu

At axiom 4, the conjunct Fst is essential. In fact, it is not enough that s and t have the indiscernible parts, as explained hitherto. The axiom says that in order to be indiscernible, s and t must also share their *form*. But concerning M-objects, we may say simply that if two of them have the same parts (remember that identity holds here) *and* form, then they are identical. This is asserted by axiom 5. Furthermore, F is a partial order.

Other defined concepts are the following ones:

Definition 2.

- (a) [*Proper Part*] PPst := Pst &  $\sim$ Pts
- (b) [*Overlap*] Ost :=  $\exists w$ (Pws & Pwt)
- (c) [*Disjointness*] Dst :=  $\sim$ Ost
- (d) [*Entanglement*] Est :=  $\exists w$ (Psw & Ptw) &  $\sim \exists u$ (Puw & (Dus & Dut))
- (e) [*Gunky Object*] Gs := PPts  $\rightarrow \exists w$ (PPwt)
- (f) [*Universe*] U := is $\Sigma$ [t: t $\equiv$ t], where i is the symbol for definite description, and [t : At] stands for the quasi-set whose elements satisfy the condition A. This definition ought to be understood as metalinguistic, for we have not introduced descriptions (but we could have done of course).
- (g) [*Atoms*]  $\Delta s$  :=  $\sim \exists t$ PPts

The first three definitions don't present problems. Concerning (d), it is motivated by the entanglement of quantum systems. Again speaking superficially, this notion, which has no counterpart in classical physics, describes the state of certain systems that are formed by two or more other systems, but so that these 'parts' are superposed in such a way that no measurement in one of these compound systems can be made without creating an influence in the others. They form wholes (even when the parts are separated by several light years of distance) that would be counted as one. Thus, definition 2(d) says that s and t are entangled iff there exists an object (w) from which both s and t are parts and no part of such an object is disjoint from both s and t.

Definition (e) enables us to consider *gunky* entities, that is, objects whose parts have proper parts. Concerning the possibility of gunk-worlds, see Williams [2006]. If we suspect that Schrödinger (see again the beginnings of his above quotation) and many others agree in that perhaps there are not ultimate particles, then it seems that any mereology suitable to cope with quantum mechanics would incorporate gunks.

Definition (f) introduces an object summing all objects, the Universe. As Tarski ([1983a], p.333n) remarked, the formal difference between mereology and

the extended system of Boolean algebra is that in mereology there is no individual corresponding to the Boolean zero (an individual which is part of every other individual). Definition (g) poses us some problems in regarding quantum physics. It says that an atomic element is an object that has no proper parts.

Theorem 1. In the above mereology, the following results obtain.

- (a) Every object is part of the Universe: for any  $s$ ,  $PsU$
- (b) No objects is a proper part of itself: for any  $s$ ,  $\sim PP_s s$
- (c) Two M-objects are identical iff they are part of one another:  $X=Y \leftrightarrow (PXY \& PYX)$

Informal Argumentation: (a) and (b) are immediate consequences of the definitions of universe and proper part. (c) results from postulate 5 and from the underlying logic (classical logic holds for M-objects).

Theorem 2. The following results *don't obtain*.

- (a) If  $x$  and  $y$  are m-objects, then  $x=y \leftrightarrow \forall z(Pzx \leftrightarrow Pzy)$
- (b) In the same vein,  $x \equiv y \leftrightarrow \forall z(Pzx \leftrightarrow Pzy)$
- (c)  $As \& t \equiv s \rightarrow At$ , where  $A$  is a formula having free occurrences of  $s$  and  $At$  results from  $As$  by substituting  $t$  in some free occurrences of  $s$ .

Informal Argumentation: (a) doesn't hold since identity does not hold for m-objects; (b) is the same proposition with  $\equiv$  for  $=$ . It doesn't hold because (informally speaking) indistinguishable objects (even being m-objects) not necessarily have the same parts. For instance, two molecules of water are formed by two Hydrogen atoms and by one Oxygen atom. But of course the atoms, which are parts of these molecules, are not the same atoms. Recall that the notion of sameness presupposes identity, and in standard mathematics, if objects count as being *two*, they are *distinct*.

Of course much should be said about these issues, and in the continuation of this discussion, our system needs improvements in the postulates and in proving the theorems, but these tasks will be postponed to another work, since (I hope) the divulgation of these ideas may incentivate other researches to contribute to the field. In the same vein, we could speak of the applications of such a mereology to quantum theory. This is a topic to make us think.

#### 4. Some further improvements

Of course the above development is still in its crude form. But I think that the divulgation of these ideas may provide me some feedback and improve the discussion. But, at least with respect to the questions posed earlier, let me say a couple of words.

It seems clear that in a quantum mereology that what can be called ‘the principle of formation’ of standard mereologies, as in our axiom 3 above (I mean, extensional mereologies) cannot hold. Roughly speaking, this principle says that for any collection of things, there is at least one object having these things as its parts. But collections are usually formalized as sets, and it would be difficult to accept that a collection containing, say, the number two, the Empire State building and the greatest fish of Lock Ness could be taken as a set having these elements as parts, for they would be abstract entities, and the Empire State Building of course is not. Furthermore, if in accepting that objects may have parts, it would be difficult to say which parts are to be regarded as *essential* parts, without which an object leaves to be *that* object (notice that this involves a theory of identity for these objects). The Tibbles’ tail is an example. Perhaps there is no criterion for distinguishing essential from *accidental* parts, but we would agree that for any object there are some parts that can be exchanged or even eliminated without changing the identity of the object. It seems to me that the problem is with respect to the concept of identity. Perhaps this question can be dealt with by assuming time and refusing to accept an identity theory for the involved objects (at least in the classical form; but recall that if we change the theory, we change the concept). An object may change its parts, as we do during our time life (say, our skin) but during short periods of time, we remain indiscernible from ourselves. Probably I am indistinguishable from myself a week ago, but for sure I am completely different from myself thirty years ago! An old friend of mine who didn’t see me for 30 years probably will not recognize me in the street.

My guess, to be developed in another paper, is to develop a mereology considering time and without identity: the objects, for short periods of time (what ‘short’ means is something to be decided –perhaps it depends on the individual itself: a short period of time for a star is a long period for us, etc.) an object continues to be indiscernible from itself, and we identify that individual by resemblance or, perhaps as Hume could say, by habit.<sup>10</sup> An alternative theory would be to discern between distinguishability and identity: an object remains identical to itself during its lifetime, but can be distinguished from itself in different times during this interval. But this entails the need to characterize the concepts of identity (and difference), and of distinguishability (and indistinguishability) independently one each other, and we really still don’t know how to do that, for indistinguishability implies identity by the Principle of the Identity of Indiscernibles. Of course there are alternatives, and quasi-set theory offers one, but this is a research program to be investigated further.

Finally, some words about atomic objects (definition 2(g)). It seems reasonable to suppose that they have a sum, so that they may be parts of some other object. Let us see the reverse problem. Usually, if two objects can be cut into identical atomic pieces, then these individuals are identical. The problem with this definition in regarding quantum mechanics is in what sense ‘identical’ is to be considered. Really, in quantum mechanics (in most interpretations) there are objects, e.g. entangled systems, which cannot be said to have ‘proper parts’

before measurement (the problem is complicated and depends on several factors; according to the mainstream interpretation, the Copenhagen view –which by its turn cannot be summarized in just one point of view for Bohr and Heisenberg, for instance, didn't agree in all the details). But, *when* a measurement is made, the state of the system is cut in one of its 'proper parts', as is well known, and in a certain sense we can speak of the parts and they may be of course atomic.

I think that in order to deal with these difficulties, we need to consider a different kind of individual, one suitable to cope with intuitions coming from quantum mechanics. Informally, let us call *entangled objects* those objects whose states are described by entangled functions. From a first point of view, entangled objects are atomic, for we cannot speak within the formalism about their constituent parts. But, from another point of view, there are certain operations holding on these objects (standing for measurements) which show them as composed by proper parts. We really don't know if we need to incorporate this concept in our formalism. By the moment, we shall keep by mentioning the idea only.

The second point is that in 'cutting' objects until arriving at atoms, in order to say that the objects are equal we need to use the notion of identity. Suppose we have the above mentioned molecules of ethanol and dimethyl ether,  $C_2H_6O$  have being decomposed in atomic elements. Suppose (for keep our argument simple), that the atoms in which we can decompose the two molecules are two Carbon atoms, six Hydrogen atoms and one Oxygen atom. Are these atomic elements of the same kind *identical*? Of course the problem of identity transfers to the constituent parts, and so on.

## References

- Artzenius, F. and Hawthorne, J. [2005], Gunk and continuous variation. *The Monist* 88 (4), 441-66.
- Domenech, G., Holik, F. and Krause, D. [2008], Q-spaces and the foundations of quantum mechanics, *Foundations of Physics* 38 (11), pp. 969-94.
- Domenech, G., Holik, F., Kniznik, L., Krause, D. [2010], No labeling quantum mechanics of indiscernible particles, *International Journal of Theoretical Physics* 49 (12), pp. 3085-91.
- French, S. and Krause, D. [2006], *Identity in Physics: A Historical, Philosophical, and Formal Analysis*. Oxford: Oxford Un. Press.
- French, S. and Krause, D. [2010], Remarks on the theory of quasi-sets. *Studia Logica* 95 (1-2), 101-24.
- Goodman, N. [1966], *The Structure of Appearance*. 2<sup>nd</sup>. Ed., Bobbs-Merrill.
- Haslanger, S. [2003], Persistence through time. In M.J. Loux and D. W. Zimmerman (eds.) *The Oxford Handbook of Metaphysics*. Oxford: Oxford Un. Press, pp. 315-54.

- Hawthorne, J. [2003], Identity, in in M.J. Loux and D. W. Zimmerman (eds.) *The Oxford Handbook of Metaphysics*. Oxford: Oxford Un. Press, pp.99-130.
- Healey, R. [2008], Holism and nonseparability in physics, at <http://plato.stanford.edu/entries/physics-holism>
- Healey, R. [2009], Holism in quantum mechanics. In Greenberg, D., Hentschel, K and Weinert, F. (eds.), *Compendium of Quantum Physics*, Berlin and Heidelberg: Springer, pp. 295-8.
- Ladyman, J. [2009], Structural realism. At <http://plato.stanford.edu/entries/structural-realism/>
- Leonard, H.S. and Goodman, N. [1940], The calculus of individuals and its uses. *Journal of Symbolic Logic* 5 (2), 45-55.
- Mahan, B. H. [1972], *Química: um Curso Universitário*. São Paulo: Edgard Blücher (translation from *University Chemistry*, Addison Wesley Pu. Co., 2<sup>nd</sup> ed., 1969).
- Nascimento, M. C., Krause, D. and Feitosa, H. de A. [2011], ‘The quasi-lattice of indiscernible elements’, *Studia Logica* 97, pp. 101-28.
- Rescher, N. [1987], Mereology. *The New Encyclopædia Britannica*, 15<sup>th</sup> edition, Vol. 23, p.288.
- Schrödinger, E. [1996], *Nature and the Greeks and Science and Humanism*. Cambridge: Cambridge Un. Press.
- Simons, P. [1987], *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- Tarski, A. [1983], Foundations of the geometry of solids. In Tarski, A. *Logic, Semantics, Metamathematics: papers from 1923 to 1938*. 2<sup>nd</sup>.ed., Indianapolis, Indiana: Hackett Pu. Co., pp. 24-9.
- Tarski, A. [1983a], On the foundations of Boolean algebra. In Tarski, A. *Logic, Semantics, Metamathematics: papers from 1923 to 1938*. 2<sup>nd</sup>.ed., Indianapolis, Indiana: Hackett Pu. Co., pp. 320-41.
- Varzi, A. [2011], Mereology. *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition, <<http://plato.stanford.edu/archives/spr2011/entries/mereology/>>.
- Williams, J.R.G. [2006], Illusions of gunks, *Philosophical Perspectives* 20 (Metaphysics 2006), pp. 493-513.

## Notes

<sup>1</sup> According to Simons [1987], p.54, the use of sets makes no substantial difference to the expressive power of the theory. The problem is that upto day we know several non equivalent set theories, and this should be considered in seeing if there are differences among mereologies based on such distinct theories. But we shall not be concerned with this topic here.

<sup>2</sup> According to Simons, there is a difference between mereologies and calculi os individuals. The former “excludes formal theories formulated in a different logical style” (Simons [1987], p.6). Here we follow this author in using the term ‘mereology’ in a broad sense. The first calculus of individuals of Leonard & Goodman [1940] also made use of set theory. Concerning the mentioned approach by Goodman, see his [1966], pp. 44 ff.

<sup>3</sup> This suggests that in a quantum mereology, an individual, when this concept makes sense (see below), cannot be identified with the sum of its parts, yet there are mereological sums even so. See also Simons [1987], p.210.

<sup>4</sup> See French & Krause [2006], chap.6; a philosophical discussion about this concept can be seen also in Hawthorne [2003]. Summing up, the standard theory of identity may be assumed to be formed by the usual first order axioms of equality (reflexivity and substitutivity) and by the axiom of extensionality. Alternatively, in higher order languages we can define identity by means of the Leibniz Law. These theories are not equivalent, but can be considered as representing identity according to 'classical' standards.

<sup>5</sup> The careful reader may be recalled that we are speaking indifferently of quantum objects and their states (usually described by rays in some Hilbert space). This overlapping of concepts does not conduce to problems if we assume that the states are states of something we wish to consider here.

<sup>6</sup> We have discussed the non-individuallity of quantum objects from other points of view: quasi-set theory deals with collections of non-individuals; Schrödinger logics discuss this notion from a type-theory like approach (see French & Krause [2006], chap.8); the characterization of a quasi-lattice of indiscernible elements was described in Nascimento et.al. [2010]; the first steps in the creation of a non-reflexive quantum mechanics, that is, a quantum mechanics founded in quasi-set theory, can be seen in Domenech et al. [2009], and in Domenech et al. [2010]. The relationships of this concept with sortal predication and with vague objects are discussed in French & Krause op.cit. and in the papers mentioned there.

<sup>7</sup> In the level  $2p$ , the magnetic number  $m$  may assume the values  $-1, 0, +1$ , while the spin  $s$  may be  $+1/2$  or  $-1/2$ , which give us six possible combinations, exhausted for instance in a sodium atom (Mahan [1972], p.321).

<sup>8</sup> This definition coincides with that of the system LG of Simons (op.cit., p.53), and with Rescher's mereology (Rescher [1987], p.288). Tarski's definition may be written as follows:  $s\Sigma\alpha := (t \in s \rightarrow Pts) \ \& \ \sim \exists w (Pws \ \& \ u \in \alpha \rightarrow Dwu)$  (Tarski [1983], p.25); intuitively speaking, every element of  $\alpha$  is a part of  $s$  and no part of  $s$  is disjoint from all elements of  $\alpha$ .

<sup>9</sup> To conform with the discussion that follows, perhaps we should take the exemple in the following way. Instead of two tiles, take one tile as in the text, cut its parts and join then again in another disposition. Then, the parts will be in fact the same parts, but the tiles are of course not identical.

<sup>10</sup> But this will push us to consider time. As we know, in standard, that is, orthodox quantum mechanics, time and space are 'classical', Newtonian, but this is not so in the relativistic theories. Which notion of time would we consider?

# DISCUSSÕES ACERCA DA VERSÃO FRACA DO PRINCÍPIO DA IDENTIDADE DOS INDISCERNÍVEIS EM MECÂNICA QUÂNTICA

JONAS RAFAEL BECKER ARENHART

*Universidade Federal da Fronteira Sul, Campus Chapecó – Santa Catarina, Departamento de Filosofia*

jonas.arenhart@uffs.edu.br

**Resumo:** Neste trabalho trataremos do problema da validade do Princípio da Identidade dos Indiscerníveis na mecânica quântica. Há atualmente um impasse na disputa concernindo seu status: alguns sustentam que o princípio é válido, outros que é inválido. Ambos os lados da disputa chamam a mecânica quântica em seu auxílio, argumentando que a teoria corrobora sua posição. Argumentaremos que uma das grandes dificuldades neste caso repousa precisamente em se supor que a mecânica quântica pode nos dar elementos decisivos para resolver este problema.

**Palavras-chave:** Identidade dos Indiscerníveis; Discernibilidade; Mecânica quântica; Metafísica.

## 1. Introdução

Segundo o Princípio da Identidade dos Indiscerníveis (PII), se os objetos  $x$  e  $y$  possuem todas as propriedades em comum, então são um e o mesmo objeto. Dito de outra forma, objetos numericamente distintos são qualitativamente distintos. Assim, se PII for válido, sempre que tratamos com *dois* objetos, podemos ter a certeza de que existe uma qualidade que os discerne, mesmo que não sejamos capazes de apontar para nenhuma tal qualidade explicitamente.

A primeira grande dificuldade nas discussões acerca do status deste princípio concerne precisamente o termo “qualidades”. O que deve ser visto como uma qualidade discernidora? Apenas propriedades? Ou será que relações também podem de algum modo discernir *qualitativamente*? É aqui que nossas escolhas se mostram decisivas para praticamente todo o debate atual em torno do status do PII. Se formos liberais e aceitarmos que relações também discernem, então em alguns contextos o princípio será válido. Se formos mais exigentes e entendermos que apenas se pode discernir qualitativamente através de propriedades, então o princípio é na maioria das vezes visto como falso.

Estes pontos, de acordo com o modo como vemos o problema, são cruciais se desejamos que a discussão possa ser conduzida com o mínimo de influência de pressupostos metafísicos dos quais não estejamos conscientes. Tornar

explícitas todas as hipóteses que assumimos (como, por exemplo, que estamos assumindo que relações discernem qualitativamente) pode evitar que alguns equívocos sejam cometidos e que expressemos compromisso com algumas proposições de status duvidoso que de outra maneira soariam infundadas. Um dos mais recentes destes equívocos, em nossa opinião, tratando justamente da validade do PII, concerne sua aplicação à mecânica quântica não-relativista.

Segundo muitos autores, dentre eles alguns célebres físicos que ajudaram a desenvolver a mecânica quântica, o PII é um princípio inválido no contexto desta teoria. De fato, as partículas quânticas seriam tais que apesar de serem numericamente distintas, possuem em algumas situações todas as suas qualidades em comum (ver French e Krause 2006 cap. 4). Assim, seria possível refutar o princípio apelando para a mecânica quântica. Mais interessante do que isto, talvez, seja o próprio fato de que apelando para uma teoria científica poderíamos verificar a validade ou invalidade de um princípio metafísico, algo que muitos filósofos têm buscado recentemente.

Outro modo de ver este problema, no entanto, surgiu recentemente, apontando na direção contrária. Segundo alguns autores, o PII é de fato válido na mecânica quântica, contrariando o que se acreditou ser o caso até pouco tempo atrás (ver Muller e Saunders 2008, Muller e Seevinck 2009). O fato crucial para se mostrar que o PII pode ser salvo na mecânica quântica, segundo estes autores, foi o esquecimento, por parte dos detratores do PII, de que relações também podem discernir. Assim, mesmo que partículas quânticas não possam ser discernidas por propriedades e nem por certos tipos de relações, existem alguns tipos especiais de relações que sempre podem ser empregados para discernir as partículas (ver a seção sobre Terminologia). Logo, a mecânica quântica vingaria o PII. Do mesmo modo que ocorreu quando se tratava de provar que PII é inválido na mecânica quântica, este fato foi visto como uma grande vitória de uma particular abordagem à metafísica, que a restringe apenas aos contextos em que se pode investigar o status de determinados problemas metafísicos em teorias científicas.

Como devemos entender a situação que surge com o antagonismo destas duas posições e seu apelo para a mesma teoria? É aqui que nossa proposta de explicitar todas as nossas hipóteses *metafísicas* sendo utilizadas em nossos argumentos entra em jogo. Como mencionamos brevemente, salvar o PII na mecânica quântica (seguindo a proposta de Muller, Saunders e Seevinck nas obras mencionadas acima) envolve a pressuposição de que relações podem discernir qualitativamente dois itens. Isto, todavia, implica em aceitar uma versão do PII que não está de acordo com aquilo que a maioria dos filósofos toma como sendo a versão mais relevante e filosoficamente interessante do PII. Assim, “salvar” o PII no contexto da mecânica quântica tem um alto preço: a versão que estamos salvando pode nem mesmo ser considerada como uma versão do PII de fato, mas apenas um princípio menos interessante, sem as principais características desejadas quando se trata de propor e defender este princípio.

A escolha parece então ficar entre, por um lado, um princípio inócuo salvo e por outro, um princípio metafisicamente robusto, mas refutado.

Neste trabalho, argumentaremos que este tipo de explicitação de nossas hipóteses pode servir para tornar ainda mais clara a relação entre a proposta metafísica destes autores, qual seja, que a metafísica deve ser naturalizada, tornada sempre relativa a alguma teoria científica, e a ideia de que há uma metafísica não-naturalizada. Argumentaremos que existem certos limites para o que se pode estabelecer em contextos como aqueles de uma teoria científica, e que hipóteses controversas, como, por exemplo, aquelas acerca da legitimidade de relações para discernir (e consequentemente a validade de PII) não podem ser decididas no contexto de teorias científicas, mas antes devem ser mantidas na metafísica “clássica”, entendida como uma investigação de primeiros princípios e de questões que de modo algum estão no escopo das teorias científicas. Utilizando a discussão acerca da validade do PII em mecânica quântica como nosso modelo e fio condutor, argumentaremos que o uso de relações para salvar o PII não está livre de controvérsias de tipo metafísico que não podem ser ignoradas quando se trata de um princípio metafísico.

Começamos na próxima seção introduzindo mais alguns pontos relevantes para a nossa discussão.

## 2. O PII

O famoso *Princípio da Identidade dos Indiscerníveis* nos garante que objetos qualitativamente indiscerníveis são de fato apenas um e o mesmo objeto. Se desejarmos uma forma rigorosa de expressar este enunciado, podemos empregar uma linguagem de segunda ordem para tentarmos captar uma versão formal do princípio do seguinte modo:

$$(F)(Fa \leftrightarrow Fb) \rightarrow a = b.$$

Aqui,  $F$  varia sobre propriedades e  $a$  e  $b$  denotam objetos particulares. Também é instrutivo expressar o princípio em sua forma contrapositiva: dados quaisquer dois objetos *numericamente* distintos, sempre existe alguma *qualidade* que faça a distinção, que serve para discerni-los. Assim, PII pode ser visto como colapsando fatos acerca da diferença numérica em fatos acerca da diferença qualitativa, algo aparentemente desejável em muitos contextos.

Uma das principais situações em que PII é muito importante e através da qual podemos melhor compreender a função metafísica desempenhada por este princípio, surge quando o associamos a uma teoria da individualidade de objetos particulares, como, por exemplo, no caso das chamadas *teorias de feixes de propriedades*. Segundo a teoria de feixes, em linhas gerais, um objeto particular é um feixe de propriedades, e sua individualidade é de certo modo explicada pela unicidade de cada um de tais feixes. Como garantir tal unicidade? É aqui que o PII entra em ação, sendo empregado para garantir que se dois feixes

coincidem em todas as suas qualidades, então são numericamente idênticos. Devemos notar que este último passo é crucial: se não garantimos que cada feixe ocorre apenas uma vez, então a teoria de feixes não desempenha seu papel de fornecer um *princípio de individualidade*, como se deseja ao se adotar esta teoria. Em geral, ataques à individuação através de feixes de propriedades são feitos precisamente através de ataques à validade do PII, um dos princípios fundamentais da teoria de feixes. E como discutiremos, não foram poucos os ataques desta natureza.

Daquilo que acabamos de dizer, fica claro que o PII possui grande relevância para o problema da individuação, através da teoria dos feixes. Independentemente desta teoria, o PII é fundamental também ao se buscar estabelecer uma forma de entendermos diferença quantitativa em termos de diferença qualitativa. No entanto, apesar destas duas importantes funções, o princípio sempre despertou muitas suspeitas, e sua validade foi frequentemente colocada em questão. De fato, é notável que apesar de ter sido defendido por eminentes filósofos como Leibniz e Bertrand Russell, o princípio da identidade dos indiscerníveis de fato nunca foi uma unanimidade, tendo gerado enorme controvérsia na história da filosofia. O grande problema envolvido com o princípio diz respeito precisamente ao fato de que aparentemente podemos conceber situações envolvendo objetos numericamente distintos mas qualitativamente indiscerníveis, como por exemplo no caso do famoso universo sugerido por Max Black (em Black 1952). Black sugeriu que podemos conceber um universo contendo apenas duas esferas de metal, a duas milhas de distância uma da outra, possuindo todas as mesmas propriedades, como o diâmetro, peso, cor, etc. Nestas circunstâncias, sugere Black, temos *duas* esferas possuindo todas as suas qualidades em comum, em clara violação do que nos prega o PII.

O universo de Black constituiu durante muito tempo um dos mais famosos argumentos questionando a validade do PII. Como todo argumento controverso, foi bastante discutido e sua aceitação ainda é objeto de muita disputa. Mas mesmo que muitos filósofos tenham expressado dúvidas acerca de seu sucesso, novos contraexemplos começaram a ser produzidos e iniciou-se uma enorme disputa acerca da própria plausibilidade destes exemplos e do efeito deles sobre o status do princípio. Diante destes argumentos, alguns filósofos passaram a sugerir que, se fosse verdadeiro, PII seria na melhor das hipóteses apenas *contingentemente* verdadeiro, ou seja, verdadeiro no mundo atual (algo de que poucos duvidavam), mas não verdadeiro em todos os mundos possíveis. Isto mudou repentinamente no começo do século 20, como veremos.

Aparentemente, não deveríamos nos surpreender com o fato de que um princípio metafísico tenha despertado tantas controvérsias, afinal, é da própria natureza dos problemas metafísicos e dos princípios neles envolvidos não poderem apelar diretamente aos fatos para buscar evidência a seu favor. Isto é, até pouco tempo apenas um reduzido número de filósofos sustentaria que uma doutrina metafísica ou um princípio metafísico pode ser avaliado de modo definitivo à luz dos conhecimentos obtidos através de uma teoria científica.

Todavia, este passo foi dado precisamente no caso da controvérsia envolvendo PII, e esta é uma atitude que vem se tornando comum entre os filósofos da ciência. A teoria utilizada em questão foi a mecânica quântica não relativista, e o princípio teve sua validade avaliada relativamente a esta teoria.

Vejamos como isto ocorreu.

### 3. PII na mecânica quântica: válido ou não?

A situação nesta controvérsia começou a tomar novo rumo no século 20 com o surgimento da mecânica quântica. Durante os desenvolvimentos da teoria, a partir de 1925 e 1926 em diante, muitos dos próprios físicos envolvidos na elaboração da teoria começaram gradualmente a sugerir que os objetos com os quais a mecânica quântica trata são tais que devem ser considerados como sendo absolutamente indistinguíveis, apesar de serem numericamente distintos (ver French e Krause 2006 cap. 3). Nestas circunstâncias, o PII não seria nem mesmo contingentemente verdadeiro, mas simplesmente falso, precisamente por encontrarmos contraexemplos para o princípio no mundo atual (ver a discussão em French 1989, French e Krause 2006 cap. 4). Claro, muitos consideraram que seria um grande avanço para a discussão o fato de que uma de nossas melhores teorias científicas nos forneça os contraexemplos para um princípio metafísico. Esta circunstância foi vista por alguns filósofos de orientação naturalista como sugerindo que a metafísica finalmente deixou para trás seus aspectos puramente especulativos e começou a voltar-se para a ciência, tornando-se assim uma disciplina muito mais respeitável aos seus olhos (para uma defesa de uma posição radical favorecendo este tipo de abordagem, ver Ladyman e Ross 2007. Para uma defesa da metafísica como investigação racional, ver Lowe 2011).

Recentemente, todavia, o debate sofreu um novo revés, com alguns filósofos sugerindo que a situação é precisamente a oposta do que se pensou até pouco tempo atrás: de acordo com alguns, o PII é de fato *válido* na mecânica quântica, é um teorema no formalismo matemático da teoria (Saunders 2006, Muller e Saunders 2008, Muller e Seevinck 2009). Segundo estes autores, a mecânica quântica *vinga* o princípio. O que se esqueceu nos debates envolvendo o PII e a mecânica quântica, segundo eles, é que relações também podem ser empregadas para discernir objetos. Assim, dados objetos numericamente distintos, basta encontrarmos relações que sirvam para discernir adequadamente e teremos que o princípio é válido. Ainda mais, segundo estes autores, a mecânica quântica sempre nos permite encontrar tais relações valendo entre os objetos com os quais ela trata. Deste modo, para objetos quânticos a identidade dos indiscerníveis deveria ser um princípio válido.

Novamente, este passo foi visto como uma vitória de uma particular abordagem científica à metafísica sobre as discussões de cunho puramente especulativo: enquanto que os filósofos especulativos em geral não são capazes

de decidir de modo completamente satisfatório a disputa acerca da validade ou invalidade do PII, se nos restringirmos a uma abordagem naturalista e considerarmos apenas teorias científicas, então podemos mostrar, segundo estes autores, que o PII é de fato válido na mecânica quântica. Segundo este modo de ver, esta pode ser tomada como forte evidência contra a metafísica especulativa e a favor de uma *naturalização da metafísica*, mesmo que não esteja completamente claro o que esta última envolveria. O ponto central seria que as teorias científicas devem ser consultadas quando se trata de disputas filosóficas, e estas disputas devem ser resolvidas no contexto das teorias científicas, já que elas são as melhores candidatas a nos dizer como o mundo de fato é na concepção de um naturalista (ou, talvez, moderando o tom excessivamente realista desta afirmação, as teorias científicas chegam muito mais perto de algo que poderíamos chamar de “verdade” acerca do mundo empírico do que as teorias metafísicas).

Todavia, no que diz respeito ao princípio que estamos discutindo, parece que chegamos a uma situação de grande impasse, com alguns autores afirmando que o PII é falso na mecânica quântica e outros sustentando a tese oposta. Para complicar ainda mais a situação, devemos mencionar um aspecto particular deste caso que dificulta ainda mais a questão: ambos os lados buscam apoio para suas afirmações na própria mecânica quântica, de modo que a teoria parece servir tanto para defender a posição de que o princípio está refutado quanto a posição de que ele é válido. Assumindo uma posição naturalista acerca da metafísica, como podemos avaliar melhor o progresso da discussão? É um princípio válido ou não?

Talvez a própria ideia de que podemos provar o princípio em uma teoria científica deva ser revisada. Nas próximas seções deste trabalho abordaremos o problema de um ponto de vista mais metafísico. Discutiremos a própria possibilidade de se provar no contexto de *teorias científicas* como a mecânica quântica um *princípio metafísico* como o PII. Argumentaremos que a controvérsia não pode, neste caso, ser resolvida de modo tão simples como querem os defensores do PII. De fato, eles devem reconhecer que hipóteses metafísicas preliminares foram assumidas em seus argumentos, hipóteses que de modo algum serão aceitas por todos e em particular por aqueles que não estavam dispostos a aceitar que o PII é um princípio válido na mecânica quântica. Neste caso, a hipótese crucial foi a de que relações podem realmente ser empregadas para distinguir qualitativamente dois objetos.

Realmente, o argumento proposto para salvar o PII no contexto da mecânica quântica faz uso de certa liberdade interpretativa: o que devemos entender por “qualidades” quando dizemos que dois itens partilham todas as suas qualidades? Devemos restringir o nosso entendimento das qualidades apenas a propriedades monádicas? Ou podemos ser mais liberais e admitir que as relações nas quais um item toma parte também podem estar contando como qualidades? De nossas respostas a estas questões dependerão o sucesso ou fracasso da defesa do PII na mecânica quântica. Realmente, argumentaremos que aceitar que a mecânica quântica prova o PII permitindo que relações figurem como qualidades pressupõe

que se aceite que o uso de relações para *distinguir* é legítimo. Como veremos, este não é o caso para a maioria dos filósofos interessados no status do PII. Mais do que isto, o problema de se aceitar ou rejeitar relações como possuindo a capacidade de discernir é o tipo de questão que a metafísica de tipo naturalista não consegue resolver (a mecânica quântica por si só não nos ajuda a resolver este tipo de problema). Na próxima seção apresentaremos parte da terminologia na qual o debate é conduzido.

#### 4. Terminologia

Vamos introduzir agora a terminologia importante para que se possa discutir alguns aspectos da questão. Seguiremos aqui o trabalho de Muller e Saunders (2008, pp. 520 - 521). Vamos supor que temos uma linguagem de primeira ordem com uma relação binária de pertinência  $\in$ , de modo que temos disponível uma teoria de conjuntos como Zermelo-Fraenkel e com ela toda a matemática necessária para se erigir o formalismo da mecânica quântica ortodoxa. Vamos supor que temos termos especiais (variáveis e constantes individuais) para objetos físicos, estados de objetos físicos e grandezas físicas. Os estados puros de sistemas físicos, como é usual, são representados por vetores unitários em um espaço de Hilbert  $H$ . As grandezas físicas mensuráveis são representadas por operadores hermitianos nestes espaços. Se um sistema físico está em um autoestado de um operador hermitiano, dizemos que ele possui a propriedade associada àquele operador, e seu valor é o autovalor correspondente.

Agora, vamos supor que um sistema físico  $a$  possui uma determinada propriedade física  $P$ . Representamos este fato em nossa linguagem através da fórmula  $Pa$ . Dados sistemas físicos distintos  $a$  e  $b$ , dizemos que são *absolutamente discerníveis* se existe uma propriedade  $P$  tal que ou  $Pa$  e não  $Pb$  ou então não  $Pa$  e ainda assim  $Pb$ . Neste caso, quando existe uma tal propriedade, escrevemos simplesmente  $Abs(a,b)$ . Se, por outro lado, existir uma relação que vale entre os mencionados sistemas físicos em apenas uma ordem possível, ou seja, se existir uma  $R$  tal que ou  $R(a,b)$  e não  $R(b,a)$  ou então tal que  $R(b,a)$  e não  $R(a,b)$ , então neste caso dizemos que  $a$  e  $b$  são *relacionalmente discerníveis*. Ainda, se existir uma relação  $R$  simétrica e irreflexiva, ou seja, tal que tanto  $R(a,b)$  e  $R(b,a)$  mas não  $R(a,a)$  e nem  $R(b,b)$ , então dizemos que  $a$  e  $b$  são *fracamente discerníveis*. Se  $a$  e  $b$  são relacionalmente discerníveis ou fracamente discerníveis, descrevemos este fato como  $Rel(a,b)$ . Tendo esta terminologia em mãos, Muller e Saunders **definem** o Princípio da Identidade dos Indiscerníveis da seguinte forma:

$$a \neq b \rightarrow (Abs(a,b) \text{ ou } Rel(a,b)).$$

Assim, se dois objetos forem distintos, segundo PII nesta versão, devem ser ou absolutamente discerníveis, relacionalmente discerníveis ou fracamente discerníveis. De fato, Muller e Saunders 2008 argumentam que apenas  $Rel(a,b)$  é

realmente relevante para a validade do PII. Vejamos brevemente seu argumento. Em primeiro lugar, se  $a$  e  $b$  são absolutamente discerníveis, então, existe uma propriedade  $P$  que as distingue. Para ver que neste caso vale  $Rel(a,b)$ , basta definir uma relação  $R$  entre  $a$  e  $b$  como “( $Pa$  e não  $Pb$ ) ou ( $Pb$  e não  $Pa$ )”. Assim,  $Abs(a,b)$  implica  $Rel(a,b)$ . Por definição, PII vale se e somente para itens  $a$  e  $b$  distintos temos  $Abs(a,b)$  ou  $Rel(a,b)$ . Agora, pelo argumento anterior, se temos  $Abs(a,b)$  temos  $Rel(a,b)$ . Com isto, o que realmente importa, segundo Muller e Saunders, é a discernibilidade relacional. Para provar que PII é inválido, é preciso provar que se  $a$  e  $b$  são distintos, então não temos  $Rel(a,b)$ .

De fato, o argumento destes autores (em Saunders 2006, Muller e Saunders 2008, e Muller e Seevinck 2009) consiste em se demonstrar que na mecânica quântica sempre podemos encontrar relações que façam o papel de discernir qualitativamente. A discernibilidade absoluta e a discernibilidade relacional estão fora de questão, como os detratores do PII queriam, pois objetos quânticos não podem ser discernidos por propriedades ou pela ordem em que entram em relações. No entanto, os inimigos do PII se esqueceram da discernibilidade fraca. Sempre podemos prover relações simétricas e irreflexivas valendo entre partículas quânticas. O exemplo clássico é o sistema de dois elétrons no estado *singlete*, com a relação “... tem spin oposto a ...” quando uma determinada orientação espacial é escolhida. Claro, nenhum elétron tem spin oposto a si mesmo, e se um elétron  $a$  tem spin oposto a  $b$ , então  $b$  também possui spin oposto a  $a$ . Isto, claro, é o suficiente para se estabelecer a validade de PII na mecânica quântica. Mas, note-se, estamos falando de PII conforme ele foi definido por Muller e Saunders, permitindo-se o uso de relações para discernir.

O que dizer da noção de individualidade através de feixes que geralmente está associada ao PII? Muller e Saunders (2008, p. 504) definem um *indivíduo* como um objeto que é absolutamente discernível de todos os outros, ou seja, que possui uma propriedade monádica que o discerne de todos os outros objetos. E o que podemos dizer de partículas como elétrons, que são absolutamente indiscerníveis, mas podem ser fracamente discerníveis? Elas são indivíduos? Seguindo Muller e Saunders, que aderem consistentemente à terminologia definida por eles, não podemos sustentar que partículas quânticas são indivíduos. Partículas quânticas como elétrons, e de fato todos os outros objetos que são apenas relacionalmente discerníveis são chamados de *relacionais*. Relacionais não são indivíduos, mas obedecem ao PII conforme definido por Muller e Saunders.

Assim, temos como consequência de aceitarmos diferentes graus de discernibilidade, uma ontologia compreendendo pelo menos dois tipos de entidades: indivíduos e relacionais. Ambos satisfazem o PII. Note que aqui nos afastamos de uma das principais motivações para se defender o PII: a teoria de feixes. De fato, agora, se aceitamos os argumentos de Muller e Saunders, podemos ter objetos satisfazendo o PII sem que sejam indivíduos, algo que contraria o espírito da teoria de feixes. Todavia, apenas garantimos que o PII é válido na mecânica quântica se aceitamos que as relações podem realente discernir. Mas

podemos perguntar: este passo é legítimo? Relações realmente discernem? Muitos filósofos pensam que a resposta a esta questão é negativa. Vejamos brevemente algumas de suas razões.

## 5. Discernibilidade e relações

Conforme ficou claro pelas definições apresentadas na seção anterior, a própria definição do PII que Muller e Saunders desejam salvar envolve a aceitação de relações para efeitos de discernir qualitativamente objetos. No caso, trata-se de relações simétricas e irreflexivas. Mas é legítimo aceitarmos que relações façam uma distinção qualitativa entre objetos do mesmo modo que propriedades monádicas o fazem, ou será que devemos rejeitar esta proposta? Existem fortes razões para se negar que relações tenham de fato a capacidade de discernir objetos.

O primeiro ponto a ser considerado consiste precisamente na natureza das relações. Em geral, relações são consideradas como sendo, de certa maneira, “externas” aos objetos. Os objetos são geralmente compreendidos como “entrando” em relações, e para tanto, devem estar ontologicamente disponíveis antes de se relacionarem. Assim, se desejamos fundamentar alguma forma de distinção qualitativa entre  $a$  e  $b$  através de uma relação irreflexiva e antissimétrica  $R$ , devemos primeiro saber que é  $a$  que está relacionado por  $R$  com  $b$ , ou seja, devemos ter alguma forma prévia de discernir os objetos para garantir que são precisamente *estes objetos* que estão se relacionando. Assim, discernir não é algo que possa ser feito pela relação, mas algo que já está pressuposto pela relação (ver Hawley 2009, Ladyman e Bigaj 2010).

Este argumento, claro, baseia-se em uma particular concepção acerca da natureza das relações e do modo como objetos particulares se relacionam, que, todavia é bastante natural e muito difundida entre os filósofos. Seu ponto de ataque é bastante claro: como o objetivo do PII é reduzir afirmações de diferença numérica ( $a \neq b$ ) a fatos acerca de diferenças qualitativas, ou seja, dar uma fundamentação da primeira noção em termos da segunda, temos que uma utilização de relações para este trabalho fracassa precisamente por pressupor aquilo que desejamos obter com a adoção do princípio em foco. Alguma forma de discernir os objetos  $a$  e  $b$  já deve ser dada de antemão para que saibamos que são eles que estão sendo relacionados por  $R$ .

Outro ponto particularmente complicado para os defensores da discernibilidade fraca como uma maneira legítima de se discernir objetos consiste no fato de que nenhuma diferença específica é apresentada neste caso (ver Ladyman e Bigaj 2010). De fato, sustentarmos que objetos  $a$  e  $b$  estão sendo discernidos mas não apresentarmos nenhuma diferença entre eles parece simplesmente uma extensão indevida da própria noção do que seja “discernir” dois objetos. Neste caso, o fato de relações simétricas e irreflexivas estarem sendo usadas torna o problema ainda mais complicado. Vamos considerar o caso do

objeto  $a$ . Se a única forma de discerni-lo de  $b$  é através de  $R$ , uma relação de discernibilidade fraca, como podemos ter certeza que estamos de fato tomando  $a$  e o discernindo de  $b$  através da relação  $R$ ? De alguma forma, como no argumento anterior, precisamos de  $a$  discernido de antemão, para somente depois podermos garantir que este objeto em particular está relacionado com  $b$ .

Uma possível solução para estas dificuldades poderia consistir em se adotar algum princípio não-qualitativo que garanta a diferença numérica, mas não a diferença qualitativa, se por diferença qualitativa entendermos apenas a distinção feita através de propriedades monádicas. Princípios como o substrato de Locke ou as essências primitivas, de Adams, poderiam desempenhar este papel (ver Adams 1979, Moreland 1998). Com o uso destes expedientes metafísicos, podemos garantir pelo menos de um ponto de vista ontológico, que  $a$  e  $b$  são distintos e que não dependem de modo algum da relação que deveria os distinguir. Todavia, o principal objetivo de se propor um princípio como o PII, e de se buscar tão arduamente salvá-lo é justamente o fato de que ele idealmente nos permite dispensar estes tipos de princípios metafísicos misteriosos. Assim, se devemos recorrer a eles, estamos de certa forma regredindo na discussão, e abandonando as motivações para se salvar o PII.

Agora, considerando o caso que nenhuma diferença específica possa ser apresentada entre  $a$  e  $b$ , como se requer, o que se ganhou com a adoção do PII na forma particular proposta por Muller e Saunders? Aparentemente, nada daquilo que se esperava do PII, e dificilmente poderíamos dizer que a versão de Muller e Saunders do PII realiza a tarefa à qual o princípio se propõe, que é precisamente dar conta da diferença numérica em termos de alguma diferença qualitativa. Além disso, como já comentamos anteriormente, ela não consegue estar associada a uma teoria de feixes da individualidade, de modo que as partículas quânticas, apesar de obedecerem ao PII, não são indivíduos. Assim, o campo fica aberto para podermos expressar mais algumas dúvidas acerca do sucesso da defesa proposta por Muller e Saunders, como veremos na próxima seção.

## 6. Metafísica naturalizada

Assim, chegamos a um impasse digno de nota. Por um lado, conforme mencionamos brevemente anteriormente, temos muitos filósofos afirmando que PII é falso em mecânica quântica, e que a própria teoria endossa esta posição (French e Krause 2006, cap. 4, Ladyman e Bigaj 2010). Estes filósofos tem em mente PII formulado da seguinte forma:

$$a \neq b \rightarrow Abs(a,b)$$

Neste caso, a mecânica quântica seria tal que podemos, com seus recursos, provar que dois objetos podem sim partilhar todas as suas qualidades monádicas. Assim, PII nesta versão é falso.

Por outro lado, temos autores como Muller, Saunders e Seevinck defendendo a conclusão oposta: PII é válido na mecânica quântica, e com os recursos da teoria podemos provar que o princípio é válido. Todavia, como vimos, a versão do PII empregada por Muller e Saunders pressupõe que relações também podem ser empregadas para discernir. Isto, como os argumentos acima mostram, não é algo que se possa tomar como universalmente aceito e não passível de discussões adicionais. De fato, o uso de relações para discernir requer uma defesa apropriada que ainda não foi apresentada pelos defensores da discernibilidade fraca. Assim, parece que encontramos um ponto em que a discussão chega a um limite, pelo menos do ponto de vista da metafísica puramente naturalizada.

O que devemos fazer nesta situação? Ora, devemos reconhecer que se nos limitarmos apenas ao que a mecânica quântica pode nos oferecer, temos duas situações possíveis. Em primeiro lugar, PII encarado da maneira tradicional, permitindo que objetos distintos sejam discernidos qualitativamente apenas por propriedades monádicas é falso. Em segundo lugar, podemos encontrar relações simétricas e irreflexivas valendo entre sistemas  $a$  e  $b$  que são tais que não é o caso que  $Abs(a,b)$ . Estas relações, claro, são fornecidas pelo próprio formalismo da mecânica quântica, sendo então totalmente autorizadas pela teoria. Todavia, se estas relações discernem qualitativamente ou não é algo que a própria teoria não nos informa.

Como vimos, existem vários argumentos *metafísicos*, oriundos em disputas da metafísica tradicional, buscando garantir que relações não discernem. No entanto, a mecânica quântica não desempenha nenhum papel preponderante nestas discussões, que são conduzidas independentemente desta teoria. Assim, tentar reduzir o problema apenas ao que a mecânica quântica pode nos informar nos deixará inevitavelmente em um impasse: PII é válido, desde que aceitemos que relações discernem, e é inválido em caso contrário. Mas a questão que sempre retorna é: relações discernem? Este é o ponto em que a discussão precisa voltar aos seus aspectos eminentemente metafísicos, pois a mecânica quântica não foi feita para resolver este tipo de problemas, trata-se de terreno próprio da metafísica.

Talvez, então, este seja um bom indício de que a metafísica deve sim buscar apoio nas ciências, e em particular na física, mas que, todavia, não devemos nos esquecer de que estas ciências não estão formuladas com o propósito específico de resolver alguns problemas que são puramente metafísicos. De fato, existe uma ampla gama de problemas de interesse metafísico, mesmo para o estudo da metafísica relacionada às teorias científicas, mas que não podem ser investigados apenas com os recursos destas teorias. Acerca destes assuntos, não podemos nos contentar apenas com aquilo que nos informam as teorias científicas, pois em geral elas não foram feitas para tratar destes problemas, e de modo ainda mais geral, pode haver mais de uma resposta possível para tais questões quando apelamos apenas ao que nos dizem as teorias científicas.

## 7. Conclusão

Neste trabalho estivemos tratando da validade de um importante princípio metafísico, o PII, concentrando nossos argumentos no contexto da mecânica quântica. Como vimos, a validade ou invalidade do princípio foi muitas vezes empregada para se sustentar que a metafísica deveria mudar seu foco e se concentrar apenas em teorias científicas, de modo que seus resultados poderiam ser então mais confiáveis e livres das disputas que desde sempre povoaram a metafísica.

Aqui argumentamos que o caso do PII é um exemplo claro de que este tipo de posição não pode ser levado aos seus extremos. Como mostramos anteriormente, a validade ou invalidade de PII dependerá de determinadas preferências ou pressuposições que não podem ser justificadas apenas ao nível da mecânica quântica. Assim, as disputas metafísicas, como é o caso da validade do PII, não podem ser resolvidas satisfatoriamente se nos restringirmos apenas ao escopo da mecânica quântica.

De fato, argumentamos que o contrário é que é verdadeiro. Se levarmos a sério a questão acerca da possibilidade de discernirmos através de relações, veremos que existe forte evidência metafísica de que isto não deveria ser aceito, e assim, apesar de podermos demonstrar no formalismo da mecânica quântica que o PII conforme definido por Muller e Saunders é uma consequência do formalismo, esta não deve ser encarada como uma versão legítima do PII. Assim, o princípio em questão não se encontra de modo algum vingado e estabelecido na mecânica quântica, mas antes o contrário é que parece evidencia-se: o princípio encontra-se sob as mesmas suspeitas de invalidade sob as quais sempre esteve, talvez, agora mais do que nunca.

## Referências

- Adams, R., Primitive Thisness and Primitive Identity, *Journal of Philosophy* **76**, pp. 5-26, 1979.
- Black, M., 1952. The Identity of Indiscernibles. *Mind* vol. **61**, no. 242, pp. 153-164.
- French, S. 1989. Why the Principle of Identity of Indiscernibles is not Contingently True Either. *Synthese* **78**, pp. 141-166.
- French, S., e Krause, D., 2006. *Identity in Physics. A historical, philosophical and formal analysis*. Oxford: Oxford University Press.
- Hawley, K., Identity and Indiscernibility, *Mind* **118**, pp. 101-119, 2009.
- Ladyman, J. e Ross, D., 2007, *Everything Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press."

Ladyman, J., Bigaj, T., The Principle of the Identity of Indiscernibles and Quantum Mechanics, *Philosophy of Science* **77**, pp. 117-136, 2010.

Lowe, E. J., The Rationality of Metaphysics, *Synthese*, **178**(1), pp. 99-109, 2011.

Moreland, J.P., Theories of Individuation: A Reconsideration of Bare Particulars, *Pacific Philosophical Quarterly* **79**, pp. 51-63, 1998.

Muller, F.A., e Saunders, S., 2008. Discerning Fermions. *British Journal for the Philosophy of Science* **59**, pp. 499-548.

Muller, F.A., e Seevinck, P. 2009. Discerning Elementary Particles. *Philosophy of Science*.

Saunders, S., 2006. Are Quantum Particles Objects?, *Analysis* **66**, pp. 52-63.

# ON THE PHYSICAL REPRESENTATION OF QUANTUM SUPERPOSITIONS

NEWTON DA COSTA

*Universidade Federal de Santa Catarina*

ncacosta@terra.com.br

CHRISTIAN DE RONDE

*Instituto de Filosofia “A. Korn” – UBA-CONICET*

*Centre Leo Apostel (CLEA) and Foundations of the Exact Sciences (FUND)*

cderonde@vub.ac.be

## 1 Introduction

Physical superpositions exist both in classical and in quantum physics. However, what is exactly meant by “superposition” in each case is extremely different. In classical physics one can have superpositions of waves or fields. A wave (field)  $\alpha$  can be added to a different wave (field)  $\beta$  and the sum will give a ‘new’ wave (field)  $\mu = \alpha + \beta$ . There is in this case no weirdness for the sum of multiple states gives as a result a new single state. In quantum mechanics on the contrary, the sum of states cannot be reduced to one single state. There is no obvious interpretation of such superposition of states. Today, quantum superpositions play a central role in the most interesting technical developments such as quantum teleportation, quantum cryptography and quantum computation (Leibfried *et al.*, 2005; Ourjoumtsev *et al.*, 2007). The question we attempt to address in this paper regards the meaning of quantum superpositions and their physical representation. There are many interpretations of quantum mechanics each of which provides an answer to this question. In the following we shall review some of these proposals. We shall argue in favor of the importance of developing a new interpretation of quantum superpositions in terms of paraconsistent logic.

Paraconsistent logics are the logics of inconsistent but nontrivial theories. The origins of paraconsistent logics go back to the first systematic studies dealing with the possibility of rejecting the principle of noncontradiction. Paraconsistent logic was elaborated, independently, by S. Jaskowski in Poland, and by the first author of this paper in Brazil, around the middle of the last century (on paraconsistent logic, see, for example: da Costa, Krause and Bueno, 2007). A theory  $T$  founded on the logic  $L$ , which contains a symbol  $\neg$  for negation, is called inconsistent if it has among its theorems a sentence  $A$  and its negation  $\neg A$ ; otherwise, it is said to be consistent.  $T$  is called trivial if any sentence of its language is also a theorem of  $T$ ; otherwise,  $T$  is said to be non-trivial. In classical logics and

in most usual logics, a theory is inconsistent if, and only if, it is trivial.  $L$  is paraconsistent when it can be the underlying logic of inconsistent but non trivial theories. Clearly, no classical logic is paraconsistent. The importance of paraconsistent logic is not limited to the realm of pure logic but has been extended to many fields of application such as robot control, air traffic control (Nakamatsu *et al.*, 2002), control systems for autonomous machines, defeasible deontic reasoning (Nakamatsu *et al.*, 2001), information systems (Akama and Abe, 2001) and medicine. In the following, we attempt to call the attention to the importance of extending the realm of paraconsistent logic to the formal account of quantum superpositions. We shall discuss the very different meaning of the term ‘superposition’ in both classical and quantum physics. In section 3, we shall present different interpretations of the quantum superposition such as those provided by van Fraassen’s empiricist interpretation, Bohmian mechanics, Many Worlds, and the Geneva approach to quantum logic. In section 4, we shall argue in favor of the importance of considering an interpretation of superposition in terms of paraconsistent logic.

## 2 What is a Quantum Superposition?

In classical physics, every physical system may be described exclusively by means of its *actual properties*, taking ‘actuality’ as expressing the *preexistent* mode of being of the properties themselves, independently of observation –the ‘pre’ referring to its existence previous to measurement. Each system has a determined state characterized mathematically in terms of a point in phase space. The change of the system may be described by the change of its actual properties. Potential or possible properties are considered as the points to which the system might arrive in a future instant of time. As noticed by Dieks:

“In classical physics the most fundamental description of a physical system (a point in phase space) reflects only the actual, and nothing that is merely possible. It is true that sometimes states involving probabilities occur in classical physics: think of the probability distributions  $\bar{n}$  in statistical mechanics. But the occurrence of possibilities in such cases merely reflects our ignorance about what is actual. The statistical states do not correspond to features of the actual system (unlike the case of the quantum mechanical superpositions), but quantify our lack of knowledge of those actual features.” D. Dieks (2010, p. 125)

Classical mechanics tells us via the equation of motion how the state of the system moves along the curve determined by initial conditions in the phase space. The representation of the state of the physical system is given by a point in phase space  $\Gamma$  and the physical magnitudes are represented by real functions over  $\Gamma$ . These functions commute between each other and can be interpreted as possessing definite values independently of physical observation, i.e. each

magnitude can be interpreted as being actually preexistent to any possible measurement. In the orthodox formulation of quantum mechanics, the representation of the state of a system is given by a ray in Hilbert space  $\mathcal{H}$ . But, contrary to the classical scheme, physical magnitudes are represented by operators on  $\mathcal{H}$  that, in general, do not commute. This mathematical fact has extremely problematic interpretational consequences for it is then difficult to affirm that these quantum magnitudes are *simultaneously preexistent*. In order to restrict the discourse to different sets of commuting magnitudes, different Complete Sets of Commuting Operators (CSCO) have to be chosen. The choice of a particular representation (given by a CSCO) determines the basis in which the observables diagonalize and in which the ray can be expressed. Thus, the ray can be written as different linear combinations of states:

$$\alpha_i | \varphi_i^{B1} \rangle + \alpha_j | \varphi_j^{B1} \rangle = \alpha_q | \varphi_q^{B2} \rangle = \beta_m | \varphi_m^{B3} \rangle + \beta_n | \varphi_n^{B3} \rangle + \beta_o | \varphi_o^{B3} \rangle \quad (1)$$

The linear combinations of states are also called quantum superpositions.

“The nature of the relationships which the superposition principle requires to exist between the states of any system is of a kind that cannot be explained in terms of familiar physical concepts. One cannot in the classical sense picture a system being partly in each of two states and see the equivalence of this to the system being completely in some other state. There is an entirely new idea involved, to which one must get accustomed and in terms of which one must proceed to build up an exact mathematical theory, without having any detailed classical picture.” P. Dirac (1974, p. 12)

The formal difference of using vectors in  $\mathcal{H}$  instead of points in  $\Gamma$  seems to imply that in quantum mechanics —apart from the ‘possibility’ which is encountered in classical mechanics— there is another, different realm which must be necessarily considered and refers, at each instant of time, to *contradictory properties*. To see this, consider the following example: given a spin 1/2 system whose state is  $|\uparrow_z\rangle$ , we let it interact with a magnetic field in the  $z$ -direction. All outcomes that can become actual in the future are potential properties of the system, in an analogous manner as all possible reachable positions of a pendulum are in the classical case. But at each instant of time, for example at the initial instant, if we consider the  $z$ -direction and the projection operator  $|\uparrow_z\rangle\langle\uparrow_z|$  as representing a preexistent actual property, there are other incompatible properties arising from considering projection operators of spin projections in other directions. For example, in the  $x$ -direction, the projection operators  $|\uparrow_x\rangle\langle\uparrow_x|$  and  $|\downarrow_x\rangle\langle\downarrow_x|$  do not commute with  $|\uparrow_z\rangle\langle\uparrow_z|$  and thus, cannot be considered to possess definite values simultaneously. Since Born interpretation of the wave function, these properties are usually considered as *possible*. However, this possibility is essentially different from the idea of possibility discussed in classical physics which relates to the idea of a *process*. If we consider that the formalism of quantum mechanics provides a description of the world, a representation of *what there is* —and does not merely make reference to

measurement outcomes—, at each instant of time the properties,  $|\uparrow_z\rangle\langle\uparrow_z|$ ,  $|\uparrow_x\rangle\langle\uparrow_x|$  and  $|\downarrow_x\rangle\langle\downarrow_x|$  must be taken into account for they all provide non trivial information about the state of affairs. In particular, the properties  $|\uparrow_x\rangle\langle\uparrow_x|$  and  $|\downarrow_x\rangle\langle\downarrow_x|$ , which constitute the superposition are *contradictory properties*.

In the quantum logic approach one of the properties, namely, the one in which we can write the state of affairs as a single term, is considered as ‘actual’ while the others are taken to be ‘potential’ properties. Potential properties can become ‘actual’ in a future instant of time according to the probability given by the numbers in square modulus accompanying the states. These properties, e.g.  $|\uparrow_x\rangle\langle\uparrow_x|$  and  $|\downarrow_x\rangle\langle\downarrow_x|$ ,  $|\uparrow_y\rangle\langle\uparrow_y|$  and  $|\downarrow_y\rangle\langle\downarrow_y|$  in our example, are always part of superpositions with more than one term and are constituted by contradictory properties. However, from a mathematical perspective, independently of their mode of existence, both potential and actual properties are placed at the same level in the algebraic frame which describes the state of affairs according to quantum mechanics: the projections of the spin in all directions are atoms of the lattice and there is no formal priority of the actual over the potential properties. In the laboratory, it is precisely this contradictory potential realm which is necessary to be considered by the experimentalist in the developments which are taking place today regarding the processing of quantum information as quantum computing and quantum communication. This seems to point in the direction that these properties have an existence which cannot be reduced to their ‘becoming actual at a future instant of time’. As noticed by Dieks (2010, p. 120): “[...] the superposition correspond to different possible indications of the measuring device (different ‘pointer positions’); that is, to different possible outcomes of the measurement. Clearly, all possible outcomes occur on an equal footing in the superposition of the final state, so that there is no sign that any one of them is more real than any other.” There are many approaches which attempt to interpret quantum superpositions. We shall discuss in the following section some of these proposals.

### 3 The Multiple Interpretations of Quantum Superpositions

As we have seen above, the formal description of quantum mechanics seems to imply a deep departure from the classical notion of possible or probable. However, there are many different interpretations of quantum mechanics in general and of the meaning of a quantum superposition in particular. In the following section we shall review some of the possible interpretations of quantum superpositions already present in the literature.

Taking into account the remark of Dieks (1988, p. 189), regarding the superposition of the Schrödinger cat (in the states ‘dead’ and ‘alive’), that: “It is the state vector which is in a superposition, not the cat itself. ‘State vector’ and ‘cat’ are two concepts at different levels of discourse.” From a general realist

perspective, which considers physics as providing an expression of the world, the question arises if this formal or mathematical representation given by superpositions, namely equation (1) —which allow us to calculate the probability of the possible measurement outcomes—, can be related conceptually to a notion which can allow us to think, independently of measurement outcomes, about the ‘superposition of states in Hilbert space’ in an analogous manner as we think of a ‘point in phase space’ (in the formal level) as describing an ‘object in space-time’ (in the conceptual level). What is describing a mathematical superposition? Can we create or find adequate concepts which can provide a representational realistic account of a quantum superposition independent of measurement outcomes? As we shall also see, from a general empiricist perspective one is not committed to answering these set of questions.

### 3.1 Quantum Superpositions as a ‘Theoretical Device to Predict Measurement Outcomes’

The idea that the quantum wave function as related to a superposition is just a theoretical device with no ontological content goes back to Bohr’s interpretation of quantum mechanics. The impossibility to interpret the quantum wave function in an ontological fashion can be understood in relation to his characterization of  $\Psi$  in terms of an algorithmic device which computes measurement results.<sup>1</sup> This position radically addressed seems to end up in the instrumentalistic account shared implicitly by many and developed explicitly by Fuchs and Peres (2000). Bas van Fraassen, whom we consider a close follower of Bohr’s ideas, has also taken an anti-metaphysical position with respect to the interpretation of the quantum wave function. His justification stands on his empiricist account of both physics and philosophy (van Fraassen, 1991; section 9.1). In tune with van Fraassen, Dieks argues in one of his latest papers in favor of a Humean position:

“The Humean maintains that we need to assume the existence of only one world, namely the ordinary actual one; that the regularities of this world are expressed in our laws and theories; and that we introduce possible other worlds and counterfactual circumstances purely as thought constructions, in order to bring out the peculiarities of the laws we have formulated. Possible worlds are mental tools and not really existing entities. Modalities, like necessity and possibility, are concepts we introduce on the basis of our theories and do not correspond to features of reality that transcend the ordinary description in terms of actual events.” D. Dieks (2010, p. 126)

From an empiricist perspective the formalism does not need to provide a description of *what there is*. Superpositions can be considered thus as a theoretical device through which one can account for the probabilities of each actual observation. Empiricism can be linked to probability in terms of the frequency interpretation which rests, contrary to the original conception of probability, not in the idea that probability describes in terms of ignorance an existent state

of affairs, but rather in the set of empirical results found in a series of repeated measurements. However, and independently of the problems encountered within such empiricist stances, if superpositions are considered just as a theoretical device, then the question of interpretation seems to lose its importance. For why should we pursue an interpretation if, like Fuchs and Peres remark, quantum mechanics does the job and already “provides an algorithm for computing probabilities for the macroscopic events”? There are other reasons which one could put forward from an empiricist perspective –as even van Fraassen (1980) argues–, however they will remain only secondary in the quest of science.

### 3.2 Quantum Superpositions as ‘Describing Quantum Fields’

As noticed by Bacciagaluppi (1996, p. 74), the hidden variable program attempts to “restore a classical way of thinking about *what there is*.” In this sense, Bohm’s proposal turns quantum mechanics into a theory which restores the possibility of discussing in terms of a definite state of affairs (described in terms of a set of definite valued properties). In Bohmian mechanics the state of a system is given by the wave function  $\Psi$  together with the configuration of particles. The quantum wave function must be understood in analogy to a classical field that moves the particles in accordance with the following functional relation:  $dx/dt = \nabla S$ , where  $S = h\delta$  ( $\delta$  being the phase of  $\Psi$ ). Thus, particles always have a well defined position together with the rest of their properties and the evolution depends on the quantum field. It then follows that, there are no superpositions of states, the superposition is given only at the level of the field and remains as mysterious as the superposition of classical fields. Given a quantum field  $\phi(x)$  the particle will move according to it. If we change the quantum field by adding another field  $\Psi(x)$  such that the new quantum field is now the superposition:  $\phi(x) + \Psi(x)$ , there is no ontological peculiarity involved for now the particle also has a well defined position and will evolve according to the new field. Presumably, due to the fact that the new field is different from the original one the particle will move in a different way and will follow a different trajectory compared to the first case. The field does not only have a dynamical character but also determines the epistemic probability of the configuration of particles via the usual Born rule.

### 3.3 Quantum Superpositions as ‘Describing Multiple Worlds’

Many worlds interpretations are no-collapse interpretations which respect the orthodox formulation of quantum mechanics. The many worlds interpretation is considered to be a direct conclusion from Everett’s first proposal in terms of ‘relative states’ (Everett, 1957). Everett’s idea was to let quantum mechanics find its own interpretation, making justice to the symmetries inherent in the Hilbert space formalism in a simple and convincing way. The main idea behind many worlds interpretations is that superpositions relate to collections of worlds, in each of which exactly one value of an observable, which corresponds

to one of the terms in the superposition, is realized (DeWitt and Graham, 1973). Apart from being simple, the claim is that it possesses a natural fit to the formalism, respecting its symmetries. The solution proposed to the measurement problem is provided by assuming that each one of the terms in the superposition is actual in its own correspondent world.

“The whole issue of the transition from ‘possible’ to ‘actual’ is taken care of in the theory in a very simple way —there is no such transition, nor is such a transition necessary for the theory to be in accord with our experience. *From the viewpoint of the theory all elements of a superposition (all ‘branches’) are ‘actual’, none any more ‘real’ than the rest.* It is unnecessary to suppose that all but one are somehow destroyed, since all the separate elements of a superposition individually obey the wave equation with complete indifference to the presence or absence (‘actuality’ or not) of any other elements. This total lack of effect of one branch on another also implies that no observer will ever be aware of any ‘splitting’ process.” H. Everett (1973, pp. 146-147, emphasis added)

Thus, it is not only the single value which we see in ‘our world’ which gets actualized but rather, that a branching of worlds takes place in every measurement, giving rise to a multiplicity of worlds with their corresponding actual values. The possible splits of the worlds are determined by the laws of quantum mechanics but each world becomes a ‘classical world’.

### 3.4 Quantum Superpositions as ‘Describing Potential Measurement Results’

The Geneva school to quantum logic and similar approaches attempt to consider quantum physics as related to the realms of actuality and potentiality in analogous manner to classical physics. According to the Geneva school, both in classical and quantum physics measurements will provoke fundamental changes of the state of the system. What is special for a classical system is that ‘observables’ can be described by functions on the state space. This is the main reason that, a measurement corresponding to such an observable can be left out of the description of the theory ‘in case one is not interested in the change of state provoked by the measurement’ but ‘only interested in the values of the observables’. It is in this respect that the situation is very different for a quantum system. Observables can also be described, as projection valued measures on the Hilbert space, but ‘no definite values can be attributed to such a specific observable for a substantial part of the states of the system’. For a quantum system, contrary to a classical system, it is not true that ‘either a property or its negation is actual’.<sup>2</sup> Continuing Heisenberg’s considerations in the new physics, Constantin Piron has been one of the leading figures in developing the notion of potentiality within the logical structure of quantum mechanics (Piron, 1976; 1983). Following (Smets, 2005), a physical property, never mind whether a classical or quantum one, is specified as what corresponds to a set of definite experimental projects. A definite experimental project (DEP) is an experimental procedure (in fact, an equivalence class of experimental procedures) consisting in a list of actions and a rule that

specifies in advance what has to be considered as a positive result, in correspondence with the yes answer to a dichotomic question.

Each DEP tests a property. A given DEP is called certain (correspondingly, a dichotomic question is called true) if it is sure that the positive response would be obtained when the experiment is performed or, more precisely, in case that whenever the system is placed in a measurement situation then it produces certain definite phenomenon to happen. A physical property is called actual in case the DEPs which test it are certain and it is called potential otherwise. Whether a property is actual or potential depends on the state in which one considers the system to be. Though in this approach both actuality and potentiality are considered as modes of being, actual properties are considered as attributes that exist, as elements of physical reality, while potential properties are not conceived as existing in the same way as real ones. They are thought as possibilities with respect to actualization, because potential properties may be actualized due to some change in the state of the system. In this case the superposition provides a measure —given by the real numbers which appear in the same term as the state— over the irrational potential properties which could become actual in a given situation. Thus, potentiality, as in the classical physical sense, can be regarded as potentiality, as referring to a future actuality.

## 4 Quantum Superpositions and Paraconsistent Logics

Although the interpretations we have discussed in the previous section from both their formal and metaphysical commitments have many differences, there is still something they all share in common: they all attempt to avoid *contradictions*. Indeed ‘contradiction’ has been regarded with disbelief in Western thought due to certain metaphysical presuppositions which go back to Plato, Aristotle, Leibniz and Kant. Even after the development of paraconsistent logic in the mid XX century and its subsequent technical progress which this theory has allowed, this aversion towards contradiction is still present today. The famous statement of Popper that the acceptance of inconsistency “... would mean the complete breakdown of science” remains an unfortunate prejudice within present philosophy of science (see da Costa and French 2003, Chap. 5).

Leaving instrumentalist positions aside, one of us has argued elsewhere (de Ronde, 2010) that one can find in the vast literature regarding the interpretation of quantum mechanics, two main strategies which attempt to find an answer to the riddle of ‘what is quantum mechanics talking about’. The first strategy is to begin with a presupposed set of metaphysical principles and advance towards a new formalism. Examples of this strategy are Bohmian mechanics, which has been discussed above (section 3.2), or GRW (Ghirardi *et al.*, 1986), which introduces non-linear terms in the Schrödinger equation. The second strategy is to accept the orthodox formalism of quantum mechanics and advance towards

the creation and elucidation of the metaphysical principles which would allow us to consider ‘what is quantum mechanics talking about’? Examples of this second strategy are quantum logic and its different lines of development such as the Geneva School of Jauch and Piron (section 3.4) and the modal interpretation (see for example Dickson and Dieks 2002; Vermaas 1999; de Ronde 2011). From this perspective the importance is to focus in the formalism of the theory and try to learn about the symmetries, the logical features and structural relations. The idea is that, by learning about such aspects of the theory we can also develop the metaphysical conditions which should be taken into account in a coherent ontological interpretation of quantum mechanics.

Quantum computation makes use of the multiple flow of information in the superposition considering even (in principle) contradictory paths. Also quantum cryptography uses the relation between contradictory terms in order to send messages avoiding classical spies. At a formal level, the path integral approach also considers the multiple contradictory paths within two points (Feynman and Hibbs, 1965). Since both the formalism and experiments seem to consider ‘contradictory elements’ within quantum mechanics, we argue that it can be of deep interest to develop a formalism which takes contradiction into account “right from the start”.<sup>3</sup> Our proposal focuses on the idea that it would be worthwhile to develop a new interpretation of quantum superpositions in terms of paraconsistent logic. We leave it for a future paper to present an explicit formal scheme for quantum superpositions (da Costa and de Ronde, 2011). However, it should be clear that we do not take paraconsistent logic to be the “true logic” which should replace classical logic; in the same way as we do not regard quantum mechanics as a theory that should replace classical mechanics (da Costa and French 2003, de Ronde 2011). In this sense, physicist should recognize the possibility to use new forms of logic –such as paraconsistent logic– which might help us understanding features of different domains of reality; features which cannot be accommodated by means of classical logic. We do not believe there is a “true logic”, but rather that distinct logical systems can be of used to develop and understand complementary aspects of reality.

Recalling the words of Albert Einstein: “it is only the theory which can tell you what can be observed”<sup>4</sup> it could be argued that only within a theory it is possible to consider and account for phenomena. From this perspective the development of the formalism can be regarded not only as a merely technical improvement, but also as a way to open new paths of understanding and even developing new phenomena. Formal development is not understood here as going beyond the theory, as improving and showing something that “was not there before” in the formalism –as it is the case of GRW or Bohmian mechanics. Rather, this development is understood as taking seriously the features which the theory seems to show us, exposing them in all their strength, “right from the start”.

## References

- Aerts, D., 1999, "Quantum mechanics; structures, axioms and paradoxes", in *Quantum Structures and the Nature of Reality: The Indigo Book of the Einstein meets Margritte Series*, D. Aerts and J. Pykacz (Eds.), Kluwer Academic, Dordrecht.
- Akama, S. and Abe, J. M., 2001, 'The role of inconsistency in information systems', *Proc. Of the 5th World Multiconference on Systemics, Cybernetics and Informatics (SCI'2001)*, pp. 355-360, Orlando.
- Bacciagaluppi, G., 1996, *Topics in the Modal Interpretation of Quantum Mechanics*, Doctoral dissertation, University of Cambridge, Cambridge.
- da Costa, N. C. A. and French, S., 2003, *Partial Truth: a unitary approach to models and scientific reasoning*, Oxford University Press, Oxford.
- da Costa, N. C. A., Karuse, D., and Bueno, O., 'Paraconsistent logics and paraconsistency', in *Handbook of the Philosophy of Science (Philosophy of Logic)*, D. Jacquette editor, Elsevier, 2007, pp. 791-911.
- da Costa, N. C. A. and de Ronde, C., 2011, 'The Paraconsistent Logics of Quantum Superpositions', in preparation.
- DeWitt, B. and Graham, N., 1973, *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton.
- Dickson, M. and Dieks, D., 2002, "Modal Interpretations of Quantum Mechanics", *The Stanford Encyclopedia of Philosophy (Winter 2002 Edition)*, E. N. Zalta (Ed.), URL: <http://plato.stanford.edu/archives/win2002/entries/qm-modal/>.
- Dieks, D., 1988, "The Formalism of Quantum Theory: An Objective description of reality", *Annalen der Physik*, **7**, 174-190.
- Dieks, D., 2010, "Quantum Mechanics, Chance and Modality", *Philosophica*, **82**, 117-137.
- Dirac, P. A. M., 1974, *The Principles of Quantum Mechanics, 4th Edition*, Oxford University Press, London.
- Everett, H., 1957, "'Relative State' Formulation of Quantum Mechanics", *Reviews of Modern Physics*, **29**, 454-462.
- Everett, H., 1973, "The Theory of the Universal Wave Function", In *The Many-Worlds Interpretation of Quantum Mechanics*, DeWitt and Graham (Eds.), Princeton University Press, Princeton.
- Feynman, R. P. and Hibbs, A. R., 1965, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York.
- Fuchs, C. and Peres, A., 2000, "Quantum theory needs no 'interpretation'", *Physics Today*, **53**, 70.
- Ghirardi, G. C., Rimini A. and Weber, T., 1986, "Unified Dynamics for Microscopic and Macroscopic Systems", *Physical Review D*, **34**, 470-491.
- Haag, R., *Local Quantum Physics*, Springer, 1992.

- Heisenberg, W., 1927, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik", *Zeitschrift für Physik*, **43**, 172- 98; reprinted as "The Physical Content of Quantum Kinematics and Mechanics, translation by J.A. Wheeler and W.H. Zurek, in *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek (Eds.).
- Krause, D., 1992, "On a quasi-set theory", *Notre Dame Journal of Formal Logic*, **33**, 402-11.
- Leibfried, D., Knill, E., Seidelin, S., Britton, J., Blakestad, R. B., Chiaverini, J., Hume, D. B., Itano, W. M., Jost, J. D., Langer, C., Ozeri, R., Reichle R. and Wineland, D. J., 2005, "Creation of a six-atom 'Schrödinger cat' state", *Nature*, **438**, 639-642.
- Nakamatsu, K., Abe, J. M. and Suzuki, A., 2001, "Annotated semantics for defeasible deontic reasoning", in W. Ziarko and Y. Yao (eds.), *RCCTC 2000*, pp. 470-478, Springer Verlag, Berlin.
- Nakamatsu, K., Abe, J. M. and Suzuki, A., 2002, "A railway interlocking safety verification systems based on abductive paraconsistent logic programming", *Soft Computing Systems: Design, Management and Applications*, Eds. A. Abraham, J. Ruiz-del-Solar & M. Koppen, Frontiers in Artificial Intelligence and Its Applications, IOS Press, pp. 775-784, Amsterdam, Ohmsha, Tokyo.
- Nakamatsu, K., Abe, J. M. and Suzuki, A., 2002, 'Defeasible deontic robot control based on extended vector annotated logic programming', In *Computer Anticipatory Systems: SASYS 2001 - Fifth International Conference*, ed. D. M. Dubois, American Institute of Physics, 490-500.
- Ourjountsev, A., Jeong, H., Tualle-Brouri, R. and Grangier, P., 2007, "Generation of optical 'Schrödinger cats' from photon number states", *Nature*, **448**, 784-786.
- Piron, C., 1976, *Foundations of Quantum Physics*, W.A. Benjamin Inc., Massachusetts.
- Piron, C., 1983, "Le réalisme en physique quantique: une approche selon Aristote", In *The concept of physical reality*. Proceedings of a conference organized by the Interdisciplinary Research Group, University of Athens.
- Plantinga, A., 2003, *Essays in the Metaphysics of Modality*, Oxford University Press, Oxford.
- de Ronde, C., 2010, "For and Against Metaphysics in the Modal Interpretation of Quantum Mechanics", *Philosophica*, **83**, 85-117.
- de Ronde, C., 2011, *The Contextual and Modal Character of Quantum Mechanics: A Formal and Philosophical Analysis in the Foundations of Physics*, PhD dissertation, Utrecht University.
- Schrödinger, E., 1935, "The Present Situation in Quantum Mechanics", *Naturwiss*, **23**, 807. Translated to english in *Quantum Theory and Measurement*, J. A. Wheeler and W. H. Zurek (Eds.), 1983, Princeton University Press, Princeton.
- Smets, S., 2005, "The Modes of Physical Properties in the Logical Foundations of Physics", *Logic and Logical Philosophy*, **14**, 37-53.
- Van Fraassen, B. C., 1980, *The Scientific Image*, Clarendon, Oxford.
- Van Fraassen, B. C., 1991, *Quantum Mechanics: An Empiricist View*, Clarendon, Oxford.
- Von Weizsäcker, C. F., 1974, *La Imagen Física del Mundo*, Biblioteca de Autores Cristianos, Madrid.

## Notas

<sup>1</sup> According to Bohr (Von Weizsäcker, 1974, p. 338) the Schrödinger wave equation is just an abstract magnitude of calculus and it does not designate in itself any phenomena.

<sup>2</sup> Private discussion with Diederik Aerts, 2010.

<sup>3</sup> In an analogous fashion as Décio Krause has developed a Q-set theory which accounts for indistinguishable particles with a formal calculus “right from the start” (Krause, 1992).

<sup>4</sup> These words, according to Heisneberg himself, led him to the development of the *indetermination principle* in his foundational paper of 1927.



III

---

FILOSOFIA DA LINGUAGEM  
E FUNDAMENTOS DA  
LINGUÍSTICA



# A FUNÇÃO DE CÓDIGO DOS NOMES PRÓPRIOS <sup>1</sup>

CÍCERO ANTÔNIO CAVALCANTE BARROSO

*Universidade Federal do Ceará - Campus Cariri*

cicero@lia.ufc.br

## 1. Introdução

Na década de 70, Tyler Burge escreveu que podia parecer surpreendente que alguém teorizasse sobre nomes próprios (cf. BURGE, 1973: 425). Afinal, *prima facie*, não parece haver nenhuma complicação na forma como eles se relacionam com o mundo. Essa primeira impressão, contudo, se desvanece quando temos ciência do número de trabalhos dedicados à temática dos nomes próprios desde os primeiros dias da filosofia analítica até hoje. A temática foi devassada de uma forma tão ampla e minuciosa que fica difícil pensar em alguma questão pertinente ao seu domínio que não tenha sido objeto de reflexão. Em face disso, hoje, o que pode parecer surpreendente é que alguém ainda se atreva a dizer algo sobre os nomes próprios. Todavia, como se vê pelo título deste artigo, esse é precisamente o meu caso. Até onde for possível, gostaria de desfazer a ideia de que não há mais nada de novo que possa ser dito sobre os nomes próprios.

Na maioria dos estudos filosóficos tradicionais sobre nomes próprios, o que se considera mais importante é explicar questões que tenham a ver com a função referencial dos mesmos, ou seja, com a função que eles têm de referir coisas. Pode-se argumentar plausivelmente que esse interesse pelos nomes próprios é derivado de um interesse mais básico pela questão das condições de verdade dos enunciados da linguagem. Aparentemente, a análise das condições de verdade de um enunciado E com ocorrência de um nome próprio n requer uma explicação acerca da função referencial de n. Com efeito, o valor de verdade de E pode se alterar em decorrência do fato de n exercer ou não uma função referencial, e as condições de verdade de E também se alteram se a referência de n se altera. Assim, é preciso esclarecer o que acontece quando n exerce uma função referencial, que fatores possibilitam que a função seja exercida etc. A elucidação dessas questões é importante para termos uma compreensão adequada das condições de verdade dos enunciados com ocorrência de nomes próprios.

Ora, mostrar interesse pela questão da verdade dos enunciados equivale a mostrar interesse pela relação da linguagem com o mundo, pela função representacional da linguagem. Acontece que a linguagem não se relaciona só com o mundo. Também existe inegavelmente uma relação entre a linguagem e os seus usuários. As pessoas usam a linguagem para dar e receber informações, e essa característica da linguagem também me parece altamente digna da atenção

dos filósofos. É preciso explicá-la. E por que não dedicar também alguma atenção ao papel que os nomes próprios desempenham nesse intercâmbio de informações? Aliás, eles desempenham algum papel?

Creio que se tivermos em mente certa concepção de *informação*, teremos que reconhecer que sim. De fato, em vista dessa concepção, todos os elementos que usamos no processo de comunicação funcionam como códigos, ou seja, funcionam como veículos de informação. Logo de saída é preciso observar que há uma distinção entre *informação sobre algo* e *informação para algo*. A primeira ideia que podemos ter a respeito dessa distinção é a de que informação-sobre está presente, por exemplo, em enunciados declarativos, e que informação-para é transmitida por enunciados conativos (ordens, instruções etc). Minha intuição, porém, é de que toda informação é informação-para, mesmo a informação-sobre é um tipo de informação-para. Ela é um tipo de informação que deflagra operações de associações de conceitos, e são essas associações que nos permitem, por exemplo, julgar se um enunciado declarativo é verdadeiro ou falso. A minha ideia é de que nomes próprios são informativos, antes de tudo, porque eles fornecem informação para nossos sistemas de decodificação de linguagem, a fim de que estes possam realizar certas operações. Nomes próprios funcionam como códigos.

É essa tese que vou defender aqui. Uma exposição mais pormenorizada da concepção de *informação* que adoto constituirá a base de minha defesa. Empreendo essa exposição na seção 3. Na seção seguinte, apresento os devidos argumentos a favor da minha tese e faço algumas observações sobre a relação entre a função referencial e a função de código dos nomes próprios. Antes de tudo, porém, quero fazer um rápido apanhado histórico para esclarecer em qual aspecto dos nomes próprios exatamente os filósofos analíticos (ou pelo menos a maior parte deles) estiveram interessados ao longo do século XX.

## **2. O que interessa explicar em relação aos nomes próprios**

O debate sobre nomes próprios no século XX envolveu duas posições principais: o descritivismo e a teoria da referência direta. A primeira diz que um nome próprio *n* refere um objeto *b* por meio de certo conteúdo (uma conotação, um sentido, uma intensão etc.); a segunda sustenta que *n* refere *b* diretamente, sem auxílio de um conteúdo. De um modo ou de outro, o propósito de ambas é explicar a forma como a função referencial de *n* é exercida.

Até aqui tudo bem. Não há nada de errado em tentar explicar a função referencial dos nomes próprios. Nomes próprios certamente têm uma função referencial e, como vimos na introdução, explicá-la é importante para esclarecer questões sobre as condições de verdade dos enunciados da linguagem. O problema é que, muitas vezes, os filósofos falam de uma forma como se a função referencial fosse a única função filosoficamente relevante dos nomes próprios

ou, pior ainda, como se ela fosse a única função dos nomes próprios. Existem várias indicações disso nos textos clássicos.

Russell, por exemplo, faz várias afirmações que indicam que ele considerava a função referencial dos nomes próprios como sua função fundamental <sup>2</sup>. Em *A Filosofia do Atomismo Lógico*, inicialmente, ele define ‘nomes próprios’ como ‘palavras para os particulares’ (cf. RUSSELL, 1985: 71). Em um segundo momento, ele substitui essa definição por uma mais abrangente e declara que:

Um nome é [...] um símbolo simples usado para designar um particular determinado ou por extensão um objeto que não é um particular mas que se trata no momento como se fosse um particular, ou que se acredita falsamente ser um particular, tal como uma pessoa (RUSSELL, 1985: 106).

Assim, para Russell, pelo menos nessa passagem, nomes próprios servem basicamente para designar particulares genuínos (no caso dos nomes logicamente próprios) ou putativos (no caso dos nomes próprios ordinários). Isso os nomes fazem de um modo peculiar. Eles designam suas denotações sem descrevê-las. Russell ressalta isso ao dizer que “o nome é simplesmente um meio de apontar uma coisa” (RUSSELL, 1985: 107). Essa função de apontar ou referir não seria apenas uma função possível do nome próprio. Russell insiste que a função referencial é a função que um nome deve ter necessariamente para ser um nome próprio. Ainda na *Filosofia do Atomismo Lógico*, ele afirma que “um nome deve nomear alguma coisa ou não é um nome” (RUSSELL, 1985: 105). Dessa forma, fica claro que, para Russell, a função necessária de um nome próprio é a referencial, algo que não desempenha essa função não é de fato um nome próprio.

Searle apresenta um ponto de vista semelhante nos *Atos de Fala*. Na seção sobre nomes próprios, lemos a seguinte pergunta: “Para que servem os nomes próprios?” A resposta é imediata e sem hesitação: “Evidentemente, para se referir aos indivíduos” (SEARLE, 1981: 228). É importante notar que Searle aqui não faz menção de qualquer outra função que os nomes próprios possam desempenhar. Para ele, é evidente que a função dos nomes próprios é por excelência a função referencial. Em contrapartida, em *Proper Names*, ele identificará outra função dos nomes próprios. Ele dirá que nomes próprios servem como pegadores nos quais penduramos descrições (cf. SEARLE, 1958, p. 172).

Kripke é outro que assume que nomes próprios servem essencialmente para referir. Toda a explicação de Kripke sobre nomes próprios se baseia na ideia de que eles são designadores rígidos. Ou seja, para Kripke, nomes servem para designar particulares, e os designam em todos os mundos possíveis em que esses particulares existem. Nesse sentido, nomes próprios funcionariam da mesma forma que as constantes de uma linguagem formal interpretada de acordo com a semântica intensional (com a diferença óbvia de que a referência de um nome não é fixada por uma função de interpretação e sim por mecanismos sociais de significação).

Mas essa visão de nomes como constantes individuais não é exclusiva de Kripke. De fato, segundo Tyler Burge, ela é uma visão generalizada entre os

filósofos. Em seu artigo *Reference and Proper Names*, depois de mostrar algumas semelhanças entre nomes próprios e constantes individuais, Burge diz: “É provavelmente verdadeiro dizer que a maioria dos filósofos, linguistas, e lógicos tem, sobre essas bases, aceitado uma visão de nomes próprios como constantes individuais” (BURGE, 1973: 426). Com isso, ele quer dizer que a maioria dos filósofos, linguistas e lógicos concorda que nomes servem basicamente para referir particulares.

As referências feitas acima parecem mostrar que os filósofos mencionados estavam interessados principalmente na função referencial dos nomes próprios. É verdade que os descritivistas acreditavam que nomes próprios podiam desempenhar outra função além da função referencial. Na medida em que postulavam um conteúdo para os nomes próprios, eles assumiam que eles eram usados para transmitir esses conteúdos (isso é sem dúvida uma função extra). Não obstante, em geral, a ideia era de que essa função extra era de alguma forma subsidiária da função referencial. Em última instância, a função referencial era que dava as cartas.

O que parece é que tanto os descritivistas quanto os adeptos da teoria da referência direta estavam, por assim dizer, enfeitiçados por certa imagem da linguagem e por isso se fixavam tão obsessivamente na função referencial dos nomes próprios. A imagem que os mantinha sob seu poder era a imagem da linguagem como instrumento de representação do mundo. Se eles tivessem dado mais atenção para o aspecto informacional da linguagem, teriam observado que as propriedades informacionais de qualquer nome próprio *n* são de fato centrais para o uso de *n*. Para que isso se torne evidente, no entanto, é preciso que tenhamos uma concepção clara do que seja informação. A seção seguinte pretende oferecer tal concepção.

### **3. Uma concepção de ‘informação’**

Temos hoje muita informação sobre *informação*, mas nem sempre foi assim. Trabalhos específicos sobre a natureza da informação começam a proliferar de modo mais efetivo só a partir do trabalho de Shannon e Weaver na Teoria Matemática da Comunicação (TMC). A partir de então, a palavra ‘informação’ passa a ser usada como um termo técnico e não apenas como um termo da linguagem comum. A teoria da computação, a psicologia cognitiva, a linguística, a física, a biologia e outras ciências passaram a se utilizar fartamente do termo. Contudo, embora esses usos conservem algumas intuições básicas, eles estão longe de ser homogêneos. Algo que deve ter contribuído para isso é a falta de uma definição de *informação* amplamente aceita. É curioso que a TMC não ofereça uma definição explícita de *informação*. Tudo o que ela faz é dar uma descrição matemática dos processos de transmissão de mensagens (cf. FLORIDI, 2010: 37 ss.).

Na filosofia, onde o termo ‘informação’ também passou a aparecer com frequência e até mesmo no título de vários trabalhos, o número dos que propõem uma definição explícita também é reduzido (DRETSKE 1981 e FLORIDI 2010 integram esse número). Além disso, as definições disponíveis não parecem retratar adequadamente aquele núcleo básico de intuições envolvidas no uso científico da noção de *informação*. As intuições que acredito serem básicas são as seguintes:

- i) **Informação é algo que um sistema envia para outro:** onde há informação, há alguém que fornece (emissor) e alguém que recebe (receptor) a informação.
- ii) **Informação é algo que vem em uma embalagem:** uma informação sempre vem inserida em um meio, contudo, ela não deve ser confundida com esse meio. (e.g., a mesma mensagem pode ser enviada por carta manuscrita, por comunicação oral etc.). Quando uma informação é inserida em um meio, diz-se que ela foi codificada, e quando é extraída desse meio, diz-se que ela foi decodificada.
- iii) **Informação é algo que sistemas não inteligentes podem acessar e processar:** o processamento de informação é um processo físico e, dessa forma pode ser realizado por qualquer sistema físico.
- iv) **Informação é algo que tem um destinatário determinado:** a informação é direcionada para certos sistemas, outros sistemas não terão acesso à informação (e.g., um texto em latim só fornece informação proposicional para quem lê latim)
- v) **Informação não é significado:** a mesma palavra pode fornecer diferentes informações para diferentes ouvintes, mas o significado da palavra é fixo.
- vi) **Informação é algo que pode causar mudanças:** um dado só é informativo se ele for capaz de causar alguma mudança em um sistema receptor apropriado. Esse é o pressuposto por trás da famosa definição de Gregory Bateson segundo a qual informação é uma diferença que faz a diferença.

Acredito que uma definição da noção de *informação* só é aceitável se expressar todas essas intuições. Como não encontrei nenhuma definição com essa característica entre as disponíveis, tive que desenvolver uma nova. A definição a que cheguei depois de algumas tentativas é a seguinte:

**Dinf.:** Informação é algo que um dado A fornece para um sistema de decodificação S se e somente se S tem uma mudança de estado quando recebe A.

De acordo com essa definição, uma informação é algo que é fornecido por um dado a um sistema de decodificação. Um dado informativo é o que chamaremos de ‘código’. É preciso chamar atenção para alguns pontos da definição e para algumas de suas implicações.

Em primeiro lugar, é necessário definir a noção de *dado*. Floridi declara que um dado é uma “carência de uniformidade” (cf. FLORIDI, 2010: 23). Aqui, eu sigo Floridi. Considero que um dado nada mais é do que um aspecto que se diferencia em um contexto <sup>3</sup>.

Pode-se postular que essa diferenciação existe na realidade independentemente de ser detectada por um ser vivo ou por uma máquina, ou pode-se considerar que ela só existe enquanto diferença detectável. Nesta última perspectiva, qualquer coisa ou aspecto que um organismo ou uma máquina possa distinguir na realidade é um dado. Na primeira perspectiva, até as coisas que não podem ser distinguidas, mas que se diferenciam numericamente das outras, são dados. Em todo caso, a definição de *dado* independe da perspectiva adotada. E, conseqüentemente, o mesmo vale para a definição de *informação*.

Em segundo lugar, é preciso esclarecer a noção de *sistema de decodificação*<sup>4</sup>. Um sistema de decodificação é qualquer sistema que implementa uma função de transição de estados. Um estado é cada uma das configurações (conjuntos de parâmetros) que um sistema apresenta ao longo do tempo, e uma função de transição de estados é uma regra que diz o que deve acontecer para o sistema passar de um estado para outro. Alguns exemplos de sistemas de decodificação são: programas, autômatos (como um editor de texto), sistemas químicos (como o procedimento usado para testar um tipo sanguíneo), sistemas biológicos (como o sistema orgânico que decodifica a informação contida em uma cadeia de DNA), dispositivos mecânicos (como uma fechadura) etc. Por esses exemplos, vê-se que não é correto supor que o processamento de informação se dá apenas em um nível abstrato e depende sempre de algoritmos elaborados, como em programas de computadores. Com efeito, é perfeitamente apropriado dizer que uma fechadura processa a informação fornecida pela chave girando, que a flauta processa a informação fornecida pelo sopro e pela digitação etc. O processamento de informação não pressupõe linguagem. Na verdade, de acordo com a minha visão, o uso da linguagem é que pressupõe processamento de informação.

É imprescindível assinalar também que a noção de *informação* não é definida simplesmente como uma propriedade de um dado. Ela é definida com base numa relação entre um dado e um sistema de decodificação. A informação fornecida por um dado é sempre uma informação para um sistema de decodificação específico. Aqui podemos observar a presença da intuição (iv): a informação tem um destinatário apropriado. Isso significa que o mesmo dado pode fornecer uma informação valiosa para um sistema de decodificação S e nenhuma informação para um sistema de decodificação S' e, inversamente, um sistema de decodificação pode ser perfeitamente adequado para interpretar certos dados e ser totalmente inútil para decodificar outros. Assim, se a especificidade da informação não for respeitada, nem o dado fornecerá informação para o sistema (é por isso que a chave do meu carro não serve para ligar outros carros), nem o sistema poderá ler a informação apropriada a partir do dado (é por isso que o meu carro só pode ser ligado com a minha chave – pelo menos eu espero). Não se deve pensar, porém, que só existe informação quando há decodificação. Para que exista informação em um dado, deve ser *possível* extrair informação dele, mas não é preciso que a informação seja *efetivamente* extraída. Um dado A pode conter informação mesmo que a informação não esteja sendo processada e mesmo que não saibamos como ela poderia ser processada ou por qual sistema. Parece natural admitir, por

exemplo, que os hieróglifos nos monumentos egípcios continham informação desde que foram escritos, muito antes portanto da descoberta da pedra de Roseta.

Outra coisa que deve ser explicada é o que um sistema de decodificação efetivamente faz. Bem, o que ele faz é ler a informação de um dado e executar uma ação. A informação é o gatilho que dispara a ação. Muitas vezes, vários sistemas podem ser combinados em série, de modo que os dados de saída de um servem como dados de entrada do outro. Nesses casos, a ação que os sistemas intermediários executam é uma ação de transformação. A informação recebida pelo sistema é transformada em uma nova informação, e essa informação, por sua vez, será lida por outro sistema. É isso que acontece, por exemplo, quando um pianista lê uma partitura. A informação fornecida pela partitura é lida por um sistema no cérebro do músico e transformada em informação mecânica, presente nos movimentos que o pianista faz para tocar. Em seguida, essa informação é lida pelo instrumento e transformada em informação sonora; e a informação sonora, por sua vez, é processada pelos sistemas auditivos dos ouvintes, e transformada em informação baseada em sinais elétricos, e assim passará a outros sistemas que a transformarão também.

Deve-se notar, como se depreende do exemplo do pianista, que é possível falar de diferentes tipos de informação. Essa classificação é conveniente, no entanto, a rigor, não é a natureza da informação que muda, e sim a natureza do meio que a contém. É o meio que é mecânico, ou sonoro, ou de qualquer outro tipo, não a informação, pois esta, conforme a intuição (ii), não deve ser confundida com sua embalagem. Assim sendo, o que chamo de ‘informação mecânica’ no exemplo do pianista, não é nada mais do que a informação fornecida por um movimento. O que chamo de ‘informação sonora’, no mesmo exemplo, é tão somente a informação veiculada por um som. E, generalizando, digo que uma informação é do tipo *x*, se *x* é um adjetivo referente ao meio que serve de veículo para a informação. Nesse sentido, é possível falar de informação sensorial, informação eletrônica, informação biológica, informação linguística e outras semelhantes.

Pode-se pensar que expressões sinônimas são dados linguísticos que fornecem a mesma informação para um usuário da linguagem, mas não é necessariamente assim. Digamos, por exemplo, que o falante A apresente seu filho para o falante B dizendo: “este é meu primogênito”. Se B não conhece a palavra ‘primogênito’, ele não vai poder decodificar adequadamente o proferimento de A. Se, em contrapartida, A disser “este é meu filho mais velho”, a informação veiculada será decodificada adequadamente. Os dois proferimentos de A são sinônimos, mas não veiculam a mesma informação para B. Isso mostra que a informação veiculada por expressões sinônimas pode variar de acordo com quem vai decodificar as expressões. Além disso, também pode haver variações na informação de acordo com a força ilocucionária do enunciado que a veicula. Por exemplo, digamos que um falante C diz de um indivíduo “aquele é nosso amado chefe” e que um falante D diz do mesmo indivíduo “aquele é nosso estimado chefe”, e, além disso, suponhamos que o proferimento de C é

interpretado literalmente e que o proferimento de D é interpretado como uma ironia. Nesse caso, podemos admitir que temos duas sentenças sinônimas comunicando informações diametralmente opostas para os seus respectivos ouvintes.

Creio que essas observações são suficientes para esclarecer Dinf. É hora de passar para a discussão da função de código dos nomes próprios.

#### **4. A função de código dos nomes próprios**

Uma ideia fundamental para mim é a de que as expressões da linguagem funcionam como códigos. De fato, podemos fazer um proferimento para alcançar diferentes objetivos, mas, até onde posso ver, qualquer que seja o objetivo, ele só será alcançado se nosso proferimento puder ser decodificado por um sistema de decodificação adequado. Se o proferimento é uma sentença declarativa, ele só pode me dar algum conhecimento se me fornecer informação; se é uma ordem ou instrução, ele só pode ser executado, seja por palavras, seja por ações, se me fornecer informação; se é uma pergunta, ele só pode ser respondido se me fornecer informação; e assim por diante. É óbvio que se esses proferimentos me forem apresentados em mandarim, eu, que não conheço nem uma vírgula de mandarim, não adquirirei nenhum conhecimento, não cumprirei nenhuma ordem, não responderei nenhuma pergunta etc. Acredito que o exame desses exemplos mostra que de fato não há possibilidade de uma expressão da linguagem ser usada de modo eficiente, seja qual for o propósito para o qual ela é designada, se não fornecer informação para um sistema de decodificação apropriado. Ora, pelo que tenho observado em minha pesquisa, o caso dos nomes próprios não é essencialmente diferente. Um nome próprio só realiza uma função se ele é usado como código. Podemos verificar isso facilmente quando o nome é usado como vocativo.

Vejam um exemplo. Eu estou andando despreocupado e de repente ouço alguém gritar meu nome atrás de mim. Em tal caso, o que eu faço em geral é me virar e procurar quem está me chamando. Por que eu faço isso? A resposta mais sensata parece ser a de que eu faço isso porque, ao ouvir meu nome, eu penso: “alguém está me chamando”, ou algo assim. O fato de eu ouvir o meu nome causa uma mudança de estado em mim. Isso indica que o nome me fornece uma informação. O resultado da decodificação dessa informação é o ato de me virar e procurar. Também em outros casos de nomes usados como vocativos, verifica-se que sempre há um efeito causado pelo nome. Em um lugar com muitas pessoas, um nome próprio pode ser usado como vocativo para selecionar o interlocutor do falante. Na sala de aula, por exemplo, eu posso perguntar: “Zezinho, qual o sentido da vida?”, e ao usar o nome ‘Zezinho’ eu seleciono o aluno que terá a árdua tarefa de responder a minha pergunta. O resultado do meu uso do nome é

que o Zezinho vai ter que responder a pergunta. Como o Zezinho sabe que lhe cabe essa tarefa? Obviamente, porque o nome 'Zezinho', ao ser proferido por mim, lhe informa disso. Em resumo, tudo indica que, quando é usado como vocativo, um nome próprio sempre fornece alguma informação para o ouvinte. Ele sempre exerce uma função de código em tais casos.

Quando um nome próprio é usado numa pergunta, como em "quem foi Napoleão?", ou em ordens, como em "vá chamar a Dalila!", também é fácil concordar que ele precisa nos fornecer alguma informação para que possamos agir de forma adequada. O próprio Kripke admite que quando alguém lhe pergunta a quem ele refere com o nome 'Napoleão', ele responde algo como "Napoleão foi imperador da França na primeira parte do século XIX; ao final, ele foi derrotado em Waterloo" (KRIPKE, 1972: 28). Isso parece indicar que, na pergunta, o nome 'Napoleão' funciona como um código que, ao ser decodificado, leva o interlocutor a dar uma certa resposta. Se a pergunta fosse "quem foi Robespierre?", a resposta seria diferente. O fato de a resposta mudar de acordo com o nome que aparece na pergunta parece indicar que cada nome fornece uma informação diferente. Mesmo quando a resposta a tais perguntas é "não sei", ainda assim o nome em questão precisou passar pelo sistema de decodificação do interlocutor. O sistema buscou certas informações na memória do interlocutor e, não as achando, produziu a resposta "não sei". Só pelo fato do sistema executar essas ações, já se pode dizer que o nome foi usado como código. Ele disparou certas operações no sistema.

Argumentos semelhantes podem ser aduzidos para o caso de ordens como as do tipo "vá chamar a Dalila!". O interlocutor só pode executar uma ordem desse tipo se ele reconhecer o nome, verificar se tem alguma informação associada ao nome e utilizar essa informação para localizar e efetivamente chamar a Dalila. E, semelhantemente ao caso anterior, pode acontecer de ele não conhecer a Dalila. Em todo caso, para poder verificar que não possui nenhuma informação sobre tal pessoa, ele precisa em primeiro lugar processar o nome. Esse processamento é o que, em tais situações, leva o interlocutor a perguntar: "quem é Dalila?". Dessa forma, fica claro que o nome sempre desempenha uma função de código.

Já quando um nome próprio é usado em uma sentença declarativa como sujeito, como objeto ou em qualquer outra função que não seja a de vocativo, pode não ser tão fácil ver de que forma o nome desempenha a função de código. Isso acontece porque o efeito do nome no interlocutor muitas vezes não é visível. Com efeito, durante a leitura de um texto eu posso deparar com uma porção de nomes próprios e continuar aparentemente impassível diante deles. Acredito, por exemplo, que minha reação ao ler o nome 'Wittgenstein' pela primeira vez não foi tão notável assim. Apesar disso, algo aconteceu em mim quando li esse nome (esse pode parecer o relato de uma epifania, mas de fato estou falando de algo bem banal). O que aconteceu em mim em tal ocasião? Para responder esse tipo de questão, devemos pensar sobre as operações que nossos sistemas internos de decodificação realizam logo que são ativados por um proferimento.

Minha hipótese é de que a primeira coisa que eles têm de fazer é realizar um tipo de *check in* de dados. Nesse *check in*, presumo que o sistema terá que fazer pelo menos três coisas: 1. reconhecer as palavras; 2. identificar a função sintática das palavras e 3. identificar a força ilocucionária do proferimento que está sendo analisado. Esse tipo de vistoria preliminar dos dados tem que ser feita antes que outros procedimentos sejam acionados. Alegoricamente falando, é como se o proferimento fosse um pacote cheio de informação, sendo que em cima do pacote há uma informação adicional que diz como a informação empacotada deve ser distribuída e a quem cada parte da informação está endereçada. Antes de repassar o conteúdo do pacote para quem de direito, o sistema deve ler as instruções na frente do pacote. Dito de forma direta, quando o sistema recebe um código, a primeira coisa que ele faz é checar as informações preliminares que orientarão o restante do processamento. Isso é o que acontece no *check in* de dados. Depois dessa primeira checagem, as outras informações contidas no código são passadas a sistemas especialistas, e aí tem início a segunda fase do processamento. É nessa fase posterior que será realizada a tarefa para a qual o código efetivamente é emitido. O *check in* de dados serve essencialmente para selecionar e classificar a informação e os parâmetros que serão processados na segunda fase.

O item do *check in* que tem maior importância na discussão sobre nomes próprios é o primeiro, o que trata do reconhecimento das palavras, por isso meus comentários vão se limitar a ele. Começemos com um exemplo já antes mencionado: um texto em latim só fornece informação proposicional para um leitor de latim. No entanto, se um analfabeto em latim por acaso depara com um texto em latim, ele sabe que não pode ler aquele texto. Ele sabe disso porque seu sistema de decodificação de linguagem tenta processar o texto e não tem sucesso. Imagino que o procedimento que o sistema usa em casos assim é mais ou menos o seguinte: palavras do texto são pinçadas e buscadas no vocabulário do leitor; uma vez que as palavras não sejam reconhecidas, o sistema produz um sinal indicando que ele não está apto a decodificar o texto. É assim que o referido leitor percebe sua incapacidade de ler o texto.

Na minha opinião, esse procedimento de reconhecimento de palavras deve acontecer sempre. Até nomes próprios precisam passar por essa etapa de reconhecimento. Não resta dúvida de que ele é essencial quando o nome aparece em perguntas como “quem é Napoleão?” e em ordens como “vá chamar a Dalila!”. Em um dos casos, o nome precisa ser reconhecido para que uma resposta seja dada; no outro caso, o nome precisa ser reconhecido para que uma pessoa seja chamada. Mas também em sentenças declarativas o reconhecimento do nome deve ser feito obrigatoriamente. Digamos, por exemplo, que eu leio o seguinte enunciado: “Wittgenstein gostava de musicais”. Quando eu leio esse enunciado, meu sistema interno de decodificação de linguagem imediatamente confere as palavras do enunciado e as relaciona com informações que se encontram registradas em minha memória, isso se as palavras forem reconhecidas. Com a palavra ‘Wittgenstein’, não é diferente. Ela também precisa passar pelo processo de reconhecimento. Se quisermos lançar mão de uma analogia, podemos dizer

que eu guardo minhas informações sobre Wittgenstein em um certo arquivo, e que o nome 'Wittgenstein' abre esse arquivo. Mas o que aconteceu quando ouvi o nome 'Wittgenstein' pela primeira vez? Digamos que tenha sido na sentença "Wittgenstein era austríaco". Posso apostar que naquele dia apliquei o mesmo procedimento de reconhecimento ao nome. A diferença é que quando meu sistema de decodificação foi verificar se já existia alguma informação que lhe pudesse ser associada, não havia nenhuma. O que eu fiz então? Simples, abri um novo arquivo para colocar informações sobre Wittgenstein e inaugurei esse arquivo com a informação que recebi por meio da sentença "Wittgenstein era austríaco". Não fiz nenhum trejeito peculiar quando fiz isso, tudo aconteceu internamente. Ou seja, o efeito que o nome provocou em mim não foi visível. No entanto, o efeito existiu e envolveu várias operações, entre as quais a operação de reconhecimento do nome. O fato de que essas operações foram deflagradas pelo nome nos autoriza a afirmar que o nome funcionou como um código. Nesse sentido, ele foi informativo.

Aqui é preciso ter muita atenção. A ideia de que um nome *n* contém informação não equivale à ideia de que *n* é sinônimo de alguma outra expressão da linguagem ou de que *n* diz algo sobre algum particular. Um nome próprio contém informação simplesmente porque causa uma mudança de estado no meu sistema de decodificação de linguagem, ele desencadeia certas operações nesse sistema. A informação que o nome fornece ao sistema é *informação-para*, ela serve para acionar certos processos mentais de computação. Entretanto, uma vez que esses processos são acionados, os nomes passam a ser associados a outras informações, e essas informações possibilitam o seu uso em novas construções linguísticas. De todo modo, deve-se observar que nomes próprios só podem ser associados a *informações-sobre* porque podem ser processados, e isso, por sua vez, só acontece porque eles sempre fornecem a nossos sistemas de decodificação *informações-para*.

Essas são, em resumo, as observações que queria fazer sobre o processo de reconhecimento de palavras. Durante esse processo, até um nome próprio precisa ser reconhecido e esse reconhecimento depende de certas informações preliminares fornecidas pelo nome. Só isso já mostra que nomes são informativos, o que equivale a dizer que eles cumprem uma função de código.

É possível, porém, particularizar mais o problema e falar das funções específicas dos nomes próprios. Com relação à função de vocativo, por exemplo, vimos que o nome só é capaz de desempenhar tal função se antes de tudo ele desempenhar uma função de código. Se a função do nome é nomear um personagem ficcional, a mesma coisa. Em relação à função referencial, não é diferente. Era essa função que o nome 'Napoleão' e o nome 'Dalila' pretendiam desempenhar em exemplos anteriores. Como acredito deve ter ficado claro, esses nomes só cumpriam sua função referencial porque antes de tudo desempenhavam uma função de código.

Mas, apesar dos argumentos já apresentados, é possível esclarecer ainda mais a dependência que a função referencial tem em relação à função de código

dos nomes próprios. Para tanto, precisamos explicitar em quais ocasiões um nome próprio é usado referencialmente. Kripke nos dá as principais indicações aqui.

Em *Naming and Necessity*, ele introduz a noção de *cadeia comunicativa*. Uma cadeia comunicativa é uma cadeia cujos elos são indivíduos que usam um nome próprio *n* com a mesma intenção de referência. O primeiro indivíduo da cadeia é alguém que estava presente quando *n* foi usado pela primeira vez para referir certo objeto *b*. O segundo indivíduo da cadeia aprendeu o uso de *n* com o primeiro, o terceiro com o segundo e assim por diante. O caso é que todos os indivíduos da cadeia usam o nome próprio *n* referencialmente. Entretanto, os critérios que o centésimo indivíduo da cadeia deve obedecer para usar o nome referencialmente não são os mesmos do primeiro. O centésimo indivíduo só usará *n* para referir *b* se sua intenção de uso der continuidade a uma certa tradição de intenções de uso, o primeiro, por outro lado, não sofre essa restrição; ele dá a *b* o nome que quiser. Assim, podemos identificar duas ocasiões em que um nome próprio *n* é usado referencialmente:

- (a) quando *n* é usado para nomear um objeto *b* em um ato de nomeação inicial; e
- (b) quando *n* é usado com a intenção de nomear *b* depois de ter sido dado a *b* em um ato de nomeação inicial, e em razão disso.

Estou convencido de que, nessas duas ocasiões, um falante *A* só terá sucesso em usar *n* referencialmente se antes usar *n* para codificar informação. Para demonstrar isso, dou argumentos específicos para os casos (a) e (b) respectivamente.

Quando um nome próprio é usado do modo descrito em (a), temos o que se pode chamar de ‘batismo’. Nessa ocasião, segundo Kripke, duas coisas podem ocorrer: “o objeto pode ser nomeado por ostensão, ou a referência do nome pode ser fixada por uma descrição” (KRIPKE, 1980: 96). Examinemos logo a segunda alternativa.

Se o batismo é realizado com o auxílio de uma definição do tipo “*n* é *D*”, onde *D* é uma descrição definida que descreve o objeto *b*, o que acontece é que *n* vai passar a ser associado à informação fornecida pela definição. Essa é a informação que será codificada em *n* em um primeiro momento. Todos os falantes que codificarem essa informação em *n* estarão, portanto, usando *n* para referir *b*, ou seja, estarão usando *n* do modo estabelecido no batismo. Os que não conectarem *n* com essa informação, simplesmente não poderão usá-lo para referir *b*, pelo menos não com base no batismo. Dessa forma, fica claro que, para cumprir seu objetivo, que é criar uma nova tradição de uso para o nome próprio *n*, o batismo depende inteiramente de que os falantes codifiquem certas informações em *n*.

Na minha opinião, o caso em que o batismo é realizado com a presença do objeto nomeado, não é diferente. Nesse caso, a eficácia do batismo também depende de que o nome seja usado para codificar informação. A informação dessa vez é principalmente informação sensorial; os indivíduos que participam do

batismo extraem essa informação diretamente do objeto b. Se eles não codificarem essa informação em n não poderão usar n para referir b posteriormente e o batismo terá sido ineficaz. Um exemplo pode deixar esse ponto mais claro.

Vamos supor que Paul Klee apresentou *Ab Ovo* pela primeira vez em uma exposição e que no começo do evento a pintura estava encoberta por uma pequena cortina. Agora, digamos que um grupo de pessoas que viu o quadro encoberto não aguentou a curiosidade e foi perguntar ao artista que quadro era aquele. Digamos então que Paul Klee tenha respondido sucintamente: “Este é meu quadro *Ab Ovo*”, sem, contudo, remover a cortina. Depois de um tempo, Paul Klee pediu a todos os presentes para se reunirem diante do tal quadro, descerrou a cortina e disse: “Este é meu quadro *Ab Ovo*”. Pergunto: quando se deu realmente o batismo? Na primeira ou na segunda vez que Paul Klee disse “Este é meu quadro *Ab Ovo*”? Parece óbvio que o ato só foi concretizado quando o artista proferiu a frase pela segunda vez. Por quê? O que aconteceu na segunda vez que não aconteceu na primeira? Obviamente, o que aconteceu foi que as pessoas viram a pintura. Com isso, elas associaram o nome ‘*Ab Ovo*’ a certas informações visuais, ou seja, elas codificaram informação no nome.

Vemos assim que não importa se o nome próprio n é introduzido por uma descrição definida ou por ostensão. Em qualquer situação, o batismo só terá êxito se n for usado como código. Isso é o que ocorre no caso (a). Mas quanto ao caso (b)?

No caso (b), podemos identificar três condições que precisam ser satisfeitas para que n seja usado referencialmente: 1. é preciso que o uso que é feito de n tenha base histórica, ou seja, é preciso que em algum momento do passado b tenha sido batizado com o nome n; 2. é preciso que o falante que usa n tenha a intenção de usar n para referir b e não outro objeto qualquer; e 3. é preciso que, de alguma forma, a intenção de uso do falante seja ocasionada pelo fato de b ter sido batizado com o nome n. Se essas três condições são satisfeitas, temos um legítimo uso referencial do nome próprio n. Acontece que, pela minha ótica, essas condições só podem ser satisfeitas se n for usado para codificar informação. Isso já foi demonstrado em relação à primeira condição nos parágrafos anteriores, resta, portanto, demonstrar o mesmo em relação às condições 2 e 3.

O requerimento 2 exige que o falante tenha certa intenção de uso a fim de usar n para referir b. De fato, é Kripke que faz essa exigência; em uma cadeia de comunicação, o aprendiz deve ter a intenção de usar o nome da mesma forma que seu instrutor o usou (cf. KRIPKE, 1972: 96). Se isso não for feito, a ligação entre o nome e sua referência não será estabelecida. Mas o que exatamente significa dizer que um falante A tem uma intenção de uso para um nome próprio n? Bem, uma coisa que pode nos ajudar a responder essa questão é examinarmos um caso concreto. Digamos que A diga “eu sou um grande fã de Sócrates”. Aqui poderia ser conveniente indagar de A: “Qual a sua intenção de uso para o nome ‘Sócrates’? Você está falando do filósofo ou do jogador de futebol?”. Se ele dissesse: “Do filósofo” ou “Do jogador de futebol”, ele esclareceria sua intenção de uso. Isso mostra que a intenção de uso de A serve como dispositivo de desambiguação.

É por meio dela que, dentre várias possíveis referências do nome próprio *n*, A escolhe uma para ser a referência de *n* em dada ocasião. Só isso já mostra o quanto a intenção de uso é importante no processo do uso referencial de um nome próprio. É importante observar, no entanto, que os falantes não lançam mão de intenções de uso apenas para resolver problemas de ambiguidade. Sempre que um falante A usa um nome próprio *n*, há uma intenção de uso envolvida. A intenção de uso é o que direciona um nome próprio para uma referência. De fato, é por isso que ela pode servir como dispositivo de desambiguação; quando a intenção direciona o nome para uma referência, ela exclui as outras possibilidades de referência. Mas voltemos ao exemplo do proferimento sobre Sócrates.

O interessante sobre o exemplo é que A esclarece sua intenção de uso associando o nome ‘Sócrates’ a certas informações. Se A responde que ele se refere ao filósofo, é porque ele codifica no nome a informação de que *Sócrates era filósofo*. Se, por outro lado, A responde que ele se refere ao jogador de futebol, é porque ele codifica no nome a informação de que *Sócrates era jogador de futebol*. Isso posto, temos já condições de responder a questão sobre o que exatamente significa dizer que um falante A tem uma intenção de uso para um nome próprio *n*. Significa dizer simplesmente que A codifica certas informações em *n*. É em virtude disso que A usa *n* em certos enunciados e deixa de usá-lo em outros. A conclusão a que chegamos então é a seguinte: exigir que um falante tenha uma intenção de uso para *n* é o mesmo que exigir que ele use *n* para codificar certas informações. Logo, o requerimento 2 só é atendido se *n* é usado como código.

A terceira condição para que um falante A use um nome *n* do modo descrito no caso (b) é a de que seu uso se conecte com uma certa tradição de uso. Isso, obviamente, ocorre quando a intenção com que A usa *n* coincide com a intenção com que os outros falantes naquela tradição usam *n*. Se pudermos determinar em que circunstâncias se dá essa coincidência saberemos em que circunstâncias essa terceira condição é satisfeita.

Isso não parece ser algo tão difícil de determinar uma vez que já esclarecemos a noção de *intenção de uso*. Se dois falantes têm intenções de uso para o nome próprio *n*, os dois codificam informações em *n*; suas intenções coincidirão se os dois conjuntos de informações codificadas em *n* compartilharem uma parte significativa de seus elementos. Certamente, os conjuntos não precisam ser idênticos. Por exemplo, um falante A pode associar o nome ‘Sócrates’ à informação de que *Sócrates foi o maior filósofo de todos os tempos* e o falante B pode associar o mesmo nome à informação de que *Sócrates foi um filósofo medíocre*, e, mesmo assim, normalmente vai acontecer deles terem a mesma intenção de uso para o nome ‘Sócrates’. O que garante que eles tenham essa mesma intenção é o fato de A codificar em *n* várias informações importantes que B também codifica em *n*. O fato de as informações não coincidirem uma a uma não prejudica a comunicação entre A e B.

Temos, assim, uma elucidação do que significa ter a mesma intenção de uso. Mas isso não é tudo que precisamos elucidar. É preciso explicitar o que deve

acontecer para que dois falantes venham a ter uma mesma intenção de uso para um nome próprio *n*. Ao que parece, tudo depende do processo de ensino e aprendizagem do uso do nome. Quando um falante B ensina o falante A a usar *n*, as informações que A vai codificar em *n* em um primeiro instante serão as que ele extrai dos proferimentos e de todos os outros códigos (gestos, desenhos, entonação, expressão facial etc.) que B lhe fornece. Se A decodificar corretamente os códigos que B lhe oferece, o conjunto de informações que ele vai codificar em *n* estará profundamente entrelaçado com o conjunto que B codifica em *n* e, por conseguinte, os dois falantes terão a mesma intenção de uso para o nome. Em contrapartida, se a decodificação que A faz dos códigos fornecidos por B não é bem sucedida, cria-se um desvio na cadeia causal de comunicação. Temos então casos como o de Madagascar. Fica claro, portanto que a conexão entre o uso que um falante A faz do nome *n* e o uso que se faz de *n* em uma certa tradição depende do processo de aprendizagem do nome. Esse, por sua vez, depende dos procedimentos de decodificação que A implementa no momento em que recebe o nome. Dessa forma, podemos concluir que a terceira condição para que um falante A use um nome *n* do modo descrito no caso (b) só é satisfeita se, no período da aprendizagem, *n* é usado como código tanto por A quanto por seu instrutor.

Isso posto, fica concluída a minha exposição a respeito do que acontece quando um nome próprio é usado referencialmente. Como vimos, sempre que um nome próprio *n* é usado assim, vários processos de decodificação e codificação são levados a efeito pelos usuários de *n*. E esses processos não acompanham o uso referencial do nome inutilmente. Na verdade, o que os argumentos apresentados nos parágrafos anteriores mostraram foi que eles são necessários para que o nome seja usado referencialmente, seja qual for o tipo de uso referencial que é feito do nome. Eles são necessários: 1. para que o batismo seja bem sucedido, 2. para que os falantes possam manifestar uma intenção de uso para o nome e 3. para que o uso que o falante faz do nome se alinhe com uma determinada tradição de uso. Para todos esses propósitos, processos de codificação e decodificação são requeridos. Isso demonstra de forma completa a precedência da função de código sobre a função referencial dos nomes próprios.

## Referências

- BURGE, T. "Reference and Proper Names". *The Journal of Philosophy*, vol. 70, nº 14, pp. 425-439, ago. 1973.
- CAPURRO, R. "Past, Present, and Future of the Concept of Information". *TripleC* (Special Edition: What Is Really Information? An Interdisciplinary Approach), vol. 7, nº 2, pp. 125-40, 2009.
- CHOMSKY, N. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press, 2000.
- DRETSKE, F. *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press, 1981.

FLORIDI, L. *Information: A Very Short Introduction*. Oxford: Oxford University Press, 2010.

KRIPKE, S. *Naming and Necessity*. Cambridge: Harvard University Press, 1980.

RUSSELL, B. “A Filosofia do Atomismo Lógico”. In: *Ensaaios Escolhidos* (Coleção: Os Pensadores). 2ª ed. São Paulo: Abril Cultural, 1985.

RUSSELL, B. “Knowledge by acquaintance and knowledge by description”. *Proceedings of the Aristotelian Society* (PAS), New Series, v. 11, pp. 108-128, 1910-11.

SEARLE, J. *Os Actos de Fala - um ensaio de filosofia da linguagem*. Coimbra: Livraria Almedina, 1981.

SEARLE, J. “Proper names”. *Mind* (New Series), vol. 67, nº 266, pp. 156-173, abr. 1958.

## Notas

<sup>1</sup> Agradeço aos membros do meu grupo de pesquisa (NEFA) pelas discussões que em muito contribuíram para a construção deste artigo. Agradeço também à Universidade Federal do Ceará por possibilitar minha participação no Principia e à Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) pelo suporte financeiro a minha pesquisa.

<sup>2</sup> É importante notar que apesar de Russell ter essa opinião, ele certamente não considera que a função referencial é a única função dos nomes próprios. Ele também mostra que pelo menos os nomes próprios ordinários têm um papel na formação de juízos na mente dos falantes (cf. RUSSELL, 1910-11: 119-20).

<sup>3</sup> A noção de *aspecto* não é definida aqui. Em todo caso, gostaria que ela fosse entendida em seu sentido coloquial.

<sup>4</sup> De fato, em vez de usar a expressão “sistema de decodificação” poderia usar simplesmente a palavra “sistema”, mas para isso teríamos que interpretar a palavra de acordo com um sentido muito estrito. Como a palavra tem outros sentidos, inclusive um sentido muito próprio do contexto filosófico (e.g., “Gilles Deleuze não tem um sistema”), acredito ser melhor usar a expressão “sistema de decodificação”.

# SIGNIFICADO E PERCEPÇÃO: UMA FALA ACERCA DAS SENSações

DEBORA FONTOURA DE OLIVEIRA

*Doutoranda em Filosofia pela Universidade do Vale do Rio dos Sinos (UNISINOS) e  
Bolsista CAPES/PROSUP.*

debbyfontoura@gmail.com

**Resumo:** Neste artigo pretende-se abordar a relação entre significado e percepção em discursos sobre as sensações. Busca-se entender como frases, que descrevem sensações, são usadas de maneira intersubjetiva. Para alcançar este objetivo, o presente trabalho divide-se em três seções: na primeira seção, tendo como base a discussão na filosofia da linguagem, serão expostas as principais distinções entre o internalismo semântico e o externalismo semântico; na segunda, será exposta a importância de um posicionamento naturalista para garantir uma intersubjetividade frente ao significado de palavras que remetem a sensações; para então, na última seção, apresentar a relação entre significado e percepção em discursos sobre as sensações - nessa seção abordar-se-á o problema da visão das cores. Ao fim, pretende-se destacar que perceber as cores e ter a capacidade de diferenciá-las não depende somente das sensações, mas também do meio e, principalmente, da linguagem.

**Palavras-chave:** Cores. Intersubjetividade. Percepção. Sensações. Significado.

## 1. Introdução

Apoiado em algumas teorias sobre as sensações, propõe-se neste artigo abordar a relação entre significado e percepção. De modo a tentar compreender o quanto objetivo pode ser um discurso acerca das sensações, leva-se em consideração a capacidade perceptiva e o questionamento sobre os significados dos termos utilizados nesses discursos. Sendo este o foco principal, busca-se explicar como palavras, utilizadas para descrever experiências sensoriais, são usadas de forma semelhante por diferentes falantes de uma linguagem, visto que realmente existe um problema de compreensão e explicação do significado dos termos sobre sensações.

Para tal articulação do tema adota-se a seguinte estratégia: primeiro, para poder compreender como é possível a comunicação acerca das sensações, faz-se necessário expor algumas teorias semânticas que se encontram na filosofia da linguagem, de modo a apresentar a concepção que cada uma possui com respeito ao significado. Em um segundo momento, busca-se examinar a relação entre a percepção e os significados. Neste estágio, tratar-se-á sobre as sensações que, comumente, são compreendidas como resultantes da capacidade de captar a

realidade que nos cerca. Como uma das problemáticas envolvidas será chamada a atenção para o problema que é posto pela visão das cores.

O que se pretende ao final do trabalho é levantar alguns questionamentos e possibilidades de como é possível compreender o significado em discursos acerca das sensações, sempre lembrando que a linguagem não é um fenômeno estático, mas sim, um processo dinâmico, no qual, o sujeito cria conceitos para estruturar seu próprio ambiente.

## **2. Teorias Semânticas: internalismo versus externalismo**

A comunicação, que se efetiva por formas de linguagem, é o que fundamenta a intersubjetividade no plano das interações sociais. A partir disso, é que a utilização de sentenças funciona como uma espécie de “ferramenta” para a transmissão do que poderíamos chamar “fluxos de pensamento”. O entendimento entre os seres humanos pressupõe, assim, frases utilizadas intersubjetivamente, as quais devem de alguma forma veicular significados que possibilitam a sua compreensão por todos. Pode-se dizer que as frases são utilizadas para expressar aquilo que queremos dizer ou nos referir, pressupondo que estamos declarando algo com determinado significado e que este, por sua vez, seja intersubjetivo.

Nos estudos sobre a linguagem, observa-se que discussões que abordam o significado de palavras sobre experiências, inseridas em juízos que expressam crenças, estão ainda em aberto. As sentenças que descrevem sensações como, por exemplo, as cores ou ainda dores, remetem a várias questões: como palavras utilizadas para descrever experiências sensoriais são usadas de forma semelhante por diferentes falantes de uma linguagem? Como determino o significado das palavras que remetem às sensações? O significado é uma representação mental ou encontra-se no mundo exterior? Como posso garantir que o significado de uma sentença seja o mesmo para dois falantes?

Para entender o papel que o significado possui em discursos sobre as sensações é importante destacar os tipos de teorias semânticas. No debate filosófico contemporâneo sobre a filosofia da linguagem podemos dizer que se formaram duas posições correntes: o internalismo semântico e o externalismo semântico.

O internalismo semântico, como entendido neste artigo, compreende a uma concepção da linguagem como privada, isto é, é uma corrente que prioriza a concepção de mente e tenta explicar e analisar o fenômeno físico da fala apelando a atividades e entidades mentais. Ou seja, recorre à noção de pensamento, idéia, etc. Essa perspectiva, concebida também como mentalista, encontrou um suporte intuitivo que estabelece uma ênfase para o ponto de vista da primeira pessoa. Entretanto, é necessário deixar claro que na história da filosofia da linguagem são encontradas duas diferentes concepções de mentalismo. Uma defende que a

correlação das palavras com os significados se dá na mente, isto é, o processo de identificação da linguagem é um ato mental, no entanto os significados não são entidades da mente, mas encontram-se em outro “mundo”, não mental nem físico; a outra concepção de mentalismo, muito mais radical, é aquela que identifica significados com entidades mentais, isto é, com estados psicológicos. Dentro da primeira concepção encontra-se Gottlob Frege (1978) como seu grande representante, para ele significados não são representações mentais de uma consciência individual, mas situam-se em um “terceiro reino”, fora da consciência e dotados de realidade objetiva. Somente é possível ter acesso aos significados quando estes estiverem relacionados a outros componentes semânticos, formando assim um pensamento completo<sup>1</sup>. Por outro lado, a segunda concepção defende que os significados estão determinados dentro da “cabeça” de um falante; tem-se como um de seus representantes o empirista John Locke (1991) que identifica os “significados” com entidades mentais, apesar das ideias, segundo esse autor, serem constituídas a partir das sensações elas possuem existência psíquica<sup>2</sup>.

Além do mais, a concepção da linguagem como mentalista, defende que seu conteúdo é determinado por fatos internos ao corpo, semelhante a uma “linguagem de pensamento”, isto é, a um quadro ou uma representação no cérebro. O conteúdo dos pensamentos prescindiria de uma avaliação no mundo exterior, ou do questionamento de outros, e seria alcançado apenas pelo ato de introspecção<sup>4</sup>. De acordo com a semântica mentalista o conceito de significado é concebido como um tipo de entidade especial que está determinada de algum modo na mente para além do que está implícito nas disposições comportamentais, identificando assim o conceito de significado com uma “ideia”. Por exemplo, o significado de ‘vermelho’ é a ideia de vermelho. Essa visão propõe que o papel da linguagem é o de dar expressão as “ideias” que são anteriores e logicamente independentes em relação à linguagem.

Contrapondo-se ao internalismo semântico, a corrente externalista postula que a linguagem deve ser fundada em evidências sensoriais, isto é, os significados não são determinados por um traço individualista de estados mentais – “os significados não estão na cabeça”. Ou seja, o conteúdo proposicional de um estado interno é constituído por fatores externos ao agente. Opondo-se e dirigindo duras críticas ao mentalismo está W. V. Quine. Este afirma que assegurar o significado como objeto mental já é uma objeção suficiente contra o mentalismo, a objeção maior, entretanto, se faz ao concebê-lo como “ideia platônica”, ou como objeto concreto denotado (OR, 1969, p. 27). O problema dos mentalistas, segundo este autor, é postular o mental como a explicação do significado para garantir assim, a sua determinação. Quine discorda completamente apoiando-se em uma posição externalista. Para ele, o significado linguístico é uma função de comportamento observável em circunstâncias observáveis, em que o significado é determinado pelo comportamento linguístico e não algo que já está na mente. Sua proposta defende uma semântica behaviorista (QUINE 1969, p. 26)<sup>4</sup>. De acordo com ele, o conceito de significado concebido pelos mentalistas não possui

critérios de avaliação satisfatórios. Se essa semântica for reformulada a partir de termos comportamentais, é possível falar em semântica científica a partir de uma visão behaviorista<sup>5</sup>.

A diferença entre internalistas e externalistas nesse caso é que os primeiros sustentam que a apreensão do significado é um processo determinado mentalmente, enquanto que os últimos defendem que elementos do meio desempenham um papel na determinação do significado.

Os externalistas como Saul Kripke (1980) e Hilary Putnam (1975), defendem que é o meio que determina a “referência” e fornece os significados das palavras. Ou seja, “significado” é algo compartilhado pelos falantes de uma linguagem, e seu conhecimento está distribuído dentro de uma comunidade de fala.

Mas, afinal, o que faz com que, ao pronunciarmos uma sentença, tenhamos a certeza de que ela será entendida da mesma forma que nós pelo ouvinte? Ao pronunciarmos uma sentença do tipo “Esta bola é vermelha” o que faz com que o pensamento particular seja sobre a bola e que se tenha as condições de verdade satisfeitas para afirmar que é vermelha? Existe um método determinado/garantido que possibilita corresponder corretamente a percepção que tenho com a representação linguística que utilizo?

### **3. Significado e intersubjetividade: o naturalismo na linguagem**

Muitos filósofos tentam responder a questões que discutem o problema da intersubjetividade do significado, principalmente, através de uma semântica naturalista, que se utiliza dos conhecimentos das ciências naturais, especialmente os da biologia e da psicologia. A semântica naturalista vem reforçar que as propriedades semânticas (propriedades intencionais) podem ser explicadas a partir de fatos naturais. Quine enfocou a importância de uma metodologia naturalista<sup>6</sup>, dado que a linguagem é uma construção social. Portanto, a semântica deveria ter como evidência o comportamento observável intersubjetivamente. De acordo com ele:

A ciência ela mesma ensina que não existe clarividência, a única informação que nossa superfície sensorial pode obter dos objetos externos deve ser limitada a projeções óticas bi-dimensionais e vários impactos de vibrações de ar sobre os tímpanos e algumas reações gasosas em passagens nasais e alguma miscelânea de semelhanças (QUINE, 1974, p. 2).

Quine explica a “constituição” da linguagem e de seu significado pela descrição da maneira como indivíduos são condicionados socialmente a usar palavras e frases em determinadas circunstâncias com as suas estimulações. O problema todo de Quine não é com uma teoria da mente, mas com a postulação de termos mentais. E segundo a sua proposta, os termos mentalistas, como crenças, desejo, e assim por diante, devem ter conteúdo empírico. Disso se segue que esses termos são, de alguma forma, aprendidos: “Tais termos são aplicados à

luz de indícios publicamente observáveis: indícios estritamente físicos de estados corpóreos (...) Sem os traços visíveis antes de tudo, termos mentalistas não poderiam ser aprendidos de qualquer modo” (QUINE, 1985, p. 5 - 6).

Para Quine, uma criança, por exemplo, não apenas ouve a sentença, vê a situação, e associa os dois. Na situação, ela também nota a orientação do falante, o gesto e a expressão facial, e desta forma ainda constata que o falante também percebe a situação. Quando uma criança faz uso da sentença, existe ainda um perceber do perceptivo<sup>7</sup>.

Todavia, opondo-se ao posicionamento quinianiano encontramos Noam Chomsky, Donald Davidson, John Searle, entre outros. Chomsky, por exemplo, vai assegurar que:

Se uma linguagem é definida como um “complexo de disposições ao comportamento verbal”, determinado simplesmente em termos de probabilidade de resposta em dadas situações sem referência a qualquer postulada teoria de competência, nós estaremos diante de um grande número de problemas (CHOMSKY, 1975, p. 196).

Defendendo uma semântica inatista Chomsky sugere o estudo da linguagem como um meio para fornecer uma perspectiva notavelmente favorável para o estudo do processo mental humano.

O aspecto produtivo do uso da linguagem, quando investigado com cuidado e respeitado pelos fatos, mostra que as noções correntes de hábito e generalização, como determinantes de comportamento ou conhecimento, são bastante inadequadas. A abstração da estrutura linguística reforça essa conclusão, e sugere mais adiante que em ambos, percepção e aprendizagem, a mente executa um papel ativo em determinar o caráter de ganho de conhecimento (CHOMSKY, 1972, p. 98 - 99).

Segundo Chomsky, é necessário postular entidades mentais, determinadas e desenvolvidas geneticamente, sem aprendizagem, as quais são caracterizadas abstratamente por uma gramática, esta sim, considerada como um componente do comportamento de um falante e ouvinte (CHOMSKY, 1959, p. 577). Chomsky, sendo um internalista, pensa que existem certas características inatas e universais a partir das quais se pode compor uma “gramática gerativa”<sup>8</sup>.

Searle, em *Indeterminacy, Empiricism and the First Person* (1987), declara que existem características essenciais para uma individuação de estados intencionais que não podem ser capturados em termos de terceira pessoa. O contra-exemplo utilizado por ele à proposta behaviorista é o argumento do Quarto Chinês - “Chinese Room”<sup>9</sup>. Segundo ele, o behaviorismo é refutado completamente por este experimento. Searle afirma:

Agora, suponha que tentássemos reduzir a sensação de dor subjetiva, consciente, de primeira pessoa, aos padrões de descargas neuronais objetivas, de terceira pessoa. Suponha que tentássemos dizer que a dor é na verdade “nada exceto” os padrões de descargas de neurônios. Bem, se tentássemos uma tal redução ontológica, as características essenciais da

dor seriam deixadas de lado. Nenhuma descrição dos fatos de terceira pessoa, objetivos fisiológicos, expressaria o caráter subjetivo, de primeira pessoa, da dor, simplesmente porque as características de primeira pessoa são diferentes das características de terceira pessoa. (SEARLE, 1992, p. 117)

Searle vai sustentar, no que ele chamou de “naturalismo biológico”<sup>10</sup>, que os processos de estados mentais não são tão diferentes dos que ocorrem em nosso estômago. De acordo com Searle, os estados mentais são tão reais quanto outros fenômenos biológicos, eles são causados e realizados no cérebro (SEARLE, 1980, p. 455).

Opondo-se ao posicionamento internalista de Searle, encontramos Putnam (1975), que sustenta que o aspecto social e público da linguagem é necessário para a semântica, pois afirma que fatores relativos ao ambiente jogam um papel essencial para a determinação do significado. Putnam parece demonstrar que é um erro considerar os “significados” como entidades teóricas – que podem ser isolados e que cumprem uma função explicativa em uma teoria científica, assim como concebê-los como entidades mentais.

Para rebater a concepção mentalista, Putnam utiliza um experimento mental em que se imagina em uma “Terra Gêmea” – semelhante à Terra – um líquido com as mesmas características que a água, porém com uma fórmula química diferente – XYZ – de H<sub>2</sub>O. Antes de conhecer a fórmula química, os terráqueos e os visitantes da Terra Gêmea pensavam estar usando o mesmo líquido, tendo a “água” o mesmo significado na Terra e na Terra Gêmea. No entanto, isso é errôneo, pois a palavra “água” possui dois significados, um usado na “Terra Gêmea” e o outro na Terra. O que Putnam quer é salvaguardar uma noção de significado, sendo ela de significados científicos. Portanto, o estado psicológico do falante não determina o “significado” da palavra. O que o leva a concluir que “‘significados’ não estão na cabeça”<sup>11</sup>.

Já Davidson (1979, 1990) aborda o problema do significado, mais especificamente a indeterminação deste, argumentando que perguntar sobre o significado de uma palavra somente é possível restringindo-se a uma linguagem particular, na qual uma teoria do significado desta seria adequada formalmente se ela satisfizesse os limites de uma teoria da verdade, ou seja, que fosse adequada empiricamente assegurando, assim, as condições de verdade desta linguagem, ajustando favoravelmente a evidência sobre a linguagem. Porém, como Davidson (1979, p. 228) observa, isso não seria suficiente para determinar uma única teoria, mas sim variadas teorias e nenhuma idêntica, permanecendo o problema da indeterminação do significado.

A indeterminação do significado se mostra porque a própria referência é indeterminada, mesmo que as formas lógicas e as verdades sejam fixadas. Demonstra-se, a partir disso, que a mesma referência pode satisfazer a dois esquemas diferentes na linguagem de um falante. As tentativas de encontrar o significado de uma palavra ao associar ao seu referente não é a saída para determinar o significado de uma palavra<sup>12</sup>.

Apoiando-se neste debate, a proposta de Wilfrid Sellars (2008) defende que a linguagem não é uma tradução de pensamentos linguísticos ou simplesmente uma forma de expressá-los, mas:

Meu problema imediato é ver se consigo conciliar a idéia clássica de pensamentos como episódios internos, os quais não são nem comportamento público nem representação verbal e aos quais é apropriadamente feita referência por meio de termos do vocabulário da intencionalidade, com a idéia de que as categorias da intencionalidade são, no fundo, categorias semânticas pertencentes a *performances* verbais públicas. (SELLARS, 2008, p. 97)

Para Sellars, o significado de uma palavra é seu uso correto no meio público, não é algo particular, mas sim o que “outras mentes” concordariam em aceitar como tal. A interiorização – criação de uma representação interna – se torna secundária para Sellars; o que é importante é a função que os pensamentos irão desempenhar no “espaço lógico das razões”. Logo, o significado da linguagem não é a explicação da intenção de um pensamento, dado que este depende dela, isto é, para pensar é necessário a linguagem. O significado dos pensamentos refere-se à fala, o significado da fala às práticas intersubjetivas. Sellars mantém, assim, “o behaviorismo verbal”, que afirma que é o comportamento verbal que permite expressar pensamentos e conceitos.

Mas a questão permanece: o que seria o significado dos termos que supostamente fazem referência às sensações? Afinal, o que faz com que, ao pronunciarmos uma sentença, tenhamos a certeza de que ela será entendida da mesma forma que nós pelo ouvinte? O que possibilita corresponder corretamente a percepção que tenho com a representação linguística que utilizo? O que está em jogo é como posso afirmar algo sobre minha percepção com verdade objetiva. Sendo assim, como é possível sustentar um discurso, sobre uma experiência sensorial, com um significado intersubjetivo? Podemos conhecer a natureza dos estados sensoriais do outro?

#### **4. Percepção e Significados: o significado dos termos que remetem a sensações**

Se defendermos que somente é possível termos acesso ao próprio estado sensorial podemos recair em um solipsismo, que nos leva a conceber o próprio sujeito como o único capaz de ter experiências, não podendo, assim, termos acesso a experiência do outro. Jonathan Dancy (1985) retoma a argumentação de Kripke (1980), acerca do argumento da linguagem privada e a análise de como um solipsista adquire a sua linguagem, e expõe que o solipsista desenvolve um conceito a partir do seu próprio caso, das suas próprias sensações, aplicando a palavra a todas as sensações semelhantes à primeira. Porém, para uma palavra ter significado é necessário que existam regras para sua utilização, e neste sentido

o solipsista poderia pensar estar seguindo uma regra, quando na verdade não estava de fato, ou ainda, pensar que a sensação parece assemelhar-se a original, mas não se assemelha. Assim, o solipsista não é capaz de criar uma regra que possa tornar objetiva a relação entre a palavra e a sensação, tudo o que quiser dizer será correto. Isso transforma o discurso em algo sem sentido e vazio. É necessário, desta forma, olhar para além do indivíduo, para a comunidade, a linguagem, deste modo, é necessariamente algo público<sup>13</sup>.

Uma saída plausível para considerar, que fornece uma relação entre o estado mental e o acesso público das sensações, é uma concepção behaviorista. Esta defende que os estados mentais, por exemplo, de ter uma dor, ver uma cor, pode ser acessado pelo comportamento, pela disposição comportamental do sujeito. Porém, descrições comportamentais, embora auxiliem no acesso ao que o outro experiencia, não conseguem dar conta de tudo aquilo que é característico a mente. O fato de um sujeito se comportar de um modo X ao observar um objeto da cor vermelha, ou se comportar de uma forma Y ao sentir uma dor, não implica que isto seja tudo que se pode dizer sobre a dor ou o vermelho.

Na filosofia da linguagem observa-se que discussões que abordam o significado de palavras provindas das experiências são um tanto obscuras. Em sentenças, que descrevem sensações, por exemplo, sobre as cores ou ainda sobre dores, qual seria o significado dos termos que referem-se à essas sensações?

Como uma das problemáticas envolvidas com respeito às sensações, chamaremos a atenção ao problema que é posto pela visão das cores. Para exemplificar este problema, algumas teorias defendem que as cores não são simplesmente obtidas pela capacidade perceptiva, mas produzem uma interação entre sensação, percepção e ainda interpretação<sup>14</sup>. Há outras teorias que sustentam que os olhos possuem certos receptores especializados no reconhecimento das cores vermelho, azul e verde. Desta forma, uma luz vermelha estimularia os receptores vermelhos, mas não estimularia os receptores azuis e verdes, já uma luz amarela estimularia os receptores do verde e do azul ao mesmo tempo. Algumas teorias argumentarão que as proposições sobre cores repousam sobre uma relação entre um esquema inato de reconhecimento de cores e as próprias cores percebidas. Outras asseguram que as cores estão nos objetos e os sentidos apenas captam as informações que já estão no mundo, isto é, as cores produzem certo tipo de sensação. Embora existam muitas teorias que tentam responder ao problema das sensações, observa-se que permanece um abismo semântico ao falarmos acerca dos termos utilizados para descrever as sensações.

Para ilustrar a problemática encontrada na relação entre linguagem e sensações, temos como exemplo os Esquimós. Na cultura dos Esquimós são encontradas, na sua língua o inuíte, mais de 15 palavras para referir-se à cor branca, dado que eles perceberiam diversas nuances da cor, dando um nome a cada uma delas<sup>15</sup>. O fato dos esquimós reconhecerem mais de 15 nuances da cor branca não significa que eles possuem olhos anatomicamente diferentes dos nossos, por exemplo, apesar de utilizarmos somente a palavra “branco”. Isso significa que a língua portuguesa limita o uso das palavras e não, por exemplo,

que um brasileiro não pudesse identificar as 15 nuances assim como o esquimó. A palavra “branco”, na língua portuguesa, reúne todas as variações em um único conceito, enquanto que a língua inuíte reconhece 15 significados distintos que fazem referência a tonalidades de branco.

Deste modo, a cor da neve que um brasileiro e um esquimó descreveriam seria, de um ponto de vista semântico, a mesma? Como este caso, há outros, por exemplo, em algumas culturas não há palavras para a cor laranja, ou ainda, as cores azul e verde são descritas por uma mesma palavra, etc. À primeira vista, parece que todos percebem a mesma cor, mas a linguagem à etiqueta de forma diferente. No caso dos esquimós, o meio exigiu que eles categorizassem a linguagem de maneira diferente, logo um falante da língua portuguesa não conseguirá identificar as 15 nuances, mas não porque é incapaz, mas porque a sua linguagem não faz isso e sua vida não o requer. Dentro da própria língua é possível encontrar casos semelhantes como, por exemplo, quando duas pessoas, ao referirem-se às águas do oceano, identificam-nas com cores diferentes: uma diz que a água é azul e a outra que é verde.

Disso se segue que perceber as cores e ter a capacidade de diferenciá-las não depende somente das sensações, mas também do meio, da aculturação e principalmente da linguagem.

Wittgenstein, em suas obras *Investigações Filosóficas* (1953) e *Gramática Filosófica* (1974), discorre sobre uso de palavras que descrevem sensações na linguagem pública. De acordo com ele, as sensações não são objetos (entidades) e os predicados de sensações não são denotações de sensações em sua extensão, mas devem ser entendidos pelo comportamento<sup>16</sup>. Sua atenção, portanto, volta-se para o modo de como sujeitos se comportam frente a determinadas cenas em uma situação de fala intersubjetiva. Isto é, os termos acerca das sensações não podem ter uma semântica denotacional/denotativa, uma vez que significaria dizer que o termo somente poderia ser aplicado ao sujeito. Para este autor, não há qualquer critério de identidade qualitativa entre as sensações, mas o conhecimento do significado das palavras deve ser dado a partir do seu uso intersubjetivo. Desta forma, toda investigação baseada no uso linguístico real procura encontrar o significado de um termo com base na sua aplicação, nas regras de uso do termo.

Nas obras de Wittgenstein, percebe-se uma grande preocupação com o uso de termos que descrevem sensações, em especial, ele discute sobre as cores:

Como ele vai saber que cor deve escolher quando ouve a palavra ‘vermelho’?  
 – Muito simples: “ele deve escolher a cor cuja imagem lhe ocorre quando ouve a palavra” Mas como ele irá saber o que significa e qual é “a que lhe ocorre quando ouve a palavra?”

Certamente há um processo como escolher a cor que lhe ocorre quando você ouve essa palavra. E a sentença “vermelho é a cor que ocorre quando você ouve a palavra vermelho” é uma definição. (WITTGENSTEIN, 1974, # 33)

Segundo Wittgenstein investigar qual é o significado de uma palavra não é uma preocupação psicológica, em tentar definir se a palavra “vermelho” traz essa cor à mente quando falada, mas sim é uma investigação gramatical do efeito de uma palavra.

O que surge diante de nossas mentes quando ouvimos uma palavra é certamente algo característico do significado. Mas, o que surge diante de minha mente é um exemplo, uma aplicação da palavra. E esse surgir diante da mente realmente não consiste em uma imagem estar presente sempre que emito ou ouço a palavra, mas no fato de que, quando me perguntam o significado da palavra, as aplicações da palavra *ocorrem* a mim. (WITTGENSTEIN, 1974, #75)

Podemos sustentar, então, que alguém compreende o que um símbolo significa quando o usa de maneira adequada. Mas como podemos falar do significado de sentenças que remetem a sensações? Como é possível que falantes aprendam a associar o mesmo significado à mesma expressão? E mais, como nós sabemos que eles fazem isso? Como ocorre, por exemplo, de duas pessoas associarem a palavra “amarelo” a certas cenas? Em Wittgenstein, essas questões são respondidas a partir do behaviorismo filosófico, porém surgem, a partir das respostas behavioristas, novas indagações acerca da possibilidade de utilizar um discurso mentalista (psicológico ou neurológico) para esclarecer o significado e o uso de expressões acerca de sensações.

O que se percebe, desde as discussões mais antigas, é que o problema acerca das sensações teria duas faces: uma biológica (da ciência) e uma conceitual (problema filosófico). Frente à perspectiva das concepções mais contemporâneas, observa-se que as propostas científicas devem estar fundamentadas em uma relação com a linguagem, cognição e comportamento. Dado que a linguagem não é algo fixo, ela não comporta em si todas as variações possíveis, além do mais, o próprio sujeito interage com o mundo e altera-o. Alguns autores reforçam o aspecto pragmático da linguagem, que implica que o significado de uma sentença é apenas parcialmente determinado pela sentença ela mesma, o que a torna dependente também das situações e fatores contextuais.

## Referências

COHEN. Color properties and color ascriptions: A relationalist manifesto. *The Philosophical Review*, 113(4), 2004, 451-506.

\_\_\_\_\_. Color Relationalism and Color Phenomenology. In. *The Red and the Real: An Essay on Color Ontology*. Oxford: University Press, Oxford. Forthcoming, 2009.

CHOMSKY. A Review of B. F. Skinner's Verbal Behavior. *Language*, 35, n. 1, p. 26-58, 1959.

\_\_\_\_\_. *Language and mind*. Enlarged edition. New York: Harcourt Brace Jovanovich, Inc., 1968 (1972).

- \_\_\_\_\_. *Reflections on language*. New York: Pantheon Books, Random House, 1975.
- DANCY, Jonathan. *Epistemologia Contemporânea*. Trad. Tereza Louro Pérez. Rio de Janeiro: Edições 70, 1985.
- DAVIDSON, Donald. Meaning, Truth and Evidence. In: BARRETT, R.; GIBSON, R. *Perspectives on Quine*. Oxford: Blackwell, p. 68–79, 1990.
- \_\_\_\_\_. The Inscrutability of Reference (1979), reprinted in D. Davidson, 1984, pp. 227–241.
- FODOR, Jerry A. *The modularity of mind*. The MIT Press, Cambridge, Mass, 1983.
- FREGE. *Lógica e filosofia da linguagem*. Tradução de Paulo Alcoforado. São Paulo: Cultrix/USP, 1978.
- HATFIELD, Gary. Sense-data and the mind-body problem. In *Perception and Reality: From Descartes to the Present*, ed. by Ralph Schumacher. Berlin: Mentis Verlag, 2004, 305–31.
- HOROWITZ, Amir. Contents just are in the head. *Erkenntnis* 54: 321–344, 2001.
- HUME, David. *Tratado da natureza humana: uma tentativa de introduzir o método experimental de raciocínio nos assuntos morais*. Tradução Déborah Danowski. São Paulo: Editora UNESP; Imprensa Oficial do estado, 2001. (Original publicado em 1739).
- KRIPKE, S. *Naming and Necessity*. Oxford: Basil Blackwell, 1980.
- LOCKE, J. *Ensaio acerca do entendimento humano*. 5. ed. Tradução de Anair Aiex. São Paulo: Nova Cultural, 1991. (Os Pensadores). (Original publicado em 1690).
- PUTNAM, Hilary. The meaning of ‘meaning’. In: *Mind, Language and Reality*. Philosophical Papers. Vol. 2. Cambridge: Cambridge University Press, 1975, pp. 215–271.
- QUINE. El Naturalismo, o el vivir por los propios médios. In. *Acerca del conocimiento científico y otros dogmas*, Paidós, Barcelona, 2001, p. 127–142.
- \_\_\_\_\_. *From stimulus to science*. Cambridge MA: Harvard University Press, 1995.
- \_\_\_\_\_. Methodological Reflections on Current Linguistic Theory. In: *Semantics of natural language*. Davidson and Harman, Reidel Publishing Company: Dordrecht – Holland, p. 442 – 454, 1972.
- \_\_\_\_\_. Mind and Verbal Dispositions. In: *Mind and language*. Samuel Guttenplan, Oxford: Oxford University Press, p. 83 – 95, 1975.
- \_\_\_\_\_. Naturalism; Or, Living Within One’s Means. In: *Dialectica*, v. 49, p. 251 – 261, 1995.
- \_\_\_\_\_. *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.
- \_\_\_\_\_. States of Mind. In: *The Journal of Philosophy*, v. LXXII, n. 1, p. 5–8, 1985.
- \_\_\_\_\_. *The roots of reference*. La Salle, III: Open Court, 1974.
- \_\_\_\_\_. *Word and Object*. Cambridge, Mass: Technology Press of the Massachusetts Institute of Technology, 1960.
- SEARLE, John R. Indeterminacy, Empiricism and the First Person. In: *The Journal of Philosophy*, v. LXXXIV, n. 3, p. 123–146, March 1987.

- \_\_\_\_\_. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: CUP, 1983.
- \_\_\_\_\_. Intrinsic Intentionality. *Behavioral and Brain Sciences* 3, p. 450-456, 1980.
- \_\_\_\_\_. *Mente, linguagem e sociedade*. Rio de Janeiro: Editora Rocco, 2000.
- \_\_\_\_\_. *Mind: a Brief Introduction*. Oxford: Oxford University Press, 2004.
- \_\_\_\_\_. *Speech acts: an essay in the philosophy of language*. Cambridge: Cambridge University Press, 1994.
- \_\_\_\_\_. *The Rediscovery of the Mind*. Cambridge Mass., London: MIT Press, 1992.
- SELLARS, Wilfrid. *Empirismo e filosofia da mente* (Stein, S.I.A., Trad.). São Paulo: Vozes, 2008. (Original publicado em 1956).
- WITTGENSTEIN, L. *Gramática filosófica*. Trad. BORGES, L. C. São Paulo: Edições Loyola, 2003. (Original publicado em 1974).
- \_\_\_\_\_. *Investigações filosóficas*. Trad. MONTAGNOLI M. G. Petrópolis: Vozes, 2004. (Original publicado em 1953).

## Notas

<sup>1</sup> Cf. FREGE, *Lógica e Filosofia da Linguagem*, 1978.

<sup>2</sup> Cf. LOCKE, *Ensaio acerca do Entendimento Humano*, 1991.

<sup>3</sup> Introspecção - método de observação e estudo, por uma dada pessoa, de seus próprios processos mentais. Faz-se aqui uma alusão aqui a Jerry Fodor, representante contemporâneo do mentalismo, que configura a mente dessa maneira em seu livro *The Modularity of Mind* (1983).

<sup>4</sup> Apesar de manter o discurso sobre estímulos. *Stimulus meaning* é o conceito semântico fundamental em Quine, e é definido em termos de ‘estimulação’, ou seja, é concebido como um par ordenado, composto por disposições a assentir ou dissentir uma sentença em consequência a uma estimulação presente, e é o que ativa a disposição, não o que a instaura. Deve-se conceber a estimulação não como um dado acontecimento particular e precisamente fechado, mas sim, como um universal, uma forma de evento que é repetível (1960, p. 33).

<sup>5</sup> O significado de uma expressão para Quine não é o objeto a qual a expressão se refere, seja ele uma ideia, uma proposição, um corpo físico, ou uma forma Platônica, mas, somente pode-se caracterizar o significado de uma expressão por descrições comportamentais.

<sup>6</sup> Naturalismo: propõe aplicar a todas as disciplinas filosóficas a metodologia das ciências naturais, afirmando que na realidade só existem coisas físicas e que o único método legítimo de estudo seria o das ciências naturais. Somente assim, pode-se verificar ou falsificar as hipóteses de maneira intersubjetiva. “A filosofia naturalista forma um contínuo com a ciência natural; se propõe clarificar, organizar e simplificar os conceitos mais amplos e básicos, assim como analisar o método científico e a evidência dentro do marco da ciência mesmo. A fronteira entre a filosofia naturalista e o resto da ciência é somente uma vaga questão de grau” (Quine, *El Naturalismo, o el vivir por los propios médios*, 1995, p. 135).

<sup>7</sup> Quine, *From Stimulus to Science*, 1995, p. 89.

<sup>8</sup> Segundo Chomsky o termo “gramática” é frequentemente usado ambiguamente para referir o sistema internalizado de regras e a descrição do linguista sobre ele, pois ela é uma representação em termos de propriedades fonéticas, semânticas e sintáticas. Entretanto, “gramática” em seus trabalhos

deve ser entendida como um sistema de regras que especifica a relação de som-significado para uma dada linguagem, ou seja, ela é um grupo de leis e princípios que determinam um grupo de sentenças com forma fixa e significado determinado. Essa gramática também pode ser chamada de “gramática gerativa”, já que se pode dizer que a gramática de uma linguagem gera um grupo infinito de “descrições estruturais”, sendo cada descrição um objeto abstrato de um tipo que determina um som particular, um significado particular, gerando assim as expressões de dada linguagem (CHOMSKY, 1972, p. 104). Chomsky ainda acredita que essa gramática é uma realização peculiarmente humana.

<sup>9</sup> É um argumento hipotético criado por Searle para refutar os teóricos da Inteligência Artificial (*Strong AI*), do funcionalismo, bem como do behaviorismo. Baseia-se na presunção de que a sintaxe (gramática) não é garantia de existência da semântica (sentido).

<sup>10</sup> “Nós podemos resumir o naturalismo biológico de maneira muito simples em duas proposições: (1) cérebros causam a mente; (2) mentes são características de nível superior do cérebro” (SEARLE, 1994, p. 545).

<sup>11</sup> Putnam, *The Meaning of the ‘Meaning’*, 1975, p. 223 – 227.

<sup>12</sup> Na tentativa de determinar o significado de uma palavra não seria correto, por exemplo, sustentar que “o significado de X é z”, onde ‘z’ é o nome do referente de “o significado de X”.

<sup>13</sup> Cf. DANCY, *Epistemologia Contemporânea*, 1985, p. 98.

<sup>14</sup> Interpretação é considerada aqui como um ato mental que toma as percepções e as organiza de modo a serem úteis para alguma finalidade. Cf. FODOR, Jerry A. *The modularity of mind*, 1983.

<sup>15</sup> Os Esquimós diferenciam no mínimo 15 nuances da cor branca, de forma a poderem identificar os tipos de neves por sua cor. Isso devido à própria necessidade de sobrevivência.

<sup>16</sup> Cf. WITTGENSTEIN, *Investigações Filosóficas*, 2004 (#253- 315).

# INDEXICALIDADE NO PORTUGUÊS BRASILEIRO: UMA ANÁLISE SEMÂNTICA BASEADA EM MUDANÇAS DE CONTEXTO\*

LOVANIA ROEHRIG TEIXEIRA

*Mestranda do Programa de Pós-graduação em Linguística - UFSC*

lovaniateixeira@gmail.com

**Resumo:** Este artigo tem por objetivo apresentar uma das mais importantes abordagens semântico-formais sobre os indexicais em língua natural, a abordagem kaplaniana, e algumas das principais reformulações dessa abordagem, postuladas por Schlenker, baseada na existência de “monstros” em linguagens não-formais. Procura-se demonstrar em quais aspectos as concepções desses dois autores se contrapõem e quais são os seus méritos ao analisar sentenças com indexicais. Por isso, na busca por um tratamento semântico mais abrangente, defende-se a abordagem baseada em mudanças de contexto que propõe a manipulação do contexto usado para a fixação dos indexicais em língua natural por um operador-monstro. Além disso, analisamos dados do português brasileiro que confirmam a existência de monstros e que mudanças no contexto não são privilégio de linguagens formais como Kaplan previa.

**Palavras-chave:** indexicais, operadores-monstros, semântica, mudanças de contexto.

## Introdução

O termo ‘indexical’ foi usado primeiramente por Pierce (1902) para designar um tipo específico de sinal, aquele que tem uma relação direta ou real com seu objeto. Desde então, o termo, em filosofia e em semântica, foi ganhando contornos diferentes e atualmente engloba as expressões que dependem do contexto para receber um valor semântico. A fim de restringir esse conceito devemos nos perguntar de qual contexto um termo indexical depende, já que contexto é um construto amplo cuja definição depende da teoria considerada, e também o que significa exatamente dependência contextual. Para um semanticista formal, o contexto é uma espécie de ênupla com algum tipo de informação que foi formalizado primeiramente por Kaplan (1989), e serve para dar conta especificamente dos indexicais e seu valor semântico. No entanto, é necessário salientar que Kaplan (1989) aborda os indexicais como expressões dependentes somente do contexto de proferimento (c\*), isso se dá por Kaplan utilizar um conjunto de concepções da lógica modal padrão (KRIKPE, 1980) que trata os indexicais como termos diretamente referencias e designadores rígidos.

Assim, uma vez que no cenário semântico-formal temos Kaplan (1989), de um lado, afirmando que, necessariamente, o contexto que deve ser considerado para a atribuição de valor semântico dos indexicais é o contexto de proferimento ( $c^*$ ), de outro lado, temos autores que afirmam que os termos indexicais podem ser avaliados em contextos diferentes do contexto de proferimento, *i.e.*, em  $c'$ , aqui podemos citar, entre outros, Predelli (2008), Anand (2006) e Schlenker (1999, 2003, 2010), cujas ideias apresentaremos neste texto. Buscaremos mostrar de que modo as reformulações propostas por Schlenker atingem a abordagem clássica para os indexicais proposta por Kaplan. Mostraremos, também, que a abordagem de Kaplan (1989) em algumas situações é muito restritiva em relação ao comportamento dos indexicais em língua natural, apesar dessa teoria ter muitos méritos, alguns dos quais serão apontados na primeira parte deste texto. Na segunda parte, trataremos das bases da reformulação schlenkeriana e, especificamente, em que aspectos ela se choca com a teoria clássica (KAPLAN, 1989). Em seguida, analisaremos algumas expressões indexicais do português brasileiro (PB) com o objetivo de fortalecer os argumentos arrolados por Schlenker a favor de sua reestruturação da abordagem de Kaplan.

## 1. Abordagem semântico-formal kaplaniana

É consenso no tratamento dos indexicais que a abordagem de Kaplan é uma das mais importantes do ponto de vista filosófico e semântico. Tal abordagem tornou-se conhecida a partir do trabalho *Demonstratives*, publicado em 1989, no qual o autor trata tanto de indexicais demonstrativos quanto de indexicais puros. Segundo o autor, indexicais demonstrativos são termos que necessitam, além das regras linguísticas, de alguma forma de ostensão para que tenham seu valor semântico determinado. Desse modo, se não houver um gesto que auxilie a determinação do referente da expressão indexical, o termo é considerado *incompleto*, o que podemos notar a partir da sentença em (1), em que sem uma pista gestual em anexo ao demonstrativo grifado, é muito difícil determinar seu referente.

(1) *Essa cadeira está com o encosto quebrado* (apontando para uma das cadeiras na sala de cinema).

Em contraste com os demonstrativos, os indexicais puros não necessitam de apontamento para que tenham seu valor semântico determinado. Segundo Kaplan (1989, p. 491), qualquer gesto que acompanha o indexical puro serve somente para enfatizá-lo ou é irrelevante. Na sentença em (2), um gesto do falante, por exemplo, para si mesmo, não tem papel fundamental para a determinação do referente do indexical 'eu', já que não há dúvida que o termo designa o falante do contexto.

(2) *Eu* não pretendo colaborar com a presidente.

Segundo Kaplan (1989), os indexicais são termos diretamente referenciais, pois esses elementos contribuem com um indivíduo numa proposição (quando se trata de um termo singular). Dessa maneira, se tivermos uma sentença como em (3) proferida pelo João, a proposição gerada por essa sentença será como a representação dada em (4).

(3) *Eu* estou doente.

(4) ⟨João, estar-doente⟩

Aliando-se à ideia de indexicais como termos diretamente referenciais, temos a concepção de que os termos indexicais são designadores rígidos, *i.e.*, um termo é um designador rígido se para todos os mundos possíveis  $w$ , o termo em questão designa o mesmo objeto  $x$  em todos os mundos  $w$  em que  $x$  existe, similarmente ao tratamento dado aos nomes próprios pela teoria da referência indireta (KRIPKE, 1980).

Na teoria kaplaniana, um termo indexical passa a ser um designador rígido a partir do proferimento da expressão ou sentença. Assim, é no momento do proferimento que os indexicais fixam o seu valor semântico, ou seja, a sentença em (3) gera a proposição singular em (4) a partir do momento em que a sentença é proferida por João. A partir disso, o referente do indexical ‘eu’ será o João para todos os mundos possíveis daquele contexto.

*Grosso modo*, o momento de proferimento das expressões indexicais inaugura o contexto relevante para a fixação de seu valor semântico e das proposições que os contêm. Especificamente para a teoria de Kaplan, o contexto relevante sempre é o contexto de proferimento ( $c^*$ ).

A concepção de contexto do ponto de vista formal da teoria kaplaniana se dá a partir de uma ênupla que possui coordenadas (as coordenadas do contexto) como: agente, ouvinte, local, tempo e mundo do contexto, *i.e.*, ⟨ $c_a$ ,  $c_h$ ,  $c_l$ ,  $c_t$ ,  $c_w$ ⟩. É a partir dessas coordenadas que as expressões indexicais recebem seu valor semântico: um referente no caso de um termo singular; uma proposição singular no caso de uma sentença.

Antes de analisarmos mais sentenças é necessário expor os outros conceitos que se relacionam ao contexto, quais sejam, o que é um contexto próprio ou bem-formado de acordo com Kaplan (1989) e a questão do significado dos indexicais. Veremos isso na sequência.

Para dar conta das sentenças com indexicais que são verdadeiras *a priori*<sup>1</sup>, mas não são necessariamente verdadeiras<sup>2</sup>, além de considerar que os indexicais são diretamente referenciais, Kaplan também estipulou que contextos próprios ou bem-formados devem respeitar certas condições como as dadas em (5)<sup>3</sup>.

(5) *Restrições sobre contextos bem-formados:*

$c$  é um contexto próprio sse o agente de  $c$  está localizado em  $c$  no tempo de  $c$  e no mundo de  $c$  ( $c_a$  está em  $c_t$  e em  $c_l$  em  $c_w$ ).

Além das restrições sobre os contextos de avaliação dos indexicais, Kaplan (1989) postula duas funções para lidar com esses termos, o *caráter* e o *conteúdo*: o caráter é uma função de contextos para conteúdos e o conteúdo é uma função de mundos possíveis para valores semânticos (o conteúdo corresponde,  *grosso modo*, ao sentido fregeano).

A partir desses conceitos (contexto próprio, caráter e conteúdo), podemos analisar as sentenças (6) e (7) (i) dados os contextos de proferimentos ( $c'$ ,  $c''$ ,  $c'''$ ) com as suas coordenadas preenchidas, ou seja, com informações de quem proferiu, quando, em que local, em qual mundo, e (ii) dado, também, que cada contexto possui o mundo relevante com o mesmo índice ( $w'$ ,  $w''$ ,  $w'''$ )<sup>4</sup>, e nesse mundo é dada a extensão do predicado 'ser-inteligente'.

(6) *Maria* é inteligente.

(7) *Eu* sou inteligente.

$c'$ = ⟨João, 10/10/2011, Florianópolis, $c'_w$ ⟩;	$w'$ = 'ser inteligente' = {Maria, Pedro, Ana};
$c''$ = ⟨Maria, 10/10/2011, Florianópolis, $c''_w$ ⟩;	$w''$ = 'ser inteligente' = {João, Pedro};
$c'''$ = ⟨Pedro, 11/10/2011, Porto Alegre, $c'''_w$ ⟩;	$w'''$ = 'ser inteligente' = {João, Maria};

A proposição que a sentença em (6) expressa, ⟨Maria, ser-inteligente⟩, será verdadeira sse o indivíduo denotado pelo nome próprio, Maria, está no conjunto de indivíduos denotados pelo predicado 'ser-inteligente' nos mundos de avaliação. Assim, o valor semântico de (6) não depende do contexto de proferimento, pois para qualquer contexto as condições de verdade serão as mesmas, já que temos um nome próprio que tem o caráter como uma função constante (não varia de contexto para contexto) e conteúdo também (indivíduo relevante não varia de mundo para mundo); logo, (6) é verdadeira nos mundos  $w'$  e  $w'''$ .

Por outro lado, em relação à sentença (7) com o indexical 'eu', que tem a função de caráter variável (varia de contexto para contexto) e o conteúdo como uma função constante (fixado um caráter, o conteúdo permanece o mesmo para os mundos de avaliação), o valor da ênupla de coordenadas é muito importante, já que é a partir do contexto de proferimento que os indexicais recebem um valor semântico.

Considere, primeiramente, a sentença em (7) no contexto  $c'$ . De acordo com a ênupla de coordenadas do contexto  $c'$ , quem proferiu a sentença foi o João. Assim, o caráter da sentença (7) é ⟨ $c_a$ , ser-inteligente⟩ e o seu conteúdo é ⟨João, ser-inteligente⟩; uma proposição estruturada que deve ser avaliada nos mundos possíveis  $w'$ ,  $w''$ ,  $w'''$ . Assim, a sentença em (7) é verdadeira nos mundos  $w''$  e  $w'''$ , pois são os mundos em que João pertence à extensão do predicado 'ser-inteligente'. No entanto, se consideramos outro contexto, por exemplo, o contexto  $c''$ , teremos condições de verdade diferentes, pois o valor semântico do indexical 'eu' será Pedro, e a sentença será verdadeira em mundos diferentes do que os do contexto  $c'$ , *i.e.*, será verdadeira somente nos mundos  $w'$  e  $w''$ .

Assim, pode-se notar que, para Kaplan (1989), os indexicais têm um caráter que depende do contexto de proferimento e, por consequência, um conteúdo resultante do contexto do proferimento. Com base nisso, entre outras coisas, Kaplan afirma que o contexto de fixação de um indexical em língua natural é sempre o contexto de proferimento ( $c^*$ ) e que somente o conteúdo dos indexicais pode ser manipulado por operadores; ele afirma ainda que operadores que manipulam o contexto de fixação dos indexicais só podem ser adicionados em linguagem formal e chama tais operadores de “monstros”.

Na seção seguinte, veremos que é justamente contra a ideia da não-existência de operadores-monstros capazes de atuar sobre o contexto em língua natural que Schlenker (1999, 2003, 2010) reúne argumentos e dados. Na sequência, corroboramos as ideias de Schlenker a partir de exemplos do PB.

## 2. Abordagem baseada em mudança de contexto

Observando dados do amárico, língua oficial da Etiópia, Schlenker (1999, 2003, 2010, 2011) notou que os verbos de atitude e *dicendi* permitem que os indexicais sob seu escopo sejam avaliados a partir do contexto reportado ( $c'$ ) e não a partir do contexto de proferimento ( $c^*$ ) da sentença, como afirmara Kaplan (1989), o que caracteriza a mudança de contexto. Considere a situação em (8):

(8) *Situação*: João disse: “Eu sou um herói”.

(a) Amárico: João Jägna näNN yt-lall

*João herói eu-sou dizer-3<sup>sg.m</sup>*

João<sub>i</sub> disse que eu<sub>i</sub> sou um herói.

Na situação em (8) o indexical de primeira pessoa (nä) está sendo avaliado a partir do contexto que está sendo reportado na sentença do amárico, ou seja, o contexto  $c'$ , e não em relação ao contexto de proferimento  $c^*$ . Note que o item ‘João’, à esquerda do verbo *dicendi*, é fixado no contexto  $c^*$ ; já os itens à direita do verbo de dizer (principalmente o indexical) recebem valor semântico a partir do contexto  $c'$ , *i.e.*, o indexical tem como referente o João (falante do contexto que está sendo reportado) e não o falante do contexto de proferimento; o que seria o caso para uma sentença do PB, já que o relato de uma situação como (8) em PB, seria feito como a sentença dada em (9), em que temos um elemento anafórico retomando João. Nesse caso, a sentença e o pronome ‘ele’ recebem valor semântico do contexto  $c^*$ , não há mudança no contexto.

(9) O João<sub>i</sub> disse que ele<sub>i</sub> é um herói.

Além do amárico, Schlenker apontou algumas línguas como o ewe<sup>5</sup> e bafut<sup>6</sup>, do continente africano, em que há elementos monstruosos após o verbo de atitude, os pronomes logofóricos. Sendo assim, quando há pronomes logofóricos

depois de verbos de atitude ou *dicendi*, há, por consequência, mudança no contexto relevante (de  $c^*$  para  $c'$ ).

Considere os exemplos em (10), de Schlenker (1999, p. 32) do ewe, em que o elemento 'yè' "[...] can only appear when the matrix element it 'corefers' with is in 2nd or in 3rd person[...]". Desse modo, em (10) o pronome logofórico só pode ser utilizado quando ele está se referindo ao sujeito da oração principal. Podemos comparar esse comportamento com o elemento 'e' da sentença em (11), que não é um logofórico e apresenta a mesma morfologia estando, ou não, sob o escopo de um operador de atitude.

(10) João be yè-dzo (pronomes logofóricos do ewe).

*João dizer LOG-sair*

João disse que ele (João) saiu.

(11) João be e-dzo (indexicais do ewe).

*João dizer ele/ela sair.*

João disse que ele ( $\neq$  de João) saiu.

Note que há uma diferença entre os indexicais monstruosos do amárico e os pronomes logofóricos do ewe e do bafut: os pronomes logofóricos são marcados formalmente (morfologia do item) como indexicais monstruosos, enquanto que os indexicais monstruosos do amárico não são diferentes morfologicamente dos indexicais não-monstruosos (a diferença de interpretação entre os indexicais do amárico só pode ser reproduzida na forma lógica).

Considere a representação abaixo, no amárico, quando há um verbo de atitude (AT), o indexical que é avaliado em um contexto mudado ( $c'$ ) tem a mesma morfologia do item que é avaliado no contexto de proferimento ( $c^*$ ). No ewe, por sua vez, quando o indexical é fixado num contexto mudado ( $c'$ ) e está sob o escopo do operador de atitude o item usado é 'yè' e quando o indexical é fixado no contexto de proferimento ( $c^*$ ) usa-se 'e'.

Amárico	Ewe
AT [ $c'$ (nä)]	AT [ $c'$ (yè)]
$c^*$ (nä)	$c^*$ (e)

Ao analisar essas línguas, Schlenker observou que quando temos um verbo de atitude há mudança no contexto relevante e os indexicais recebem valor semântico a partir de um contexto diferente do contexto de proferimento. Assim, Schlenker postula que os verbos de atitude e os verbos *dicendi*<sup>7</sup> são operadores-monstros, ou seja, podem manipular o contexto em que os indexicais sob seu escopo são avaliados; se isso estiver correto, há monstros em línguas naturais, *contra* Kaplan (1989).

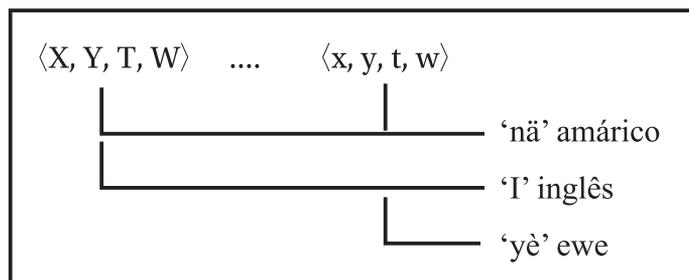


Figura 1: Esquema de fixação dos indexicais ao contexto (SCHLENKER, 1999, p. 13)

O esquema na Figura 1 descreve a dependência contextual dos indexicais nas sentenças subordinadas, ou sentenças de relato, em que um verbo de atitude vem seguido por indexicais que podem ser monstruosos ou não. Vemos que o indexical 'nä' do amárico pode depender tanto do contexto de proferimento, representado pela ênupla em letras maiúsculas, ou do contexto reportado, representado pela ênupla em minúsculas, quando está sob o escopo do operador de atitude. O indexical do inglês 'I' só pode ter seu valor determinado a partir do contexto de proferimento, estando ou não sob o escopo do operador de atitude; o pronome logofórico do ewe, 'yè', só pode receber um valor semântico do contexto que está sendo reportado (c'), quando está sob o escopo de um operador de atitude.

A partir dessas análises, Schlenker (2003, 2010) estabeleceu quatro critérios mínimos para que indexicais sejam considerados monstruosos, são os seguintes:

- (i) o elemento é semanticamente dependente do contexto;
- (ii) o elemento deve ser um indexical estrito (puro cf. Kaplan), *i.e.*, lexicalmente especificado como dependente do contexto;
- (iii) deve-se excluir a possibilidade de que a oração encaixada esteja sendo citada;
- (iv) o elemento investigado deve ser interpretado *de se* quando é avaliado a partir do contexto mudado<sup>8</sup>.

Esses critérios serão explorados de maneira mais detalhada na seção seguinte, na qual também analisaremos indexicais do PB a fim de verificar se eles podem ser ou não ser indexicais monstruosos quando se seguem de um verbo de atitude ou *dicendi*.

### 3. Análise de dados do PB de acordo com a proposta de mudança de contextos

Analisaremos nesta seção expressões temporais indexicais do PB em relação aos quatro critérios estabelecidos por Schlenker (2003, 2010). As expressões indexicais monstruosas devem ser avaliadas em um contexto diferente do contexto

de proferimento quando estão sob o escopo de um operador de atitude ou *dicendi*. As expressões que serão objeto de análises são ‘em dois dias’, ‘o dia depois de amanhã’, ‘o dia antes de ontem’, ‘dois dias atrás’. Essas expressões temporais apresentam comportamento semelhante quando não estão sob o escopo de um operador de atitude ou *dicendi*. Considere as sentenças de (12) a (15).

(12) Maria encontrou uma nota de 50 *dois dias atrás*.

(13) Maria encontrou uma nota de 50 *no dia antes de ontem*.

(14) A reunião dos pais acontecerá *em dois dias*.

(15) A reunião dos pais acontecerá *no dia depois de amanhã*.

As sentenças (12) e (13) quando proferidas, *e.g.*, na terça-feira, se referem ambas ao domingo; as duas têm orientação temporal para o passado. As sentenças em (14) e (15) têm orientação temporal para o futuro, no entanto, elas também são dependentes do contexto, pois quando são proferidas, *e.g.*, no domingo, se referem ambas à terça-feira.

A partir disso, munidos das expressões indexicais e dos critérios para que um indexical seja considerado um indexical com comportamento monstruoso, analisaremos na sequência as expressões segundo os quatro critérios citados na seção anterior.

### 3.1 Dependência do contexto

O primeiro passo para que um termo seja considerado monstruoso, ou que possa receber um valor semântico a partir de um contexto diferente do contexto de proferimento ( $c^*$ ), é assegurar que esse termo/expressão seja um indexical, *i.e.*, seja uma expressão dependente do contexto. Para verificarmos esse aspecto considere, primeiramente, as sentenças em (16) e (17).

(16) O concerto da Orquestra Sinfônica será realizado *em dois dias*.

(17) O concerto da Orquestra Sinfônica será realizado *no dia depois de amanhã*.

Caso Pedro encontre essas sentenças escritas em um aviso colado em uma parede, não saberá quando ir ao concerto, visto que não há como fixar a referência dos indexicais temporais. A situação em questão seria facilmente resolvida se ele soubesse quando as sentenças foram produzidas. Com isso notamos que sem fixar o momento de produção das sentenças não se pode atribuir o referente adequado aos indexicais. Por conta desse comportamento, as expressões grifadas nas sentenças em (16) e (17) são indexicais.

Agora considere as sentenças (18) e (19) escritas em um aviso, sem data, na porta de um laboratório.

(18) As pessoas que realizaram exames *dois dias atrás* devem retornar ao laboratório.

(19) As pessoas que realizaram exames *no dia antes de ontem* devem retornar ao laboratório.

Do mesmo modo que as sentenças em (16) e (17), as sentenças em (18) e (19), quando escritas em um aviso fixado na porta de um laboratório, não conseguem ter sua coordenada de tempo preenchida por um dia específico ou adequado à situação, pois no aviso não há identificação do dia da produção das sentenças. No entanto, novamente, basta que o aviso esteja datado, por exemplo, com 15 de outubro de 2011, para que haja a determinação do referente das expressões indexicais. Assim, ambas as sentenças passam a se referir ao dia 13 de outubro de 2011.

Portanto, as expressões do PB ‘em dois dias’, ‘dois dias atrás’, ‘o dia depois de amanhã’ e ‘o dia antes de ontem’ respondem positivamente ao primeiro dos critérios vistos acima, *i.e.*, são dependentes de um contexto.

### 3.2 Indexical estrito

De acordo com Schlenker (2003, 2010), para que um indexical tenha um comportamento monstruoso, ou tenha a capacidade de ter seu contexto de avaliação mudado por um operador-monstro, como o verbo de atitude, ele precisa ser um indexical estrito (lexicalmente especificado como indexical). Por conta disso, não deve ter a capacidade de se referir, por exemplo, a algum momento saliente no discurso, ou seja, ter um comportamento parecido com o das expressões anafóricas ‘dois dias depois’ e ‘dois dias antes’, que se ligam a um evento ou momento saliente.

Quando os anafóricos, como ‘dois dias depois’ e ‘dois dias antes’, apresentam usos indexicais, eles são considerados indexicais por acidente (cf. SCHLENKER, 2003, p. 64). Os usos indexicais dos anafóricos se dão quando eles se referem a um momento saliente no tempo que coincide com o momento do proferimento. Assim como podemos observar em (20) e (21).

Tais expressões, contudo, podem ser analisadas como portadoras de um argumento temporal, parafraseado como ‘dois dias depois de *t*’ e ‘dois dias antes de *t*’, como exemplificado por (22) e (23), não sendo, portanto, indexicais estritos.

(20) O João disse, no domingo, que me encontraria *dois dias depois*.

(21) O João disse, ontem, que me viu no centro *dois dias antes*.

(22) O João recebeu o dinheiro *dois dias depois* do fim do prazo.

(23) O João recebeu o dinheiro *dois dias antes* do fim do prazo.

Tendo em vista o comportamento dos indexicais acidentais, podemos testar se as expressões ‘em dois dias’, ‘dois dias atrás’, ‘o dia depois de amanhã’ e ‘o dia antes de ontem’, que passaram pelo primeiro critério estabelecido por Schlenker (2003), são indexicais estritos. Para tanto, essas expressões temporais não podem formar sentenças aceitáveis quando estiverem substituindo os anafóricos

legítimos. Isso ocorre, pois “[...] [os indexicais] can never have unambiguously anaphoric readings” (SCHLENKER, 2010, p. 21).

A verificação se dá da seguinte forma: no conjunto de sentenças de (24) a (27)<sup>9</sup> (que são sentenças gramaticais ou naturais com expressões anafóricas) testamos todas as expressões temporais que estão sob investigação, colocando-as no lugar dos anafóricos. Para que os indexicais sejam classificados como estritos, eles devem tornar as sentenças estranhas ou agramaticais, visto que os indexicais estritos estariam exercendo uma função que não lhes é peculiar, qual seja, referirem a um momento saliente. Considere as sentenças a seguir com base nas sentenças em (22) e (23).

(24) \* O João recebeu o dinheiro *em dois dias* do fim do prazo.

(25) \* O João recebeu o dinheiro *no dia depois de amanhã* do fim do prazo.

(26) \* O João recebeu o dinheiro *dois dias atrás* do fim do prazo.

(27) \* O João recebeu o dinheiro *no dia antes de ontem* do fim do prazo.

Dado que as sentenças de (24) a (27) não são bem-formadas, as expressões do PB que estão sob análise podem ser consideradas indexicais estritos, respondendo positivamente ao segundo critério estabelecido por Schlenker.

### 3.3 Não estar sendo citado após o verbo de atitude ou *dicendi*

O terceiro critério para que uma expressão possa ser classificada como indexical monstruoso é ela estar sendo usada, e não citada, após o verbo de atitude ou verbo *dicendi*. A partir disso, é de se esperar que se estabeleça uma relação, na sentença, entre um indivíduo e uma proposição, aspecto que é característico em um relato de proferimento; e não se firme uma relação entre um indivíduo e um conjunto de palavras, como é permitido no discurso direto.

Dessa maneira, para que excluamos a possibilidade de que a oração encaixada, que contém a expressão indexical, esteja sendo mencionada basta que tenhamos na sentença analisada o complementizador ‘que’, pois “Em português, o complementizador ‘que’ não pode aparecer antes de uma citação direta [...], mas é obrigatório em citações indiretas [...]” (CUNHA, 2006, p. 74). Assim sendo, considere as sentenças em (28) e (29), que serão os nossos objetos de análise.

(28) O João tem me dito repetidamente, ao longo dos anos, *que* devolverá meu dinheiro...

(i) em dois dias.

(ii) no dia depois de amanhã.

(29) O João tem me dito repetidamente, ao longo dos anos, *que* estava doente...

(i) dois dias atrás.

(ii) no dia antes de ontem.

Como pode ser observado em (28) e (29), as sentenças apresentam em suas estruturas o complementizador ‘que’ e, de acordo com Cunha (2006), esse item só está presente quando vem seguido por uma citação indireta; desse modo, pode-se afirmar que as orações encaixadas estão sendo usadas. Caso as orações encaixadas que contêm as expressões indexicais estivessem sendo citadas, elas deveriam ser como as dadas em (28’) e (29’), ou seja, sem o complementizador e usando as mesmas palavras que foram proferidas pelo João (discurso direto).

(28’) O João tem me dito repetidamente, ao longo dos anos: “Eu devolverei teu dinheiro...

(i) em dois dias.

(ii) no dia depois de amanhã.

(29’) O João tem me dito repetidamente, ao longo dos anos: “Eu estava doente...

(i) dois dias atrás.

(ii) no dia antes de ontem.

Em resumo, conforme vimos a partir de (28) e (29), as expressões em análise estão sendo usadas, e não citadas, após o verbo de atitude ou verbo de dizer.

### 3.4 Interpretação *de se* após o verbo de atitude ou *dicendi*

Os indexicais que se comportam como a teoria kaplaniana prevê não podem ser avaliados em um contexto diferente do contexto de proferimento  $c^*$  quando estão sob o escopo de um operador de atitude. Os indexicais monstruosos ou modificados, por sua vez, são mais versáteis e podem ou não ter seu contexto de avaliação manipulado pelo operador de atitude. Assim, os indexicais monstruosos podem tanto se referir ao contexto de proferimento ( $c^*$ ), como os bem-comportados, quanto se referir ao contexto reportado ( $c'$ ), o comportamento que os torna monstruosos. Por conta disso, as expressões temporais indexicais que buscamos, quando estão sob o escopo de um operador de atitude, se ligam tanto à coordenada de tempo do contexto de proferimento ( $c^*_t$ ) quanto à coordenada de tempo do contexto reportado ( $c'_t$ ).

Além disso, dentre as expressões que estamos analisando procuramos qual delas será interpretada *de se* em relação ao tempo do contexto reportado, ou seja, o tempo será calculado a partir do tempo em que o agente do contexto reportado proferiu as sentenças. A partir disso, teremos os primeiros indexicais monstruosos identificados no PB, ou seja, indexicais sendo avaliados em contextos mudados.

Separamos as expressões indexicais por orientação temporal para analisá-las aos pares. Assim, ‘em dois dias’ e ‘no dia depois de amanhã’ formam um par, pois representam um momento posterior ao tempo do contexto; ‘dois dias atrás’ e ‘o dia antes de ontem’ formam outro par, pois representam um momento anterior ao tempo do contexto.

Considere a sentença em (30) com a expressão indexical ‘em dois dias’.

(30) O João tem me dito repetidamente, ao longo dos anos, que devolverá meu dinheiro *em dois dias*.

Em (30), a expressão indexical temporal deve ser avaliada em relação ao contexto reportado ( $c'$ ); em relação aos vários proferimentos ou promessas do João, o agente de  $c'$ . Ou seja, o que o João vem me falando ao longo dos anos é a sentença (31).

(31) Eu devolverei teu dinheiro *em dois dias*.

Assim, a expressão indexical 'em dois dias' que é usada no meu proferimento não toma a coordenada de tempo do contexto do meu proferimento,  $c_t^*$ <sup>10</sup>. Na verdade, a expressão indexical temporal se refere à coordenada de tempo do contexto reportado por mim,  $c'_t$ , ou seja, se liga ao tempo dos contextos em que João promete me pagar em dois dias. Por isso, 'em dois dias' é considerado um indexical que apresenta um comportamento monstruoso, porque há modificação no contexto considerado, *input* da função carácter, e deve ser interpretado *de se* (o tempo deve ser avaliado a partir do tempo do proferimento do João ( $c'$ )).

Analisaremos, agora, o comportamento da expressão indexical 'o dia depois de amanhã'. Para isso, considere a sentença em (32).

(32) O João tem me dito repetidamente, ao longo dos anos, que devolverá meu dinheiro *no dia depois de amanhã*.

Em (32), a expressão indexical só pode ser avaliada em relação ao proferimento, ou em relação ao contexto  $c^*$ ; isso mostra que 'no dia depois de amanhã' é uma expressão indexical kaplaniana que respeita a estipulação do autor, na qual a avaliação de um indexical sempre se dá em relação ao contexto  $c^*$  e não há mudança no contexto relevante, e, por consequência, sempre se liga à coordenada  $c_t^*$ .

Podemos tornar as coisas mais claras utilizando datas. Por exemplo, se eu profiro (32) no dia 25 de dezembro de 2010, o ouvinte preencherá o conteúdo do indexical 'no dia depois de amanhã' com o dia 27 de dezembro de 2010, porque ele toma como ponto de partida o tempo  $c_t^*$ , dia 25 de dezembro de 2010, dia do meu proferimento. Desse modo, o que João vem me dizendo todos os anos é algo como a sentença em (33).

(33) Eu devolverei o teu dinheiro *em 27 de dezembro de 2010*.

Por conta disso, a expressão indexical 'no dia depois de amanhã' não é um indexical monstruoso<sup>11</sup>, já que a avaliação da expressão está ligada ao tempo do contexto de proferimento, como Kaplan (1989) havia estipulado para os indexicais em sua teoria.

Partimos para a análise do par 'dois dias atrás' e 'o dia antes de ontem'. Considere, primeiramente, a sentença em (34) com a expressão 'dois dias atrás'.

(34) O João tem me dito repetidamente, ao longo dos anos, que ele estava doente *dois dias atrás*.

Em (34), a expressão indexical temporal ‘dois dias atrás’ deve ser avaliada em relação aos vários contextos reportados e em relação ao tempo desses contextos,  $c'_t$ . Assim, ‘dois dias atrás’ recebe seu valor semântico em relação aos vários proferimentos do João, que é o agente do contexto reportado  $c'$ , pois o que o João vem me falando todos os anos é algo como a sentença em (35).

(35) Eu estava doente *dois dias atrás*.

Desse modo, a expressão indexical ‘dois dias atrás’ que é usada no meu proferimento não se refere à coordenada de tempo do meu proferimento  $c^*_t$ . Na verdade, a expressão toma a coordenada de tempo do contexto reportado por mim,  $c'_t$ , os contextos em que João queixa-se de ter estado doente dois dias atrás. Assim sendo, a expressão temporal ‘dois dias atrás’ é monstruosa porque há modificação no contexto considerado, *input* da função caráter, e, por isso, é interpretada *de se*, ou seja, o tempo é avaliado a partir do tempo do contexto reportado  $c'_t$ , do tempo do proferimento do João.

Finalmente, vamos analisar o comportamento da expressão ‘no dia antes de ontem’. Para isso, considere a sentença em (36).

(36) ?<sup>12</sup> O João tem me dito repetidamente, através dos anos, que estava doente *no dia antes de ontem*.

(36) é estranha porque a expressão indexical ‘no dia antes de ontem’ só pode ser avaliada em relação ao momento do proferimento, e em relação a  $c^*_t$ . Nesses termos, (36) causa estranhamento pois a situação e o uso da palavra ‘repetidamente’ impõem uma interpretação de proferimentos iterativos, ou seja, força uma interpretação relacionada a proferimentos que se deram em contextos diferentes de  $c^*$ , ou seja, em contextos mudados. Isso pode ser um indício de que essa expressão só pode ser avaliada adequadamente quando toma a coordenada de tempo do contexto de proferimento  $c^*_t$ . Podemos então afirmar que o contexto de avaliação de ‘no dia antes de ontem’ não pode ser manipulado por operadores de atitudes, o que demonstra que essa expressão é um indexical kaplaniano.

Portanto, aplicando os critérios que Schlenker (2003, 2010) oferece para encontrar monstros e indexicais monstruosos chegamos à seguinte conclusão: o operador de atitude é um MONSTRO, pois é um quantificador sobre contextos, e quando ‘em dois dias’ e ‘dois dias atrás’ estão sob seu escopo podem ter seus contextos de avaliação manipulados.

Na Figura 2, apresentamos um resumo dos achados no PB, ou seja, as expressões que estavam em nossa listagem inicial (‘no dia depois de amanhã’, ‘no dia antes de ontem’, ‘em dois dias’ e ‘dois dias atrás’) estão classificadas em relação aos critérios para indexicais monstruosos. As expressões indexicais temporais do PB que receberam ‘OK’ para todos os critérios são monstruosas. Os indexicais kaplanianos receberam ‘OK’ para os critérios (a), (b) e (c), só não se

adequaram ao critério (d), já que não podem ser interpretados *de se* em relação ao tempo.

<b>Crítérios:</b>	<i>dois dias atrás em dois dias</i>	<i>o dia antes de ontem o dia depois de amanhã</i>
a) dependente do contexto	OK	OK
b) indexical estrito	OK	OK
c) não ser citado	OK	OK
d) interpretado <i>de se</i>	OK	Não
	<b>indexical monstruoso</b>	<b>indexical kaplaniano</b>

Figura 2: Critérios para indexicais modificados no PB.

## Considerações Finais

Conforme foi visto, a abordagem de Kaplan (1989) para os indexicais dá conta de várias interpretações relacionadas a esses itens por meio da consideração de contextos no caráter e de mundos possíveis no conteúdo. Ela também modela de maneira satisfatória um contexto bem-formado para a avaliação dos indexicais. No entanto, a ideia de Kaplan (1989) de que em língua natural não existe um elemento, um operador-monstro, que pode manipular o contexto dos indexicais não resiste aos exemplos trazidos por Schlenker (1999, 2003, 2010, 2011), que mostrou, através de línguas como o amárico e com os pronomes logofóricos do ewe, que os indexicais podem ter seu contexto manipulado por verbos de atitude permitindo que o contexto em que os indexicais se fixam mude de  $c^*$  para  $c'$ . Isso transforma os verbos de atitude e os verbos *dicendi* em operadores-monstros, pois são capazes de mudar o contexto em que os indexicais sob seu escopo se fixam. Além disso, os pronomes logofóricos têm papel importante na concepção de que há operações semânticas envolvidas nos casos de indexicais avaliados em contexto mudados ( $c'$ ) e não implicações pragmáticas. É correto afirmar isso, pois os elementos que são avaliados em outro contexto, por exemplo, o  $c'$  (contexto reportado), são marcados morfologicamente no ewe e no bafut, apesar de em outras línguas naturais, como no PB e no amárico, essa característica não estar pronunciada. Por conta disso, no PB e no amárico o mesmo item lexical é usado em contextos modificados ( $c'$ ) e em contextos de proferimento ( $c^*$ ), a diferença semântica só pode ser observada na linguagem formal. Isso ocorre, por exemplo, com os elementos que foram analisados e que mostraram um comportamento monstruoso quando estavam no escopo do verbo *dicendi*, que foram os indexicais temporais 'em dois dias' e 'dois dias atrás'. Assim sendo, podemos postular que há indexicais monstruosos e que o verbo 'dizer' do PB, que pode fixar a avaliação dos indexicais em um contexto diferente do contexto de proferimento  $c^*$ , é um

operador-monstro. Além disso, é correto afirmar que a previsão de Kaplan (1989) de que as línguas naturais não possuem indexicais avaliados em contextos mudados não se aplica ao PB.

## Referências

- ANAND, P. *De De Se*. Ph. D. Dissertation, Massachusetts Institute of Technology: USA, 2006.
- BASSO, R. M.; TEIXEIRA, L. R. Monstros no discurso (meta)ficcional. *Revista Letras*, 2011. Aceito para publicação.
- CUNHA, M. A. F. O complemento dos verbos de enunciação. *Linguística*, v. 2, n. 1, p. 69-84, 2006. UFRJ.
- FREGE, G. *Lógica e filosofia da linguagem*. Tradução: Paulo Alcoforado. São Paulo: Editora da Universidade de São Paulo, 1978.
- KAPLAN, D. Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In: *Themes from Kaplan*. ALMOG, J; PERRY, J; WETTSTEIN, H. (ed.). New York: Oxford University Press, p. 481-563, 1989.
- KRIPKE, S. *Naming and Necessity*. Harvard University Press: USA, 1980.
- PIERCE, C. S. Logic as semiotic: the theory of signs. 1902. (Reprinted in: BUCHLER, J. (Ed.) *Philosophical Writings of Pierce*. New York: Dover, 1955, p. 98-119.).
- PREDELLI, S. I am not here now. *Analysis*. v. 58, p. 107- 115, 1998.
- \_\_\_\_\_. Modal Monsters and Talk about Fiction. *Journal of Philosophical Logic*. v. 37, p. 277-297, 2008.
- SCHLENKER, P. A plea for monsters. *Linguistics and Philosophy*. Netherlands: Kluwer Academic Publishers. v. 26, p 29-120, 2003.
- \_\_\_\_\_. *Indexicality and De Se Reports*. Disponível em: <[https://files.nyu.edu/pds4/public/Schlenker-Indexicality\\_and\\_De\\_Se.pdf](https://files.nyu.edu/pds4/public/Schlenker-Indexicality_and_De_Se.pdf)>. Acesso em: Outubro de 2010.
- \_\_\_\_\_. *Propositional Attitudes and Indexicality: A Cross-Categorical Approach*. Massachusetts Institute of Technology: USA, 1999.
- \_\_\_\_\_. Indexicals. 2011. Disponível em: <<https://files.nyu.edu/pds4/public/Indexicals.pdf>>. Acesso em: Maio 2011.

## Notas

\*Agradeço a leitura atenta e as sugestões do Prof. Dr. Renato Miguel Basso, que em muito contribuíram para a clareza do texto. Todos os equívocos são de minha responsabilidade.

1 'Eu estou aqui agora' é uma sentença verdadeira *a priori* já que não precisamos saber nada sobre o mundo para saber que ela é verdadeira. Assim, o contexto é responsável por fixar os valores semânticos dos indexicais e pelas verdades *a priori*.

2 'Necessariamente, eu estou aqui agora' não é necessariamente verdadeira, pois basta imaginar qualquer situação contrafactual, por exemplo, você, leitor, poderia estar agora participando de um campeonato de pesca no Amazonas, e não necessariamente lendo este texto onde quer que você esteja. Assim, os mundos possíveis são responsáveis pelas verdades necessárias das sentenças com indexicais.

3 Os *contextos próprios* se contrapõem aos *contextos impróprios* analisados por Predelli (1998), entre outros. Um exemplo bem conhecido de contexto impróprio são as mensagens gravadas em secretárias eletrônicas ('Eu não estou aqui no momento'). Nessa mensagem, o agente do contexto ( $c_a$ ) não está no local ( $c_l$ ) e no tempo ( $c_t$ ) do contexto.

4 Para simplificar, considera-se que o mundo possível relevante é o mundo possível do contexto. Isso não é um procedimento necessário, mas nos auxilia na sistematização das explicações realizadas.

5 Uma das línguas faladas em Togo, Benim e Gana.

6 Língua oficial de Camarões.

7 Schlenker afirma que operadores de atitude são monstros e utiliza exemplos de sentenças que mostram que nessa classe de operadores devem ser também incluídos os verbos *dicendi*. Para evitar esse comprometimento teórico, discriminamos os grupos de verbos sem assumir a posição de Schlenker.

8 Note que os quatro critérios estabelecidos por Schlenker (2003, 2010) formam dois grupos: os critérios de (i) a (iii) são exigidos por todas as línguas naturais que possuem indexicais sendo avaliados em contextos mudados, independentes da teoria considerada. No entanto, o critério (iv) não é uma exigência das línguas naturais e sim da teoria de Schlenker (2003, 2010) e dos que compartilham dela como Anand (2006), entre outros. Um caso de indexical avaliado em contextos mudados, mas que não possui a leitura *de se* é o do indexical 'eu' da sentença 'Eu acho que eu poderia morar mais perto dela' dada como resposta por uma atriz, após o término da peça de teatro, à pergunta 'O que você acha que poderia ser feito para que a peça melhorasse?', que é discutido por Basso e Teixeira (2011).

9 O símbolo '\*' indica em (24)-(27) a agramaticalidade da sentença.

10 Assumimos que o tempo do contexto de proferimento  $hic et nunc = t^* = c^*_t$ , do mesmo modo que o tempo do contexto reportado  $= t' = c'_t$ .

11 Esta expressão tem apenas interpretação *de re*.

12 Nesse caso, o sinal de interrogação marca que a sentença foi julgada como "não-natural" por alguns falantes do PB.

# ALGUNAS CUESTIONES SOBRE LA TRIANGULACIÓN DAVIDSONIANA COMO RESPUESTA A LAS DIFICULTADES EVIDENCIADAS EN EL ARGUMENTO SOBRE “SEGUIR UNA REGLA”

RICARDO NAVIA

*FHUCE - Depto. de Historia de la Filosofía*

*Universidad de la República - Uruguay*

naviamar@adinet.com.uy

Resumo: Desde el argumento wittgensteiniano sobre “seguir una regla”, la filosofía teórica contemporánea está enfrentada al problema de lograr una explicación sobre la naturaleza y surgimiento de la normatividad y especialmente de la idea de verdad. Todo lo cual está en el fondo de cuestiones decisivas sobre los problemas del escepticismo, del realismo y del externalismo semántico, que nuestro grupo (CSIC) ha estado investigando. Davidson ha intentado una respuesta centrada en lo que él ha llamado “triangulación” que procura a la vez ser una explicación no-metafísica y no reductivista de la normatividad.

Según Davidson es en la situación de triangulación que los seres humanos fijan los significados como base de su capacidad lingüística y forjan el concepto de verdad objetiva como base del pensamiento. Sin embargo, varias objeciones (Pagin, 2001; Talmage, 1997; Verheggen, 1997) han surgido hacia la idea de la triangulación, tanto porque se cuestiona la necesidad y el efecto real de la misma, como, sobre todo, porque se ha objetado que requiere de cierta comunicación cuando, se suponía, ella debía explicar los significados; o, que requiere de pensamientos cuando se postulaba que ella creaba las condiciones para el surgimiento mismo del pensamiento.

Asimismo, se plantea la dificultad de la desambiguación (identificación) de la causa común (o mismidad del referente), aspecto fundamental en la discusión del realismo, problema que se agudiza desde que Davidson acepta la inescrutabilidad de la referencia. Estimamos que ambos problemas son de vital importancia porque, de resolverse, permitirían avanzar en el tratamiento de las dificultades evidenciadas por el argumento de seguir una regla que hacen a una concepción no metafísica de la normatividad.

En este artículo nos proponemos:

- 1 - reconstruir sumariamente la situación de triangulación, su necesidad y sus efectos.
- 2 - revisar las acusaciones de circularidad; acusaciones que me parecen originadas en cierta incompreensión de la naturaleza de la triangulación
- 3 - explorar el significado filosófico de la misma para una concepción esencialmente social del lenguaje, así como para una explicación no reductiva ni metafísica de la normatividad básica del lenguaje y del conocimiento que nos permitiría salir de la situación en que nos ubica el célebre argumento wittgensteiniano.

**Palabras claves:** triangulación, objetividad, desambiguación, seguir una regla, normatividad

"La triangulación no es una cuestión de una persona que capte un significado que esté ya ahí, sino de una actuación que . . . confiere un contenido al lenguaje"(Davidson, D. Subjetivo, intersubjetivo y objetivo, p. XV)

"Tenemos muchos vocabularios para describir la naturaleza cuando la consideramos no mental y tenemos un vocabulario mentalista para describir el pensamiento y la acción intencional; de lo que carecemos es de una manera de describir lo que ocurre entre estas dos cosas." (Davidson, D. La aparición del pensamiento, p. 182)

## **Introducción**

El argumento de la triangulación es introducido desde 1982 por Davidson en varios escritos con el propósito de mostrar las condiciones mínimas necesarias para el surgimiento del pensamiento y del lenguaje, evidenciando el carácter esencialmente social de ambos.

Sin embargo, en la medida en que el argumento desarrolla y defiende dos ideas previas fundamentales en Davidson, a saber: que las creencias requieren del concepto de verdad objetiva y que el concepto de verdad objetiva requiere interacción lingüística con otras personas, el argumento tiende a probar bastante más que una descripción genética del surgimiento del pensamiento. En efecto, involucra también: a) el origen y la función del concepto de objetividad, que ya había inquietado a la filosofía del lenguaje desde el problema wittgensteiniano de "seguir una regla"; b) una posible respuesta al problema de la doble ambigüedad de la referencia de las palabras; c) un fundamento para su externismo semántico; d) una concepción naturalista de las actitudes proposicionales que intenta no ser reductivista y e) la mencionada idea sobre la naturaleza esencialmente social del pensamiento y del lenguaje.

La sumaria mención de estas derivaciones evidencia el destacado significado filosófico de la triangulación, que rebasa por cierto los límites de esta comunicación. En ella nos proponemos limitarnos a:

- I - Una reconstrucción del argumento de la triangulación
- II - Un análisis de algunas objeciones planteadas al argumento, como forma de tratar de evidenciar su verdadera naturaleza y su rol argumental
- III - Algunas consideraciones sobre dos aspectos del significado filosófico del mencionado argumento, a saber:
  - a - el origen y la función del concepto de objetividad
  - b - la viabilidad de basarse en él para una concepción naturalista pero no reductivista de las actitudes proposicionales

## I - Reconstrucción de la triangulación como argumento (1)

La introducción de la triangulación para dar cuenta de las condiciones mínimas de surgimiento del pensamiento y del lenguaje se inicia en su artículo “Rational Animals” de 1982 y adquiere una especial relevancia en “La emergencia del pensamiento” y en “Las condiciones del pensamiento” de 1988.

(A) Davidson comienza el tratamiento de este tema explicitando la diferencia entre la capacidad ya existente en los animales para discriminar entre distintos objetos y la capacidad humana para usar conceptos. Empieza considerando una forma primitiva de triangulación: en ella ciertos animales simplemente aprenden a reaccionar frente a cierto estímulo del medio. Aquí hay un esbozo de triangulación porque se aprende a mirar algo en común y a advertirlo frente a una reacción también común de un par de la especie. Pero estas criaturas no manejan conceptos ni actitudes proposicionales desde que no están describiendo nada sino en todo caso identificando una señal. (2001b, 294).

(B) Un nivel algo superior está dado por el aprendizaje infantil de una primera lengua. (*Second Person*, p.262) Allí, al niño se le estimula a producir cierta emisión lingüística constante frente a la presencia de objetos o eventos similares (la emisión “mesa” ante la presencia de mesas). Un aspecto importante de este aprendizaje es que implica la capacidad del niño de advertir que ciertos objetos son similares a otros en ciertos aspectos relevantes (aún habiendo alguna diferencia en aspectos no jerarquizados en ese contexto). La captación de esta similaridad va a jugar un papel fundamental en el surgimiento del pensamiento.

Sin embargo, repara Davidson, hay aquí un problema a aclarar respecto a la localización del estímulo. ¿Cuál sería la razón para sostener que el objeto mesa es el estímulo que provoca la emisión “mesa” y no las ondas lumínicas que llegan al ojo del niño o aún meramente las modificaciones periféricas en los receptores del sujeto? Escribe:

... de hecho, si hemos de elegir, parece que la causa próxima de la conducta posee los mejores títulos para recibir la denominación de estímulo, . . . puesto que estas pautas(estimulativas) siempre producen la conducta mientras que las mesas sólo la producen en condiciones favorables.(2)

Pero en seguida se responde:

¿Por qué, sin embargo, parece natural decir que el perro responde al timbre y el niño a las mesas?

Encontramos similares las emisiones de la palabra “mesa” que el niño lleva a cabo, y las cosas del mundo que acompañan a esas emisiones y que clasificamos juntas de forma natural son precisamente mesas. (3)

Claro que uno tiene derecho a preguntarse de dónde emana esa “naturalidad” de la clasificación. Dicha “naturalidad” parece justificarse no por una similaridad en las pautas acústicas o visuales que no las podemos observar fácilmente y si lo pudiéramos hacer sería difícil determinar qué las hace similares(4), si no que la localización del objeto se concreta sobretudo en virtud

de una triple similaridad que se produce en este proceso de triangulación. El aprendiz reúne como similares a una clase de objetos (mesas por ejemplo), el maestro hace el mismo agrupamiento y el intérprete a su vez agrupa como similares las respuestas del niño a las mesas. (5)

Es este uno de los lugares clásicos de definición de la triangulación donde expresa:

. . . una línea parte del niño en dirección a la mesa, otra línea parte de nosotros en dirección a la mesa y la tercera va de nosotros al niño. El estímulo relevante se halla allí donde convergen las líneas del niño a la mesa y de nosotros a la mesa. (6)

A esta altura, el niño no tiene aún capacidad conceptual ni lenguaje, pero según nuestro autor está sentando las bases para su proceso constitutivo. Esta interacción básica entre dos personas y un objeto así identificado posibilita la definición de un contenido para el pensamiento que comienza a constituirse. Aquí es fundamental el papel de la segunda persona. Sin otra persona con la cual "triangular" no habría contenidos para los pensamientos, no habría pensamientos. Si solo tuviéramos criaturas aisladas, sus respuestas por complejas que pudieran ser no constituirían actitudes proposicionales, no podrían tener la capacidad de "reaccionar a" o "pensar en" algo situado a cierta distancia en lugar de sobre sus propios receptores.

En *La emergencia del pensamiento*, Davidson dice que para una persona aislada, la causa estaría doblemente indeterminada. En el sentido de que desconoceríamos que aspecto de la causa de una preferencia es relevante para su significado (la forma, el color, el movimiento, etc.) y también ignoraríamos si la causa son las pautas estimulativas en los órganos sensoriales (estímulo proximal) o el objeto que a distancia provoca ese tipo especial de estimulaciones (estímulo distal).

Como dice Davidson "*El mundo del solipsista puede tener cualquier dimensión, lo que equivale a decir que no tiene dimensión alguna, que no es un mundo*" (7) Esto es, la interacción hasta aquí descrita en la triangulación es una condición necesaria para la posibilidad del pensamiento, pero no es aún una condición suficiente para el mismo.

El mero hecho de que se trate de reacciones compartidas no indica el aspecto relevante. Por si solas, las reacciones compartidas a los objetos del medio no determinan las causas más que las reacciones individuales, es el llamado problema de la "ambiguación". Como dice C. Verheggen "*Para que S signifique algo mediante palabras . . . S debe también estar en posición de reconocer al menos algunas de las causas como tales. . . pensar en esas causas como existiendo independientemente del hablar sobre ellas*" (8) Y eso solo es posible mediante la comunicación, solo ella consigue la desambiguación.

(C) Sin embargo, la triangulación al mismo tiempo tiene otro importantísimo efecto sobre los participantes: los introduce en la idea y posibilidad del error. Cuando dos personas correlacionan sus reacciones con objetos o eventos del

mundo exterior y con las reacciones de otras personas, se ha establecido una norma o estándar para reaccionar de cierta forma ante determinados objetos o hechos, de modo tal que se está “acordando” que si no se reacciona de ese modo, se está cometiendo un error. Se han introducido las categorías de “acertado” y de “no acertado” (antecedentes inmediatos de la categorías de verdadero y falso). Davidson considera que esa es la capacidad distintiva de la conceptualización y del pensamiento. En uno de sus últimos escritos sobre la relación entre el lenguaje, la creencia y el concepto de objetividad, Davidson escribió:

Si yo creo que lo que estoy viendo es una jirafa, estoy empleando el concepto de una jirafa en el sentido de que estoy clasificando lo que veo. . . Saber esto es saber que algunas clasificaciones son verdaderas y algunas son falsas. Si yo no me doy cuenta de la posibilidad de la clasificación equivocada, no estaría teniendo un pensamiento proposicional (Responses, 2003, p.698)

(D)Sin embargo, nuestro autor se apresura a aclarar que la constitución del pensamiento no está aún completada: no alcanza con que el objeto haya sido localizado mediante la intersección de dos tipos de respuestas de similaridad. En el nivel de triangulación primitiva los significados aún no pueden ser fijados porque para ello no alcanza con una mera asociación entre causas reiteradas de una preferencia y dicha preferencia, las respuestas son compartidas pero aún comportan cierta ambigüedad.

Para que “dos perspectivas privadas converjan para marcar una posición en el espacio intersubjetivo” <sup>(9)</sup>, es necesario además que las dos reconozcan que están coincidiendo en la definición de ese punto o estándar. Para ello, “la única forma de saber que la criatura que ocupa el segundo punto está reaccionando al mismo objeto es saber que ésta está pensando acerca del mismo objeto” (Sinclair, 714a), para ello se requiere de la comunicación lingüística. <sup>(10)</sup>

Davidson sostiene entonces que nuestra captación de objetos y pautas en un espacio objetivo – esto es, independiente de nuestros pensamientos – “surge a través del reconocimiento, provisto por la comunicación lingüística, de otro ser que está compartiendo pensamientos sobre objetos” y pautas que existen independientemente de esos pensamientos. Como sostiene Sinclair:

Las conexiones causales son descriptibles en términos intencionales, y por tanto abren espacio para el error, solo cuando las líneas causales convergen mediante la interacción social, y solo cuando las respuestas a tal convergencia son mutuamente relevantes para las criaturas involucradas en esa interacción. El estatus intencional de los pensamientos de un hablante, que surge mediante el proceso de comunicación lingüística, es así dependiente de las dinámicas causal y social presentes en la triangulación.<sup>(11)</sup>

De este modo, Davidson expone su tesis de que pensamiento y lenguaje son mutuamente dependientes y se desarrollan simultáneamente, idea que desarrollaría hacia el final de su obra (Davidson 2001, p. 293).

## **II) Objeciones a la triangulación**

De las varias objeciones que se han planteado al argumento de la triangulación, me interesa ver la que sostiene que la triangulación quedaría sin sentido desde el momento en que requiere de pensamiento cuando se suponía que ella creaba las condiciones para el surgimiento mismo del pensamiento. Cuando Pagin (2001) formula esta objeción está aludiendo a que según Davidson la triangulación solo se consume cuando la persona dispone del concepto de verdad objetiva, que ya constituye un pensamiento o incluso que el concepto de verdad objetiva requiere del concepto de objeto o evento, lo cual también supone haber ya accedido al nivel del pensamiento.

Creo que esta acusación se podría encarar por al menos tres vías. La principal respuesta tendría que ver con el carácter no reductivo del naturalismo de Davidson que no pretende explicar sin hiato cómo se genera el pensamiento desde el no-pensamiento, sino que solamente intenta identificar las condiciones mínimas y las circunstancias en que éste surge.

Nuestro autor aclara especialmente que la triangulación no pretende explicar el pasaje desde reacciones que no implican pensamiento a comportamientos mediados por pensamiento proposicional, para empezar porque no disponemos de un vocabulario apto para tal explicación, si no que solo pretende mostrar las condiciones mínimas necesarias para tal transición <sup>(12)</sup>, con lo cual está estableciendo la base de su no-reductivismo.

Una respuesta complementaria sostendría que ambas ideas se alcanzan a la vez. Por último, una tercera posibilidad, no necesariamente excluyente con las anteriores, podría interpretar que el ingreso a la comunidad de pensamiento y lenguaje requiere de varios niveles de triangulación. En un primer nivel se registran similitudes, en un segundo nivel eso da lugar a un concepto (que sintetiza esas similitudes registradas), el que se consolida como tal con la captación de la intención de los pares, accediéndose así a la idea de acierto o error (independiente de la creencia) que caracteriza a la idea de objetividad.

Una segunda objeción casi paralela a la recién analizada señala que la triangulación requiere de interacción lingüística cuando se suponía que nos indicaba las condiciones mínimas para el surgimiento del lenguaje y especialmente para la determinación del significado de las preferencias y de los pensamientos.

Creo que ésta perplejidad puede resolverse si consideramos que el establecimiento de una pauta objetiva requiere que las dos personas conozcan lo que la otra piensa a través de la comunicación pero no necesariamente desde el comienzo esto se produce a través de una comunicación lingüística, sino de una conducta comunicativa que comienza a unirse significativamente a un objeto y cuando esto ocurre ya tenemos la significación y por tanto el lenguaje. Es decir, nuevamente aquí estamos sosteniendo que ambas logros se alcanzarían a la vez.

En el fondo, parece como que este tipo de objeciones se originan en un enfoque que presupone que existe alguna clase de prioridad entre los ordenes

del conocimiento objetivo, del conocimiento subjetivo y del conocimiento intersubjetivo; pero precisamente la novedad y la clave del enfoque de Davidson consiste en sostener que estos tres ordenes de conocimiento son interdependientes y se fundan a la vez. Además, en este enfoque están también operando otras dos tesis del autor: el holismo de lo mental y su antirreductivismo. Por su holismo, Davidson tiene conciencia de que es difícil decir algo sobre la aparición del pensamiento porque hay varios conceptos interdependientes involucrados en ello.<sup>(13)</sup> Por su antirreductivismo sabe que carecemos de un lenguaje que nos permita hablar con propiedad de las fases de transición.<sup>(14)</sup>

Dos pasajes de Davidson pueden cerrar esta respuesta a las objeciones:

¿...por qué según mi punto de vista es tan difícil intentar decir algo interesante y profundo acerca de la aparición del pensamiento? La razón es que hay muchos conceptos que debemos abordar para poder describir o hablar sobre el pensar, actuar por una razón, creer o dudar, todos los cuales dependen unos de los otros. Este es holismo de lo mental. (La aparición del pensamiento, p. 177)

... hay un estadio en que no hay pensamiento que está seguido de otro estadio en que sí lo hay. Describir la aparición del pensamiento sería describir el proceso que conduce del primero al segundo de estos estadios. Carecemos de un vocabulario satisfactorio para describir los estadios intermedios. (La aparición del pensamiento, p. 181)

## **A - Origen y función del concepto de objetividad**

Como ya vimos, es en la situación de triangulación que los seres humanos fijan los significados de sus pensamientos y preferencias como base de su capacidad lingüística y forjan el concepto de verdad objetiva como base de su pensamiento y de sus actitudes proposicionales en general.

Davidson define el concepto de objetividad como “la conciencia, no importa cuan inarticulada, del hecho de que lo que es pensado o dicho puede ser verdadero o falso”<sup>(15)</sup>. Recalca que solo decimos que alguien tiene una creencia si tiene el concepto de verdad objetiva, esto es, si es consciente de que esa creencia puede ser verdadera o falsa por razones que son independientes de sus creencias, y esto porque para concebir una proposición es necesario conocer lo que son y lo que no son sus condiciones de verdad.

La dedicación que Davidson presta desde 1990 a la reconsideración de los conceptos de verdad, pensamiento y objetividad en buena medida se centró en explicar la captación del concepto de verdad objetiva, cómo llegamos a reconocer la diferencia entre lo que creemos y lo que es el caso. (DD., 1995a, 204).

Reconoce en Wittgenstein ha quien adelantó la tesis de que “la fuente del concepto de verdad objetiva es la comunicación interpersonal”<sup>(16)</sup> en tanto el Argumento del Lenguaje Privado precisamente defiende que a menos que un lenguaje sea compartido con otras personas no hay modo de distinguir entre usarlo correctamente o usarlo incorrectamente. A partir de eso Davidson

pretende llevar esta tesis sobre el origen de la corrección lingüística al concepto general de objetividad.

Como ya señalamos, siguiendo la descripción de la triangulación: a partir de la constatación de similitudes forjamos y usamos conceptos para clasificar las cosas y eventos de nuestro entorno y a partir de ese momento, se abre la posibilidad de que lo que nuestros pensamientos y preferencias clasifican esté acorde o no con aquella categorización. Claro que dicha "constatación de similitudes" tiene que trascender el ámbito subjetivo de cada persona y ubicarse en un espacio público, lo cual se realiza por el mutuo reconocimiento comunicativo que cierra el triángulo visto.

Queda así en evidencia la necesidad de la triangulación para el surgimiento de la objetividad. Dice Davidson

"La única forma de saber que... la segunda criatura está reaccionando al mismo objeto que uno mismo es saber que la otra persona tiene en mente el mismo objeto para que dos personas sepan que sus pensamientos están relacionados así, se requiere que estén en comunicación" (La Segunda Persona, p.174)

Davidson enfatiza que es la posibilidad de aplicar mal el concepto (en una preferencia) y por tanto la posibilidad del error (proposicional) lo que distingue la conceptualización de la mera discriminación. De este modo nace un estándar (de objetividad) que está más allá de nuestras creencias. Dicho estándar, por su lado, se conecta con el concepto de verdad: puesto que saber aplicar esas clasificaciones o conceptos es saber bajo qué condiciones esos juicios son verdaderos. Pero, a su vez, el concepto de verdad se conecta con el de objetividad: porque conocer bajo qué condiciones un juicio es verdadero implica el reconocimiento del hecho de que un juicio es verdadero o falso por un criterio que es independientemente de nuestras creencias. (1995a, p.210)

Vale recordar que el momento fundamental de instituir una pauta, que partiendo de dos subjetividades se ubica empero en un espacio público con lo cual rebasa tales puntos de partida, solo puede cumplirse mediante un "reconocimiento mutuo" en torno a algo del entorno que solo puede verificarse a través de la comunicación. Escribe Davidson:

A menos que la línea de base del triángulo, la línea entre dos agentes, sea reforzada al punto donde ella pueda implementar la comunicación de contenidos proposicionales, no hay forma de que los agentes puedan hacer uso de la situación triangular para formar juicios acerca del mundo. Solo cuando el lenguaje está en su lugar pueden las criaturas apreciar el concepto de verdad objetiva (1999a, *The Emergence*, p.14)

El concepto de objetividad emerge cuando en la interacción lingüística llegamos a reconocer que otro individuo comparte con nosotros pensamientos sobre objetos y propiedades, que ambos reconocemos como que existen independientemente de nuestros pensamientos (Sinclair, p.714)

Sobre este paso final de la triangulación Davidson en “*Tres variedades del conocimiento*” se plantea la pregunta crucial de porqué esa medida interpersonal ha de constituir la única forma de establecer una norma objetiva, aún cuando así lo presuponga la comunidad. En lo personal, la respuesta en ese ensayo no me resulta del todo clara pero vale la pena analizarla. Para empezar aclara que va a responder a partir del análisis de la situación de interpretación radical. Quizás el pasaje más significativo es cuando afirma:

No tenemos razones para atribuir a una criatura la distinción entre lo que se piensa que es el caso y lo que es el caso a menos que esa criatura posea la norma que un lenguaje compartido *proporciona* ...(*Tres variedades del conocimiento*, p. 286)

Lo cual podría entenderse como que si cada sujeto correlacionara con el mundo a su antojo no habría comunicación; por tanto, por una especie de argumento trascendental, si hay comunicación exitosa es que los participantes se han sujetado a una pauta objetiva de algún modo alojada en el lenguaje y en la estructura total de creencias.

En su ensayo “*La segunda persona*” Davidson aborda el problema de la objetividad desde el punto de vista de la pauta objetiva que es requerida por el lenguaje, el terreno común que nos permita saber cómo interpretar las preferencias de otro, incluso saber – con buena probabilidad – cómo va a continuar un hablante del que ya hemos interpretado emisiones anteriores, la típica situación descrita en el argumento de “seguir una regla”.

El problema se torna agudo desde que Davidson sostiene en *A Nice Derangement of Epitaphs* y en *Second Person* que – si bien toda comunicación se da en un entorno social – no puede recurrirse a los conceptos de convención o regla, desde el momento en que pueden comprenderse interlocutores que parten de convenciones lingüísticas diferentes. Desarrolla la tesis de que comunicarse, hablar una lengua: “solamente requiere que cada hablante se haga intencionalmente interpretable al otro” (*La segunda persona*, p. 167). Plantea que la intención comunicativa cubre la condición de delinear de un modo “indefinidamente amplio” un estándar objetivo sobre el cual se puede acertar o errar. Permite de este modo una explicación alternativa a la comunitarista para el problema de la normatividad, clásica perplejidad que nos legó el argumento de “seguir una regla”.

El argumento ratifica la tesis Wittgenstein-Kripke de que la primera lengua no puede ser un lenguaje privado, pero en el argumento de Davidson, la naturaleza social del lenguaje es entendida de modo diferente. En el argumento de Kripke se requería de una segunda persona o de la comunidad para incorporar una rutina que en adelante sería el vínculo de comunicación; en cambio en el argumento de Davidson la interacción de al menos dos hablantes/interpretes es la fundante de la rutina. En Kripke una uniformidad social, externa pretendía explicar el entendimiento, lo cual reenvía al predicamento señalado por Wittgenstein (¿en base a qué pautas entenderíamos dicha rutina?), en cambio en

Davidson el entendimiento interactivo de hecho es el que funda la pauta que consolida un entendimiento previo, con lo cual no reaparecerá la pregunta porque la interpretación está ya dada y es anterior a todo hecho que demande interpretación.

## Bibliografía

- Davidson, Donald "Rational Animals", en *Subjective, Intersubjective, Objective*, Oxford, Clarendon Press, 2001, pp.95-105.
- Davidson, Donald "The Problem of Objectivity", en *Problems of Rationality*, Oxford, Clarendon Press 1995, pp. 3-18.
- Davidson, Donald "The Second Person", en *Subjective, Intersubjective, Objective*, Oxford, Clarendon Press, 2001, 107-121.
- Davidson, D., Las condiciones del pensamiento, en Davidson, D. *Mente, mundo y acción*, Barcelona, Paidós, 1992.
- Davidson, D., Tres variedades de conocimiento, en Davidson, D., *Subjetivo, intersubjetivo, objetivo*, Madrid, Cátedra, 2003.
- Davidson, D., La aparición del pensamiento, en Davidson, D., *Subjetivo, intersubjetivo, objetivo*, Madrid, Cátedra, 2003.
- Davidson, D., Responses in
- Davidson, D., Reply to Føllesdal, en Hahn, Lewis(ed.), *The Philosophy of Donald Davidson*, La Salle: Open Court, p. 729-32.
- Verheggen, Claudine 'Davidson's Second Person', *The Philosophical Quarterly*, 47 (1997), pp.361-9
- Verheggen, Claudine "Triangulating with Davidson", *The Philosophical Quarterly*, Vol. 57, nº 226, (2007), pp. 96-103.
- Talmage, C. "Meaning and Triangulation", *Linguistics and Philosophy*, 20 (1997) pp.139-45.
- Pagin, P. "Semantic Triangulation", en *Interpreting Davidson*, edited by Petr Kotatko, Peter Pagin and Gabriel Segal, 2001, Stanford: CLSI Publications, pp.199-212.
- Sinclair, Robert, The Philosophical Significance of Triangulation: Locating Davidson's Non-Reductive Naturalism, en *Metaphilosophy*, Vol. 36, nº5, Oct. 2005, pp.708-728

## Notas

<sup>1</sup> El estatuto filosófico de la triangulación como argumento que combina elementos trascendentales con descripción naturalista es todo un tema digno de estudio.

<sup>2</sup> Davidson, D., *Las condiciones del pensamiento*, p. 158

<sup>3</sup> Op. Cit. p.158

<sup>4</sup> Idem

<sup>5</sup> Op. Cit. P.158-59

<sup>6</sup> idem

<sup>7</sup> Idem p.159

<sup>8</sup> Verheggen, C.: *Triangulating with Davidson*, p. 99.

<sup>9</sup> Idem

<sup>10</sup> Op. Cit. p. 161

<sup>11</sup> Sinclair, R., *The Philosophical Significance of Triangulation*, p. 714

<sup>12</sup> Reply to Føllesdal, en Hahn, p.731

<sup>13</sup> Davidson, *La aparición del pensamiento*, p.177

<sup>14</sup> Op. Cit.,pág. 181

<sup>15</sup> Davidson, *The Problem of Objectivity*, pág. 4

<sup>16</sup> Davidson, *Tres variedades de conocimiento*, p. 286.

# A TESE DA DESIGNAÇÃO RÍGIDA

VALDETONIO PEREIRA DE ALENCAR

*Professor assistente na UFC*

valdetonio\_alencar@yahoo.com.br

A tese da Designação Rígida é uma tese sobre o comportamento semântico dos Nomes Próprios (NP's) em contextos modais defendida por Kripke (1980). Segundo ele, NP's são designadores rígidos, isto é, um NP designa o mesmo objeto em todos os mundos possíveis, ao contrário das Descrições Definidas (DD's), que possuem um comportamento semântico em contextos modais distinto dos respectivos NP's. DD's não designam de forma rígida (pelo menos, na maioria dos casos). Essa distinção semântica entre NP's e DD's é utilizada por Kripke para criticar o chamado "Descritivismo". O objetivo deste texto é explicitar a tese da designação rígida, apontar críticas e respostas. Eu pretendo, assim, analisar até que ponto Kripke é bem sucedido em sua refutação do Descritivismo.

Kripke aponta no prefácio de *Naming and Necessity* (1980, p. 3) que sua teoria dos NP's advém de sua semântica da lógica modal (1980, p. 3). O principal passo consistiu na formulação de um modelo de mundos possíveis.

Existem várias abordagens sobre mundos possíveis. O modelo de Kripke constitui apenas um dentre vários modelos alternativos. Três abordagens são mais destacáveis segundo Susan Haack (2002, pp. 253-4):

- a. Abordagem linguística: todo conjunto maximal consistente de sentenças, estabelece um mundo possível (HINTIKKA, 1969).
- b. Abordagem conceitualista: mundos possíveis são estabelecidos a partir do mundo atual. Como a abordagem linguística, todo mundo possível pode ser descrito através de um conjunto maximal consistente de sentenças. Contudo, nem todo conjunto maximal consistente de sentenças estabelece um mundo possível. As combinações possíveis estão limitadas pelos objetos (e propriedades) do mundo atual.
- c. Abordagem realista de David Lewis (1986): "...aceita o discurso sobre mundos possíveis em seu valor ostensivo, como um discurso sobre entidades reais, abstratas, inteiramente independentes de nossa linguagem ou pensamento" (HAACK, 2002, p. 254).

Através de sua abordagem conceitualista, Kripke define as noções modais de necessidade ( $\Box$ ) e possibilidade ( $\Diamond$ ) da seguinte maneira:

- Def. " $\Box$ ": " $\Box p$ " (é necessário que p) é verdade se e somente se "p" é verdadeiro em todos os mundos possíveis relativos à m1 (mundo atual).

- Def. “ $\diamond$ ”: “ $\diamond p$ ” (é possível p) é verdade se e somente se “p” é verdadeiro em pelo menos um mundo possível relativo à m1.

Ao utilizar essas definições, Kripke conseguiu contrastar o comportamento semântico em contextos modais das DD's em relação aos NP's. No prefácio de *Naming and Necessity*, Kripke utiliza a teoria descritivista de Russell para explicitar a tese da designação rígida. Russell defende duas teses:

- 1) NP's são abreviações de DD's. Essa tese russelliana está fundada na distinção entre conhecimento por familiaridade e conhecimento por descrição. Através de termos indexicais, podemos ter conhecimento por familiaridade de seus portadores. O meu conhecimento dos portadores dos NP's é meramente descritivo. Por isso, NP's são tratados como DD's, os portadores dos NP's são apenas feixes de dados sensoriais.
- 2) DD's são tratadas de forma quantificacional através da teoria das DD's de Russell (TD).

A conjunção das teses acima implica:

- 3) NP's são tratados de forma quantificacional.

Kripke concorda com a TD de Russell, mas questiona a tese 1. Tome-se a seguinte sentença:

(1) É necessário que Aristóteles seja Aristóteles.

Tome-se que o conteúdo de “Aristóteles” para um falante A seja unicamente o significado da DD “o mestre de Alexandre Magno” ou, se pensarmos a partir de Searle, poder-se-ia supor que esse conteúdo descritivo é o único associado por todos os membros de uma comunidade<sup>1</sup>. Por (1) e pela substituição de “Aristóteles” por “o mestre de Alexandre Magno”, temos as seguintes possibilidades:

- (1)\* É necessário que Aristóteles seja o mestre de Alexandre Magno.
- (1)\*\* É necessário que o mestre de Alexandre Magno seja Aristóteles.
- (1)\*\*\* É necessário que o mestre de Alexandre Magno seja o mestre de Alexandre Magno.

Aceitando-se que a substituição de termos co-referenciais preserva verdade (ou falsidade), o valor de verdade de (1), (1)\*, (1)\*\* e (1)\*\*\* deveria permanecer o mesmo, pois apenas substituímos termos co-referenciais e seguimos o critério epistêmico dos descritivistas. Segundo esse critério, o significado de um NP é dado pelo significado descritivo que um falante ou que uma comunidade associa. Além de possuírem o mesmo valor de verdade, as proposições expressadas pelas sentenças acima seriam a mesma proposição, já que “Aristóteles” expressa o significado de “o mestre de Alexandre Magno”. Essa afirmação não seria propriamente de Russell (para o qual termos sub-sentenciais não expressam

significado)<sup>2</sup>, mas Frege e Searle diriam que termos sub-sentenciais expressam significado.

Mesmo sendo possível que Aristóteles não fosse filósofo ou não fosse grego, Aristóteles continuaria a ser Aristóteles. A identidade de um objeto é preservada nos mundos possíveis nos quais o mesmo existe, por isso (1) é verdadeira. É intuitivo que o mestre de Alexandre Magno poderia ter sido outro, por isso é óbvio que (1)\* e (1)\*\* são falsas. *Ser o mestre de Alexandre Magno* não é (intuitivamente) uma propriedade necessária de Aristóteles. De forma semelhante, Kripke diz que (1)\*\*\* é falsa: como o mestre de Alexandre Magno poderia ter sido outro, então é possível que o mestre de Alexandre Magno não fosse o Mestre de Alexandre Magno. Kripke, pois, afirma:

...although someone other than the U.S. President in 1970 might have been the U.S. President in 1970 (e.g., Humphrey might have), no one other than Nixon might have been Nixon. In the same way, a designator rigidly designates a certain object if it designates that object wherever the object exists; if, in addition, the object is a necessary existent, the designator can be called *strongly rigid* (KRIPKE, 1980, pp. 48-9).

Não temos, assim, preservação de verdade de (1) para (1)\*, (1)\*\* ou (1)\*\*\* e menos ainda sinonímia entre as mesmas. Se a substituição *salva veritate* vale em contextos modais, então o significado dos NP's não é dado pelo conteúdo descritivo associado pelos falantes. NP's possuem um comportamento semântico distinto em relação às DD's: NP's são designadores rígidos, isto é, designam o mesmo objeto em todos os mundos possíveis nos quais esse objeto existe, enquanto DD's não apresentam essa constância. A referência desses termos muda através dos mundos possíveis. Por isso, NP's não expressam o significado de DD's. Um descritivista poderia replicar que o princípio de substituição *salva veritate* é válido apenas para contextos extensionais. Contudo, restaria explicar o porquê desse princípio não funcionar em contextos não extensionais. É interessante levantar duas críticas iniciais a Kripke, que permitem entender melhor a tese da designação rígida.

O NP "Aristóteles" designa o mesmo objeto em todos os mundos possíveis nos quais ele existe. Contudo, Aristóteles poderia não se chamar "Aristóteles". Nesse mundo possível, "Aristóteles" não se referiria a Aristóteles. Logo, "Aristóteles" não é um designador rígido. Bem, o argumento de Kripke não está interessado nas mudanças linguísticas existentes em outros mundos possíveis. A partir do mundo atual, estabelecem-se os mundos possíveis de Kripke. Ao perguntarmos se Aristóteles refere-se ao mesmo objeto em todos os mundos possíveis, utilizamos a linguagem do mundo atual para descrever todos os mundos possíveis estabelecidos a partir do mundo atual. Quando descrevemos os mundos possíveis através da linguagem do mundo atual, temos que o NP "Aristóteles" refere-se ao mesmo objeto, ainda que em outros mundos possíveis os falantes possam utilizar outro NP para Aristóteles. Mantemos nesses mundos a referência desse NP. Quando descrevemos esse mundo com nossa linguagem, o termo "Aristóteles" designa o mesmo objeto.

Outro problema é o do escopo dos operadores modais levantado por Dummett (1973). É óbvio que Aristóteles não poderia não ser Aristóteles, mas é óbvio que o mestre de Alexandre Magno pudesse não ser o mestre de Alexandre Magno? Temos duas sentenças:

- (2) Aristóteles poderia não ser Aristóteles.  
 (3) O mestre de Alexandre Magno poderia não ser o mestre de Alexandre Magno.

Claramente (2) é falsa, mas existe uma ambiguidade no operador modal em (3)<sup>3</sup>. Tendo-se “M” para representar a propriedade *ser mestre de* e “a” para representar Alexandre Magno, tomem-se as seguintes sentenças:

$$(3)^* \diamond \exists x (\forall y (Mya \leftrightarrow x=y) \wedge \neg Mxa)$$

$$(3)^{**} \exists x (\forall y (Mya \leftrightarrow x=y) \wedge \diamond \neg Mxa)$$

A sentença (3)\* é uma falsidade, pois é falso que existe um mundo possível no qual há um objeto que é e não é mestre de Alexandre Magno. A sentença (3)\*\* é verdadeira, pois existe um objeto que é mestre de Alexandre Magno no mundo atual, mas existe um mundo possível no qual esse objeto não é mestre de Alexandre Magno. Substituindo-se em (2) “Aristóteles” por “o mestre de Alexandre Magno”, ter-se-ia a sentença (3). Em uma das leituras de (3), temos que termos co-referenciais são substituíveis. De (2) para (3)\* não há problema (pois ambas são falsas), mas de (2) para (3)\*\* o valor de verdade é alterado. Contudo, nada privilegia uma leitura em relação à outra. A questão é: dado que (2) é uma sentença falsa e que “Aristóteles” expressa o significado de “o mestre de Alexandre Magno”, podemos inferir logicamente que (3) é uma sentença falsa? Segundo uma leitura, temos preservação no valor de verdade, segundo outra não temos. Logo, não podemos inferir que (3) é falsa, pois essa é uma sentença de valor de verdade ambíguo. Criticando Kripke, Dummett propõe convencionar uma leitura do escopo dos operadores modais que preserve o valor de verdade quando substituirmos em uma sentença um NP por uma DD co-referencial (Dummett, 1973, pp. 113-115). Segundo ele, “the question may be resolved for all contexts by adopting some uniform convention determining the scope” (Dummett, 1973, p. 115). Contudo, essa solução não funciona. Tomemos:

- (4) Aristóteles poderia não ser o mestre de Alexandre Magno.

(4) é uma sentença verdadeira, caso aceitemos que não é necessário a Aristóteles ter ensinado Alexandre. Substituindo-se em (4) “Aristóteles” por “o mestre de Alexandre Magno”, temos (3). Contudo, (3) tem um valor de verdade ambíguo, ao contrário de (4). Como Dummett sugeriu, poderíamos convencionar que a leitura de (3) fosse sempre (3)\*\*, pois quem afirma (3) está querendo dizer (3)\*\* (Dummett, 1973, p. 114). Nesse caso, o valor de verdade é preservado. Substituindo-se em (2) “Aristóteles” por “o mestre de Alexandre Magno”, temos também (3). A sentença (2) é falsa, e a (3) é verdadeira já que deve ser lida como (3)\*\*. Não temos preservação de verdade de (2) para (3) segundo a leitura (3)\*\*.

O problema permanece. Nossa convenção não resolve todas as substituições possíveis. Quando resolvemos um exemplo da substituição (de (4) para (3)), deixamos em aberto o outro caso de substituição (de (2) para (3)) e vice-versa. Uma solução arbitrária seria convencionar de maneira *ad hoc* diferentes interpretações (dependendo da substituição efetuada), para que a verdade seja sempre preservada.

Outro desafio para a teoria de Kripke é o caso de DD's rígidas. Um exemplo de DD rígida ocorre na Matemática. Como números possuem apenas propriedades necessárias, as DD's na Matemática são todas rígidas<sup>4</sup>: "Note that here both ' $\pi$ ' and 'the ratio of the circumference of a circle to its diameter' are rigid designators, so the arguments given in the metric case are inapplicable" (KRIPKE, 1980, p. 60). O contraste pretendido por Kripke entre os NP's e as DD's em contextos modais não é alcançado na matemática. A tese descritiva (pelo menos em Frege e Searle) aponta que o conteúdo descritivo associado pelos falantes compõe o significado dos NP's. Kripke afirma não ter argumentos contra esse Descritivismo no âmbito da matemática. Será que ele tem razão? Bem, em geral, seres humanos não possuem muito conhecimento matemático. Às vezes, utilizamos DD's que não são rígidas para fixar a referência de um número. Podemos associar ao numeral "5" o significado da DD não rígida "o número de dedos da minha mão direita": em outro mundo possível, o número de dedos da minha mão poderia ter sido outro. A partir da tese da designação rígida, não se conclui que o significado de um NP não é dado pelo significado de suas respectivas DD's. A argumentação kripkeana permite-nos concluir que não é o caso: se um falante associa um certo conteúdo descritivo, então esse conteúdo é sempre o significado do NP em questão, ou compõe o feixe descritivo que expressa o significado desse NP. O fato de um falante (ou uma comunidade) associar ao numeral "5" o significado da DD "o número de dedos da minha mão direita" não implica que o significado dessa DD seja o significado de "5" ou mesmo componha o significado desse numeral. Portanto, a tese da designação rígida alcança seu objetivo mesmo na matemática.

A existência de algumas DD's rígidas não constitui um problema a princípio. Mas, e se todas as DD's fossem rígidas, haveria algum problema? Apesar da tese da designação rígida perder sua força, isso não implicaria que o Descritivismo estaria correto. Como podemos tornar rígida uma DD? "O mestre de Alexandre Magno" não designa o mesmo objeto em todos os mundos possíveis. Para tornar essa DD rígida precisamos de um operador. Podemos introduzir o operador de atualidade em qualquer DD para torná-la rígida. "O mestre de Alexandre Magno neste mundo (no mundo real ou atual)" é uma DD que possui a mesma referência em todos os mundos possíveis. Temos um problema, pois qualquer DD a que um falante associe é capaz de receber o operador de atualidade. Logo, aquela DD associada por um falante a um certo NP possui o mesmo comportamento em contexto modal do que o NP em questão desde que essa DD seja rigidificada. Nenhum contraste em termos de comportamento modal é constatável entre NP's e DD's. Conclusão: a tese da designação rígida não é verdadeira. A resposta de

Kripke consiste em distinguir entre designadores rígidos *de jura* e designadores rígidos *de facto*:

...the distinction between '*de jure*' rigidity, where the reference of a designator is *stipulated* to be a single object, whether we are speaking of the actual world or of a counterfactual situation, and mere '*de facto*' rigidity, where a description 'the x such that Fx' happens to use a predicate 'F' that in each possible world is true of one and the same unique object (e.g., 'the smallest prime' rigidly designates the number two) (KRIPKE, 1980, p. 21 n. 21).

NP's são designadores rígidos *de jure*, mas DD's não são. A rigidez dos NP's advém da estipulação de sua referência no mundo atual, enquanto a rigidez das DD's advém do fato delas expressarem uma propriedade instanciada pelo mesmo objeto em todos os mundos possíveis. Kripke espera, assim, ainda conseguir distinguir o comportamento semântico dos NP's em relação às DD's. Mas não é claro como isso é conseguido através da distinção entre rigidez *de jure* e *de facto*. Kripke apenas observou que a rigidez dos NP's é distinta da rigidez das DD's. DD's referem-se expressando atributos unicamente instanciados, NP's referem-se por mera estipulação. Contudo, a distinção entre rigidez *de jure* e *de facto* independe, em um certo sentido, do fenômeno da rigidez. Poderíamos distinguir entre dois tipos de termos singulares: termos singulares *de jure* e termos singulares *de facto*. Os primeiros referem-se por mera convenção (indexicais e NP's), enquanto os últimos referem-se por meio de atributos unicamente instanciados (DD's). O contraste entre a maneira como os NP's e as DD's referem-se estaria explicitado. Temos, assim, um critério para afirmar que NP's não expressam o significado de DD's<sup>5</sup>. A preocupação de Kripke com a rigidez de algumas DD's indica que ele não deseja apenas refutar o Descritivismo dos seus antecessores, mas que pretende refutar a tese mais geral de que NP's não expressam o significado de DD's, sendo os primeiros termos meramente denotativos. Se ele pretende mesmo isso, claramente a tese da designação rígida não é suficiente. Por isso, Kripke distinguiu entre rigidez *de jure* e *de facto*. Contudo, se Kripke pretende refutar uma versão mais fraca de Descritivismo<sup>6</sup>, então a tese da designação rígida é suficiente. A rigidez das DD's não implica que essas DD's serão associadas pelos falantes no seu uso dos NP's. Nem sempre (talvez nunca) usaremos DD's rigidificadas para identificar o portador de um NP. Logo, sustenta-se que não é o caso: se um falante associa um certo conteúdo descritivo, então esse conteúdo é sempre o significado do NP em questão.

Outra réplica contra a objeção das DD's rigidificadas é questionar se DD's rigidificadas são realmente DD's. Tomemos o predicado "ser cadeira". A partir desse termo predicativo, podemos formar um termo singular anexando um indexical: "esta cadeira". Não temos mais um termo predicativo. O mesmo ocorre com relação ao operador de atualidade. A aparente DD "o mestre de Alexandre Magno neste mundo (no mundo real)" é um termo indexical, sua rigidez deriva do caráter indexical do operador de atualidade, não dos predicados descritivos presentes na aparente DD "o mestre de Alexandre Magno neste mundo (no

mundo real)”. Esse termo designa o mesmo objeto em todos os mundos possíveis (Aristóteles), mesmo em mundos possíveis nos quais Aristóteles não é mestre de Alexandre Magno. Toda DD tornada rígida é apenas um indexical. Afirmar que um NP expressa o significado de uma DD tornada rígida é simplesmente afirmar que um NP expressa um significado de um termo indexical. Sem possuir predicados descritivos relevantes na determinação do seu significado, o significado de um indexical seria o objeto referido no contexto de seu proferimento. Temos, nessa situação, que NP’s teriam como significado seus portadores. Essa conclusão é uma mera consequência da admissão de que NP’s expressam o significado de DD’s tornadas rígidas. Quem não defende isso não está comprometido com aquela conclusão. Além disso, é questionável se a rigidez das DD’s tornadas rígidas é uma rigidez *de facto*, pois a propriedade que uma DD desse tipo representa não é instanciada pelo mesmo objeto em todos os mundos possíveis. Um exemplo claro de rigidez *de facto* é o caso das DD’s na Matemática. Uma DD como “o sucessor de 8” designa o mesmo objeto em todos os mundos possíveis, pois a propriedade *ser sucessor de 8* é instanciada pelo mesmo objeto em todos os mundos possíveis. Não há uma estipulação nesse caso. No que diz respeito às DD’s tornadas rígidas, temos que sua rigidez advém de uma estipulação. Uma DD como “o melhor jogador de futebol de toda a História neste mundo (no mundo real)” designa o mesmo objeto em todos os mundos possíveis por causa de um indexical, pois a propriedade *ser melhor jogador de futebol de toda a História* é instanciada por diferentes objetos através dos mundos possíveis. Portanto, DD’s tornadas rígidas teriam o mesmo tipo de rigidez dos NP’s: rigidez *de jure*. Logo, essa distinção kripkeana não consegue contrastar o comportamento semântico dos NP’s em relação às DD’s (com a exceção das DD’s no campo da matemática).

Deixemos de lado por enquanto esse argumento e aceitemos que a distinção entre a rigidez *de jure* e a *de facto* alcançou seu objetivo. NP’s são designadores rígidos *de jure*. DD’s não são designadores rígidos ou são designadores rígidos *de facto*. Em qualquer caso, DD’s possuem um comportamento semântico distinto dos NP’s. Conclusão: NP’s não expressam o significado de DD’s. Tentei argumentar que essa conclusão é devido à distinção entre rigidez *de jure* e a *de facto* e que a distinção entre *de jure* e *de facto* é independente do fenômeno da rigidez. Portanto, essa conclusão não é devedora da verdade da tese da designação rígida. Coloco-me, por ora, a seguinte indagação: qual a função da tese da designação rígida dentro das investigações kripkeanas acerca do fenômeno da nomeação? Bem, retomemos uma importante consequência do argumento que sustenta a tese da designação rígida: não é o caso se um falante associa um conteúdo descritivo qualquer, então esse conteúdo é sempre o significado dos NP’s. A tese da designação rígida, contudo, não nos permite concluir que NP’s não expressem o significado de DD’s. Na matemática, isso já foi apontado, mas o mesmo ocorre com todos os NP’s. Uma fundamental conclusão da tese da designação rígida consiste em distinguir entre significar e fixar a referência, pois não é o caso que se um conteúdo descritivo é associado pelo falante ou mesmo

por toda a comunidade em seu uso de um NP que esse conteúdo descritivo seja o significado dos NP's. A princípio, portanto, não me é permitido afirmar que NP's não expressem o significado de DD's. Contudo, se a argumentação de Kripke é correta, temos: se buscarmos construir uma teoria do significado dos NP's, então essa teoria não deve ser buscada no conteúdo descritivo associado pelos falantes ou por toda a comunidade. A argumentação em torno da tese da designação rígida permite-nos afirmar que o problema do significado não é dependente do problema da fixação da referência. Apesar disso, se os NP's possuem um significado distinto da referência, então esse significado é fixador da referência por afirmar um certo conteúdo descritivo do portador de um NP qualquer. No prefácio de *Naming and Necessity*, Kripke aponta a distinção entre significar e fixar a referência como uma das consequências da tese da designação rígida: "... this [NP's são designadores rígidos] showed that epistemic questions should be separated from questions of necessity and contingency, and that to fix a reference is not to give a synonym" (1980, p. 15). Em outro momento do prefácio de NN, Kripke, curiosamente, parece entender o contrário que o fenômeno de rigidez seria uma consequência da distinção entre significar e fixar a referência:

...the names of ordinary language are rigid designators can in fact be upheld. Part of the effort to make this clear involved the distinction between using a description to give a meaning and using it to fix a reference. Thus at this stage I rejected the conventional description theory as an account of meaning, though its validity as an account of the fixing of a reference was left untouched (1980, p. 5).

A tese da designação rígida implica que o significado de um NP é independente de como o falante fixa a referência. Por isso, devemos distinguir entre significar e fixar a referência. Descritivistas, como Searle, afirmaram que o significado de um NP é dado pelo significado das DD's utilizadas pelos falantes em uma comunidade para fixar a referência. Caso concordemos com a tese da designação rígida, o Descritivismo de Searle, Frege e Russell é simplesmente falso. Contudo, ainda permanece em aberto a questão central: qual o significado dos NP's? Essa questão não é solucionável positivamente apenas através da tese da designação rígida. A partir dessa tese sabemos que o significado de um NP não é o que os descritivistas clássicos pensavam.

## Referências

- DUMMETT, Michael. **Frege: philosophy of language**. London: Duckworth, 1973.
- HAACK, Susan. **Filosofia das lógicas**. São Paulo: editora da UNESP, 2002.
- HINTIKKA, Jaakko. "modality and quantification" in: **Models for modalities**. Dordrecht, Boston: Reidel, 1969.

KRIPKE, Saul. **Naming and necessity** (1972). Cambridge, Massachusetts: Harvard University Press, 1980.

\_\_\_\_\_. "Identity and necessity" in: MUNITZ, M. (ed.). **Identity and individuation**. New York: New York University Press, 1971.

LEWIS, David. **On the plurality of worlds**. Oxford: Blackwell, 1986.

RUSSELL, B. **The Principles of Mathematics** (1903). New York e London: W. W. Norton & Company, 1996.

\_\_\_\_\_. "On denoting" (1905) in: **Logic and knowledge**. Nottingham: Spokesman, 2007.

## Notas

<sup>1</sup> Tomo essa suposição por motivo de simplificação. De fato, o significado que os membros de uma comunidade linguística relacionam a um NP é fluido e variável de falante. Mas isso não afeta a argumentação que eu estou expondo.

<sup>2</sup> Russell manteve uma teoria semelhante à de Frege em *The Principles of Mathematics* (1903). Em *On denoting* (1905), Russell atacou duramente a concepção de sentido de Frege. Segundo Russell, termos sub-sentenciais não expressariam sentido/significado.

<sup>3</sup> Existe também uma ambiguidade no operador de negação, mas como não faz diferença para a argumentação, preferi não levantar mais essa complicação.

<sup>4</sup> Isso se restringe a DD's que exibem propriedades formais. DD's como: "o número preferido de Maria" não são rígidas, já que não exibem nenhuma propriedade numérica.

<sup>5</sup> Essa é, basicamente, a teoria de John Stuart Mill.

<sup>6</sup> "versão mais fraca de Descritivismo": NP's expressam o significado de DD's associadas pelo falante ou o significado de um feixe disjuntivo de DD's que representam propriedades comumente atribuídas ao portador do NP.



IV

---

ÉTICA E POLÍTICA



# PRINCÍPIOS BIOÉTICOS E A ATUAL LEI DE BIOSSEGURANÇA BRASILEIRA<sup>1</sup>

DARLEI DALL'AGNOL

UFSC/CNPq

darlei@cfh.ufsc.br

Na presente comunicação, vou fazer algumas reflexões procurando repensar a atual Lei Brasileira de Biossegurança que proíbe a engenharia genética em células germinativas humanas, em zigotos e em embriões humanos. Participei da discussão pela sua aprovação em 2005 e da defesa de sua constitucionalidade em 2008, mas considero, hoje, que ela precisa ser aperfeiçoada. Nesse sentido, o biodireito precisa ser mais ágil e inspirar-se em princípios bioéticos sólidos para regular e aplicar normas jurídicas relacionadas com o desenvolvimento da biotecnologia. O meu *questionamento* da referida lei será feito a partir dos próprios princípios bioéticos que estabelecem parâmetros para as pesquisas envolvendo seres humanos. A intenção principal é apenas a de refletir sobre certas condições restritivas da lei e algumas de suas proibições que parecem não fazer sentido tendo em vista certos desenvolvimentos da biotecnologia aplicada à reprodução humana.

Não há dúvida de que o progresso na biotecnologia, considerada aqui no seu sentido amplo, dá origem a perplexidades morais e gera aparentes dilemas clamando por algum tipo de reflexão ética. Basta lembrar as discussões sobre os riscos, sobre os benefícios e malefícios, trazidos pelos transgênicos (desde os animais e plantas transgênicos e a possibilidade de criação de novas espécies sujeitas a patenteamento até a produção de formas artificiais de vida produzidas com “DNA sintético” que poderão funcionar como bioreatores despoluindo o meio ambiente). Assim como a biotecnologia aplicada aos animais não-humanos e às plantas gera reflexões éticas, também a engenharia genética em seres humanos gera temores e é frequentemente associada a práticas eugênicas que levam a algum aprimoramento e, então, as discussões tornam-se ainda mais acaloradas pela associação com práticas nazistas. Existem também questões muito delicadas sobre patenteamento de genes humanos ou mesmo sobre os perigos da produção de “novas raças” (sobrehumanos, pós-humanos etc.) que levantam preocupações similares. Há ainda os perigos do bioterrorismo e até mesmo temores sobre os efeitos potencialmente catastróficos da biopirataria. Portanto, questões de biossegurança em relação ao bem-estar humano (por exemplo, segurança alimentar), segurança animal e ambiental (por exemplo, o perigo do descontrole genético) trazem muitas preocupações sociais e éticas clamando por novas reflexões e novas normas jurídicas. Temer um fim apocalíptico produzido pela

tecnologia e defender a volta a um primitivismo naturalista, todavia, não faz sentido.

Não há dúvida, então, que a sociedade precisa regular a biotecnologia para minimizar os riscos e maximizar os benefícios protegendo pessoas, animais e o meio ambiente. É claro que a legislação da biossegurança deve incluir todas as tecnologias biológicas e não somente a engenharia genética. Uma legislação eficaz sobre essas questões deve estabelecer os requerimentos para pesquisa, manuseio e eventual comercialização dos produtos resultantes. Claramente, a finalidade ética da biossegurança é proteger a saúde humana e animal e o meio ambiente. A questão principal é: como fazer isso? Em outras palavras, a biossegurança deve ser assegurada por instituições específicas (estatais ou independentes) e por leis públicas, mas elas são realmente capazes de refrear o progresso biotecnológico quando ele é danoso e permiti-lo quando é benéfico? E o que a ética, especialmente a bioética, pode fazer para auxiliar nessa tarefa?

A principal agência regulatória no Brasil em termos de biossegurança, além é claro dos poderes legislativo e judiciário, é a CTNbio (Comissão Técnica Nacional de Biossegurança), cuja principal função é aconselhar o Ministro da Ciência e da Tecnologia. A CTNbio é composta por vários membros, majoritariamente por cientistas de diferentes áreas das ciências naturais e ambientais, mas também por representantes civis. A CTNbio reúne-se mensalmente para certificar laboratórios e autorizar a pesquisa científica envolvendo produtos geneticamente modificados. A ANVISA também desempenha um papel importante seja em questões alimentares seja no controle de remédios. Nos Estados Unidos, a segurança alimentar é regulada pelo FDA (Departamento de Alimentação e Remédios) e pelo DHHS (Departamento de Saúde e Proteção Ambiental) e pela APHIS (Serviço de Inspeção da Saúde Animal e das Plantas), ambas ligadas ao Ministério da Agricultura. Na Inglaterra, também há agências e, sobre o principal tema a ser discutido neste trabalho, destaca-se a respeitadíssima HFEA (*Human Fertilization and Embriology Authority*) que recentemente envolveu-se na discussão de uma reforma legislativa parecida com a que será defendida aqui. Não vou entrar aqui numa discussão sobre o papel e a eficácia dessas agências.

O principal referencial ético usado pela CTNbio é o enfoque bioético baseado em 4 princípios fundamentais. A Instrução Normativa nº 9, de 1997, estabelece que qualquer experimento que conduz à modificação genética de seres humanos deve ser considerada “pesquisa envolvendo humanos” e, portanto, ser regulada pela famosa e polêmica Res. 196/96 do Ministério da Saúde que seguirá, então, os princípios *prima facie* da autonomia, da não-maleficência, da beneficência e da justiça. Esses princípios afirmam que a moralidade da pesquisa científica envolvendo seres humanos é assegurada considerando-se as seguintes normas:

- a) consentimento livre e esclarecido dos indivíduos-alvo e a proteção a grupos vulneráveis e aos legalmente incapazes (autonomia);

- b) ponderação entre riscos e benefícios, tanto atuais como potenciais, individuais ou coletivos (beneficência), comprometendo-se com o máximo de benefícios e o mínimo de danos e riscos;
- c) garantia de que danos previsíveis serão evitados (não maleficência);
- d) relevância social da pesquisa com vantagens significativas para os sujeitos da pesquisa e minimização do ônus para os sujeitos vulneráveis, o que garante a igual consideração dos interesses envolvidos, não perdendo o sentido de sua destinação sócio-humanitária (justiça e equidade).

Atualmente, esse aparato bioético é um dos melhores disponíveis visto que ele é capaz de resolver muitos problemas concretos, mas os princípios devem ser contrabalançados de modo equilibrado para que sejam bem aplicados. Caso contrário, ele leva a decisões arbitrárias. Os princípios do respeito pela pessoa e da beneficência (e não-maleficência) serão considerados aqui normas éticas fulcrais no sentido wittgensteiniano, ou seja, são proposições gramaticais constitutivas do próprio sentido da moralidade. Desse modo, não será assumida nem uma teoria ética puramente consequencialista ou utilitarista, como a de Savulescu, nem uma teoria puramente deontológica (do tipo habermasiano ou rawlsiano como a que inspirou a crítica de Baldwin ao princípio da beneficência procriativa). Ao contrário, a ética principialista é um enfoque normativo misto.

A adoção, de forma séria e não simplesmente como “jeitinho,” ou “para inglês ver,” desse aparato bioético deve nos levar a repensar a Lei de Biossegurança Brasileira (Lei 11.105/2005). Essa norma é uma estranha combinação de questões envolvendo OGMs e pesquisa com células-tronco embrionárias. Teria sido melhor separar essas questões. Por isso, vou concentrar-me, aqui, somente no segundo ponto. O artigo 5º da referida lei torna permissível a pesquisa envolvendo células-tronco embrionárias, mas estabelece duas condições: primeira, os embriões devem ser inviáveis para reprodução; segunda, eles devem estar congelados há 3 anos ou mais. Em qualquer circunstância, o consentimento dos pais é necessário e essa é uma clara aplicação do princípio da autonomia citado acima. Institutos de pesquisa, clínicas, universidades etc. devem também aprovar seus projetos envolvendo células-tronco em seus respectivos comitês de ética para julgar se os outros princípios estão sendo realmente resguardados.

A Lei de Biossegurança, como é talvez amplamente sabido, foi questionada pelo Procurador-geral da República, Cláudio Fonteles. Ele argumentou que o Art. 5º seria inconstitucional. A razão principal por ele apresentada foi a de que a Constituição Federal está, já no seu *primeiro* artigo, baseada no princípio da dignidade humana. Mas o que significa “dignidade humana”? Trata-se de um valor transcendente ou ele pode ser descrito de forma completamente naturalizada? Em ambos os casos, há pressupostos metafísicos que não podem ser *comumente partilhados numa moralidade pública*. Além disso, ele argumentou que a constituição garante o direito à vida (no seu Art. 5º) e – a partir daqui o procurador argumentou por conta própria revelando algumas de suas

crenças subjetivas básicas – visto que a vida começa no momento da concepção (supostamente, quando o espermatozóide fecunda o óvulo), a Lei 11.105/2005 seria, por conseguinte, inconstitucional. Em maio de 2008, o Supremo Tribunal Federal decidiu que a pesquisa com células-tronco embrionárias é legal, que a Lei de Biossegurança não é inconstitucional, baseado na tese de que o *embrião não é uma pessoa*. Afinal, a Lei de Biossegurança tinha sido aprovada, em 2004, por 90% dos parlamentares que claramente representam os interesses gerais da sociedade brasileira, incluindo, por exemplo, dos deficientes físicos, e não apenas da Monsanto. A principal questão ética, todavia, não é se a Lei de Biossegurança é constitucional ou não (nesse sentido, a argumentação *amicus curiae* apresentada pelo MOVITAE – Movimento em Prol da Vida – contra o Procurador-geral foi totalmente adequada), mas se o próprio artigo primeiro da constituição federal não está associado a uma crença moral particular. Ele deve expressar um requerimento de uma Moralidade Comumente Partilhável, ou seja, de uma moral pública baseada nas normas fulcrais do respeito mútuo entre as pessoas e no cuidado para com os indivíduos vulneráveis.

É importante, agora, esclarecer o estado atual dessas pesquisas com células-tronco, principalmente, salientar que elas são bastante *experimentais* ainda. Daí a importância da discussão séria sobre os princípios fundamentais da bioética. Já há transplantes de células-tronco sendo feitas em hospitais brasileiros para pessoas com problemas físicos. Por exemplo, no hospital São Rafael, de Salvador, alguns pacientes vêm aparentemente recuperando movimentos devido a tais tratamentos (cf. em [www.hsr.com.br](http://www.hsr.com.br)).

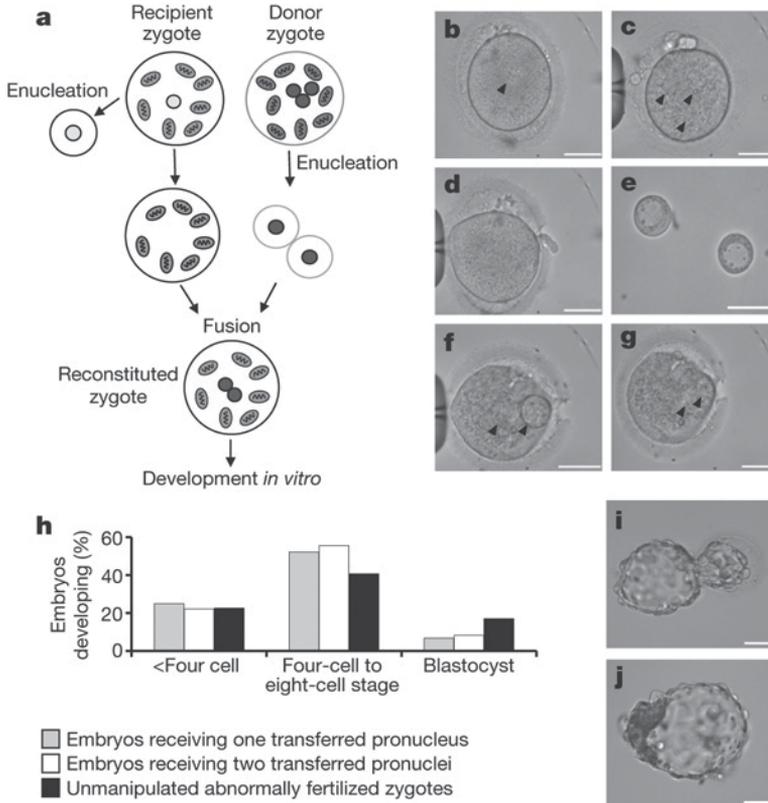
No Brasil, uma das pesquisas mais confiáveis é a do Centro do Genoma Humano, da USP, dirigido por Mayana Zatz. Denunciando a propaganda enganosa e comércio fácil dos “negócios da China”, ou seja, as terapias com células-tronco altamente duvidosas praticadas, por exemplo, pelo Dr. Huang, a geneticista Zatz mostra, em seu livro *Genética*, que não há ainda como garantir que, uma vez injetadas, as células-tronco não se comportem como “crianças malcriadas”: ao injetá-las, por exemplo na medula de uma pessoa, é possível perder o controle e assim elas “decidem” que, em vez de neurônios, vão formar ossos. Há outros problemas: “Ainda não temos certeza, por exemplo, de que as células-tronco vão chegar ou permanecer no órgão-alvo. E não sabemos qual é o risco de formação de tumores, ou seja, a sua replicação indefinidamente.” (ZATZ, 2001, p.136) É por essa e por outras razões que os *tratamentos experimentais* com células-tronco embrionárias com seres humanos só foram autorizadas *há poucos dias* em países como os EUA. A Empresa de biotecnologia Geron obteve a autorização governamental para tratar pacientes com lesões na medula espinhal. Como ressalta Zatz, há muitas questões que precisam ser ainda respondidas e, segundo o seu relato, ela está injetando células-tronco humanas obtidas de diferentes fontes (cordão umbilical, polpa dentária etc.) em *modelos animais* que apresentam doenças neuromusculares. Por conseguinte, ela adverte que está longe ainda de iniciar terapias gênicas em pacientes humanos para fins de tratamento no nosso país. Precaução e cautela, são, portanto, fundamentais.

De qualquer forma, as duas condições restritivas do Art. 5 deveriam ser eliminadas. Mayana Zatz trabalha apenas com células-tronco mesenquimais, mas as embrionárias, totipotentes, são potencialmente muito mais promissoras em termos de medicina regenerativa, de terapia celular. Por isso, deveria ser permissível o uso de embriões viáveis para fins de pesquisa e não somente aqueles congelados há muito tempo. Em outros termos, se é adotado seriamente o princípio do respeito pela autonomia como norma fulcral, deve-se permitir aos pais que gostariam de usar os embriões excedentes para fins de pesquisa científica que fiquem livres de fato para escolher entre os benefícios e riscos, sempre sujeitos à avaliação pessoal, da terapia gênica. Considerando que tal pesquisa foi possibilitada no Brasil somente recentemente e que o tratamento com células-tronco tornou-se possível há menos tempo ainda (mesmo que somente em modelos animais) e que os resultados são promissores, as duas condições restritivas acima mencionadas são injustificáveis tanto sob o ponto de vista do princípio da beneficência quanto do respeito pela autonomia, o qual deve ter peso maior numa sociedade verdadeiramente pluralista e democrática. A Lei de Biossegurança é, então, anti-ética no sentido de não possibilitar um uso maior dos embriões congelados para, via pesquisa e tratamento com células-tronco embrionárias, melhorar significativamente a qualidade de vida de muitos seres humanos que estão presentemente sofrendo.

Essa não é, todavia, a principal preocupação da presente comunicação. Minha reflexão, hoje, está focada noutro ponto. A Lei de Biossegurança proíbe, no seu Art. 6º, a produção de embriões humanos para comercialização; proíbe a *clonagem* humana e proíbe o uso de *engenharia genética* em células germinativas humanas, zigotos e embriões humanos. Essa proibição também não é contrária ao princípio da beneficência ou até mesmo um desrespeito à autonomia? É necessário, por conseguinte, ter clareza sobre o fato de que a Lei de Biossegurança não apenas não é inconstitucional nos termos nos quais está decretada, mas também que ela é muito restritiva considerando o progresso que se está alcançando na biotecnologia recentemente. O argumento principal que será aqui apresentado está baseado na necessidade de contrabalançar melhor os princípios bioéticos adotados na própria lei, mas especificamente no mencionado princípio da autonomia, o qual deveria ocupar um lugar de maior destaque em relação aos outros princípios.

Se essa afirmação faz sentido, então a proibição da clonagem e da engenharia genética feita no Art. 6º precisa ser revista. A intenção, aqui, não é sustentar um ponto de vista definitivo e dogmático sobre essas questões, mas somente fazer algumas reflexões a partir de certos desenvolvimentos científicos e tecnológicos. Para ilustrar, é necessário chamar a atenção para uma discussão importantíssima sobre métodos para evitar doenças mitocondriais que está sendo feita na Inglaterra e nos Estados Unidos. Imaginem a seguinte situação: um casal quer ter filhos, mas precisa fazer fertilização *in vitro*. Num diagnóstico pré-implantação (DPI), descobre-se que os pré-embriões possuem uma anomalia no DNA mitocondrial que levará ao desenvolvimento de alguma doença degenerativa grave. Pesquisadores ingleses da Universidade de Newcastle, na Inglaterra,

descobriram técnicas para evitar tais doenças genéticas. O procedimento pode ser ilustrado a partir da revista *Nature*, de 2010, que publicou trabalhos sobre essa questão (vide referências):



A figura ilustra a transferência de núcleos entre zigotos humanos e o desenvolvimento inicial do embrião resultante.

Para compreender melhor o que está sendo proposto, é necessário fazer alguns esclarecimentos. Uma célula é composta por dois tipos de DNA, a saber, o DNA Nuclear e o DNA Mitocondrial. A organela chamada “mitocôndrias” é responsável pela produção de energia das células. Ora, enquanto o DNA Nuclear é o maior responsável pelas nossas características físicas e está até mesmo associado a alguns aspectos da personalidade, o DNA Mitocondrial é distinto e uma anomalia pode causar uma série de doenças. Costuma mencionar-se, nesse sentido, não apenas Huntington, Parkinson, possivelmente Alzheimer, mas até mesmo a epilepsia, a cegueira, a surdez, a demência, a distrofia muscular etc. Estima-se que das 1700 doenças conhecidas que possuem origem genética, mais de 60 podem ser causadas por anomalias do DNA Mitocondrial. Visto que no “momento da fecundação,” ou melhor, quando o espermatozóide penetra o óvulo, somente o DNA Nuclear do pai é fundido com o DNA Nuclear da mãe, só a genitora transmite as doenças associadas a anomalias no DNA Mitocondrial. As técnicas propostas pelos pesquisadores ingleses impedem a transferência da

anomalia. De qualquer modo, no final do procedimento tem-se um embrião com um pai e duas mães biológicas. Vou chamar isso de “fertilização tripla”.

O que os pesquisadores da Universidade de Newcastle estão propondo, então, é a transferência do núcleo do zigoto produzido pelo casal para a base citoplasmática de outro zigoto produzido a partir do óvulo de uma doadora que tenha um DNA Mitocondrial sadio. É claro que várias questões surgem aqui, algumas técnicas e outras éticas. Foi por isso que, recentemente, o secretário de saúde britânico Andrew Lansley, solicitou à HFFA uma audiência pública para discutir evidências da eficácia e segurança de técnicas que evitem doenças mitocondriais através da concepção assistida (*in vitro*). O que acontece é que, na Inglaterra, o ACT 1990 permite somente que células germinativas e embriões que não tiverem seu DNA Nuclear ou Mitocondrial alterado possam ser usado. Esse é o caso também no Brasil, mas diferentemente do que acontece no nosso país, a Grã-Bretanha permite a criação de embriões somente para fins de pesquisa. Quando através de testes pré-implantação é diagnosticada uma doença com base mitocondrial, o embrião não pode ser usado para reprodução. Assim, a transferência nuclear tem o potencial de evitar a transmissão de sérias doenças mitocondriais. Somente para clarificar: a transferência nuclear (*pronuclear transfer*) envolve a transferência de núcleos de um zigoto (um pré-embrião) com mitocôndrias não sadios para um zigoto com mitocôndrias sadio. O uso dessa técnica não é atualmente permitido na Inglaterra, pois ela altera o DNA do embrião. Daí a necessidade de mudar a legislação para permitir a técnica e evitar que sejam gerados descendentes com doenças gravíssimas.

Há, também, problemas de ordem ética. A técnica acima descrita e ilustrada resulta na destruição de alguns pré-embriões ou zigotos que, segundo algumas convicções morais e religiosas, é uma pessoa em potencial. Todavia, no caso brasileiro, a discussão sobre a utilização de células-tronco embrionárias mostrou que a jurisprudência brasileira não considera o embrião uma pessoa. Desse modo, a técnica não deveria ser proibida por essa razão. Mas, também no nosso país, não é permitido aplicar engenharia genética em zigotos ou embriões humanos. Por conseguinte, também nós precisamos mudar a nossa legislação, principalmente, remover os empecilhos da Lei de Biossegurança.

Alguém poderia objetar que a decisão do Supremo Tribunal de não considerar o embrião uma pessoa foi equivocada e argumentar que num futuro próximo poderia existir outra composição de juízes e outra decisão ser tomada no sentido contrário tornando, portanto, proibidas tanto as pesquisas com células-tronco embrionárias quanto técnicas de fertilização tripla. Esse cenário, todavia, poderia não ser necessariamente contrário a um procedimento de fertilização tripla que não envolvesse destruição de zigotos ou embriões. Esse é o caso, se for usada a técnica “*spindle transfer*” que implica em retirar o núcleo de um óvulo materno e transferi-lo para um óvulo enucleado de uma doadora. Nesse caso, não existiria destruição de pessoa alguma.

Se, e somente se, essas técnicas são realmente efetivas e seguras, então nós brasileiros devemos repensar a Lei de Biossegurança e remover as proibições do

Art. 6º, especialmente, a proibição de engenharia genética, para possibilitar que métodos que evitem a transmissão de doenças genéticas (que ironicamente deixariam de ser geneticamente transmissíveis) possam ser usados. Pelos menos, é necessária também em nosso país uma regulação diferente sobre esse assunto para permitir aqueles interessados em ter descendentes livres de certas doenças causadas por anomalias no DNA mitocondrial que possam fazer as suas escolhas.

Outra objeção que alguém poderia fazer é que tais procedimentos levam a um tipo de melhoramento genético inapropriado, ou seja, que se constituem em práticas eugênicas. Todavia, a eugenia *negativa*, para fins exclusivamente curativos (numa lista de doenças a ser especificadas em lei), é moralmente aceitável e até mesmo uma exigência dos princípios *prima facie* da não-maleficência e da beneficência. Portanto, a eugenia negativa é eticamente aceitável numa Moral Comumente Partilhável, numa moralidade pública, pluralista, democrática. Ela não pode ser associada com práticas nazistas sendo, inclusive, o seu verdadeiro oposto.

Há, obviamente, outros princípios que devem ser considerados como, por exemplo, o da justiça distributiva. Não há dúvida de que a engenharia genética pode produzir, a partir de uma filosofia política libertarianista, a base para novas desigualdades sociais e, portanto, produzir um mundo ainda mais injusto. Há, todavia, formas de evitar esse problema seja usando somente a eugenia negativa, curativa, com doenças *claramente listadas e reguladas pelos órgãos competentes*, seja adotando princípios de justiça distributivos baseados na necessidade e disponibilizando as técnicas para todos igualmente nos sistemas de saúde públicos. Não vou aprofundar esse ponto aqui e é claro que existem também questões comparativas de justiça. Além disso, se se mantém presente que a biotecnologia é aplicada não somente para pesquisa envolvendo seres humanos, mas, mais amplamente, também para o bem-estar animal e ambiental, então é necessário outro princípio *prima facie*, a saber, o da reverência ao valor intrínseco da vida como um todo. Infelizmente, não é possível discutir esses princípios nessa breve comunicação, mas parece que eles não colocam dificuldades substanciais para o argumento que foi apresentado acima.

Para finalizar, então, sem evidentemente pretender concluir a reflexão, é necessário fazer duas questões para continuar refletindo: por que não usar técnicas que evitem doenças transmissíveis via DNA Mitocondrial para pais que gostariam de dar um ponto de partida melhor para seus filhos? Se não há razões fundamentadas para não fazê-lo, então, por que não mudar a Lei de Biossegurança brasileira permitindo engenharia genética, com técnicas derivadas de clonagem, em células germinativas, em zigotos e embriões humanos mesmo que isso signifique eugenia negativa?

## Referências

- BALDWIN, T. Choosing Who: What is Wrong with Making Better Children? In: SPENCER, jr. *Freedom and Responsibility in Reproductive Choice*. Oxford: Hart Publishing, 2006. p.15-30.
- BEAUCHAMP, T. & CHILDRESS, J. *Principles of biomedical ethics*. Oxford: Oxford University Press, 2009.
- BORÉM, A. & SANTOS, F.R. & BOWEN, D. E. *Understanding Biotechnology*. New Jersey: Prentice Hall PTR, 2003.
- BUCHANAN, A. *et al.* *From chance to choice*. Genetics & Justice. Cambridge/New York. Cambridge University Press, 2000.
- CRAVEN, L. *et al.* Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease. *Nature*. V. 465, 2010. p.82-85.
- DALL, AGNOL. D. *Bioética: princípios morais e aplicações*. Rio de Janeiro: DP&A, 2004.
- DALL, AGNOL. D. *Princípios Bioéticos e a Lei de Biossegurança*. Cf. <http://www.cfh.ufsc.br/~darlei/biossegu.html> (14/04/2005).
- FRIAS, L. *A ética do uso e da seleção de embriões*. Florianópolis: Edufsc, 2012 (Série Ethica, no prelo)
- SAVULESCU, J. Procreative Beneficence: Why we Should Select the Best Children. *Bioethics*. V.15, N.5-6, 2001. p. 413-426.
- ZATZ, M. *Genética: escolhas que nossos avós não faziam*. São Paulo: Globo, 2011.

## Notas

<sup>1</sup> Comunicação apresentada no encontro do GT de Ética da ANPOF durante o VII Simpósio Internacional Principia em 17/08/2011. Uma versão preliminar, em inglês, foi lida no II Colóquio Internacional NEPC - Biotecnologia e Regulações, realizado em Belo Horizonte, em 28/04/2011, num mesa redonda com Mark Sherman, da Universidade de Oxford, sobre "*Justifying regulation and public involvement in biotechnological research*." Isso explica o modo como alguns temas são abordados aqui. Agradeço aos participantes dos encontros pelas contribuições, especialmente Hugh Lacey em BH e Lincoln Frias em Floripa, e ao CNPq pelo apoio à minha pesquisa atual sobre as bases metaéticas da bioética.

# O “GRUPO DE SETEMBRO” E O “11 DE SETEMBRO”<sup>1</sup>

FERNANDO CÉSAR COSTA XAVIER

*Universidade Federal de Roraima – UFRR*

fxavier010@hotmail.com

**Resumo:** O objetivo do presente artigo é analisar o “11 de setembro” como um problema de justiça internacional, captando-o sob a perspectiva teórica do marxismo analítico (corrente ligada ao chamado “Grupo de Setembro”). Uma vez que a corrente marxista analítica recupera o materialismo histórico-dialético como ferramenta epistêmica, e ainda propõe a justiça como tema caro à ciência política “à esquerda”, acredita-se que essa corrente pode lançar novas e instigantes luzes sobre as razões que deram causa ao 11 de setembro. No geral, até hoje, as chaves de análise do evento enfatizaram o aspecto cultural-religioso como móbil, e não apenas para a compreensão do ataque ao World Trade Center, senão da maioria dos conflitos internacionais contemporâneos. Nada obstante, o presente texto propõe a análise do fato – do “fato da história” *11 de setembro* – a partir de uma teoria social e histórica autorreferenciada como um *paradigma*, tal como John Roemer apresenta o marxismo analítico. Essa corrente teórica propõe um novo horizonte heurístico à medida que retoma antigas categorias da ciência política marxista a partir de uma relativamente nova epistemologia – a filosofia analítica anglosaxônica. A tese em questão é que, caso alguns autores do marxismo analítico se manifestassem especificamente acerca das razões que motivaram os ataques do 11 de setembro, diriam tratar-se, em grande medida, de reações contrárias às novas formas estruturais de opressão e alienação impostas à periferia do capitalismo global.

**Palavras-chave:** Grupo de Setembro; 11 de setembro, 11/9; justiça internacional; culturalismo; terrorismo global.

## Introdução

Às vésperas de completarem uma década, os atentados de 11 de setembro de 2001, contra as torres do World Trade Center, em Nova Iorque, marcaram tão fortemente a política internacional, que foram considerados, desde o primeiro momento, como um autêntico *fato da história*<sup>2</sup>.

Embora se acredite que, à luz da racionalidade pública, a imoralidade dos motivos que conduziram ao atentado é incontestável, vale lembrar que, em muitas partes do globo, muitas pessoas celebraram a investida da Al Qaeda contra as torres gêmeas, considerando-a um “ato de justiça”. E não é de todo incorreto afirmar que parte da esquerda mais ortodoxa, se não se esforçou para legitimar o atentado, ao menos não fez coro às críticas contra a ação empreendida pela organização de Osama Bin Laden<sup>3</sup>.

Muitas controvérsias se instalaram, e toda a repercussão que esse evento traumático causou nos anos seguintes permite dizer que ele ainda desafia os analistas que tentam desvendar-lhe as razões, ainda que alguém possa acreditar que estas já teriam sido fartamente expostas nos últimos anos pelos noticiários, ou com mais rigor pelos estudos acadêmicos publicados na última década.

Ante isso, o presente artigo pretende analisar o “11 de setembro” (doravante referido como ‘11/9’<sup>4</sup>) como um fato da história, consoante a filosofia da história marxista, sob a suposição de que, se há ainda estudiosos que insistem nessa chave de leitura, é porque ela ainda pode servir para aprimorar análises críticas. A remissão ao Grupo de Setembro será então tanto mais apropriada, à medida que se recorde que muitos pensadores reunidos nesse grupo se ocuparam, e se ocupam, da questão da justiça a partir de premissas marxianas.

Antes de se enveredar pela análise do 11/9 a partir das lentes do marxismo, porém, importa recapitular a leitura predominantemente proposta para o evento, para em seguida afirmar a necessidade de substituí-la pelo marxismo analítico.

## **1. Muitas globalizações: a análise “culturalista” do 11/9**

No geral, até hoje, as chaves de análise do evento enfatizaram bastante o aspecto cultural-religioso como móbil, e não apenas para a compreensão do 11/9, senão da maioria dos conflitos internacionais ocorridos em fins do séc. XX – sobretudo após 1979, com a Revolução Islâmica no Irã.

Com o fim da Guerra Fria, a eclosão de conflitos entre cosmovisões foi favorecida, e a chamada corrente culturalista, cujo expoente aqui indicado é Samuel P. Huntington, encontrou terreno fértil para defender que existiriam no mundo concorrentes culturas que disputariam entre si pela imposição de padrões civilizatórios. A publicação de *O choque das civilizações*, de Huntington, em 1993<sup>5</sup>, contribuiu sobremaneira para disseminar, ao longo da década de 90, essa abordagem culturalista, a tal ponto que, nos anos que precederam o 11/9, o receio de muitos – inclusive de acadêmicos – diante de “culturas”<sup>6</sup> como a islâmica e a asiática, encontrava-se já bem sedimentado no “Ocidente”<sup>7</sup>.

Em seu comentado livro, vale lembrar, Huntington propõe um debate que orbita “em torno da hipótese de que a política mundial está passando por uma modificação [...] na natureza dos conflitos, que [...] serão cada vez mais baseados em motivações culturais e religiosas, e não em diferenças ideológicas ou econômicas” (BORRADORI:2004:76).

Tratava Huntington de reconfigurar a tensão Ocidente-Oriente, dando ênfase para uma alegada divergência cultural-religiosa entre os hemisférios. O esforço do autor para depreciar o Leste se tornava patente à medida que, para ele, a democracia e os direitos humanos seriam incompatíveis com os valores orientais: “Os esforços ocidentais para propagar tais ideias [individualismo, liberalismo, constitucionalismo, direitos humanos, igualdade, liberdade, império

da lei, democracia, mercados livres, separação da Igreja e Estado] acabam produzindo uma reação contra o ‘imperialismo dos direitos humanos’ e uma reafirmação dos valores nativos, como se pode ver no apoio ao fundamentalismo religioso pelos mais jovens nas culturas não-ocidentais” (HUNTINGTON *apud* PRZEWORSKI *et alli*:2003:28-29).

Essa análise enviesada não deixou de parecer um vaticínio para muitos, à medida que organizações islâmicas fundamentalistas passaram a assumir o antiamericanismo antes propagandeado por organizações de extrema-esquerda, e a se especializar em ataques terroristas de largo alcance. O alvo preferencial, evidentemente, foram os Estados Unidos e seus aliados militares, países estes que, não apenas para Huntington, mas também outros autores culturalistas (desde Almond e Verba, passando por Inglehart, Granato e Leblang), seriam detentores de uma exemplar cultura democrática.

Apesar da proeminência que a hipótese culturalista alcançou, não tardou para que muitos analistas passassem a apontar equívocos contidos nela. Przeworski *et alli* (2009), mitigando a relevância das divergências culturais, fizeram questão de afirmar que “a maioria dos estudiosos do fundamentalismo religioso islâmico atribui o seu crescimento à deteriorização das condições econômicas das massas urbanas” (op. cit:29).

Um autor influente como Habermas, longe de ser analítico mas compactuando com a premissa da primazia da estrutura, também divergiu amplamente da tese culturalista de Huntington, e sustentou que o mote da difusão do terrorismo islâmico estaria ligado a questões econômicas, e não culturais. Segundo Giovanna Borradori (2004:77), comentando Habermas, um impacto significativo em nível estrutural seria a forma como o consumismo capitalista “explode” na periferia da sociedade internacional, inclusive no mundo árabe.

Embora muitos tenham buscado – com êxito – refutar a hipótese culturalista de que nos novos tempos o “Oriente fundamentalista” se levantava contra o “Ocidente democrático”, é curioso notar que muitas premissas do culturalismo ainda guiam o exame do 11/9. Talvez porque mesmo o *mainstream* islâmico aproveitou-se da ilustração culturalista-religiosa para dar conta de uma luta entre “povos fiéis” e “nações hereges”.

Mas até que ponto a nova dimensão do conflito Ocidente-Oriente não [mais] encerra um confronto entre “classes” (no sentido marxista)? Eis uma das questões que estimula a retomada da teoria marxista em uma de suas versões contemporâneas.

## 2. O Grupo de Setembro e o marxismo recuperado

O *Grupo de Setembro* representa um círculo restrito de estudiosos destacados pela retomada da teoria marxista a partir da abordagem tipicamente anglo-saxônica da filosofia analítica<sup>8</sup>.

A *teoria da história de Karl Marx* (1978), de Gerald A. Cohen, é considerada o livro que inaugura a linha de pesquisa do grupo: o marxismo analítico<sup>9</sup>. A partir de 1979, alguns pensadores de diversas áreas, instigados pela publicação do texto de Cohen, passaram a se reunir anualmente, nos meses de setembro, para compartilhar pesquisas que envolviam o marxismo e o apresentavam como tema de interesse comum. O resultado que adveio foi um conjunto de trabalhos que, se agrupados, representam uma teoria social e histórica autorreferenciada como um *paradigma* (ROEMER:1986), a partir do qual os problemas advindos do capitalismo e da exploração do homem pelo homem deveriam ser (re)analisados.

Ao reavivar antigas categorias da ciência política marxista (exploração, alienação, emancipação, produção), o Grupo de Setembro buscou apropriar-se de uma metodologia muito influente durante a Guerra Fria, para com ela abordar de modo mais preciso – e “sofisticado” (idem) – aquelas categorias típicas do marxismo.

A filosofia analítica e a teoria da escolha racional propunham a exposição clara e lógica de conceitos, a submissão constante das hipóteses a testes e ao confronto com o empírico, o postulado do agente racional individualmente considerado etc., e todos esses postulados convenceram muitos marxistas de que, seguindo essa linha, poderiam suplantar a ideologização e dar ao legado de Marx o “máximo rigor científico” (GARGARELLA:2008:105). Jon Elster, um dos fundadores do Grupo, rejeitando a dogmática marxista, assentou que um marxista analítico deveria estar disposto a abandonar proposições marxistas em prol de confutações lógicas e empiricamente confirmadas (idem).

Desse modo, conforme Arnsperger & Parijs (2003), o “marxismo analítico se funda sobre um duplo projeto: de um lado, buscar na tradição marxista um certo número de ideias centrais [...] de outro lado, aplicar a elas o equipamento intelectual herdado [...] da filosofia analítica e da economia matemática (p. 55).

Embora o Grupo de Setembro seja associado ao marxismo analítico, alguns dos membros, como Philippe Van Parijs e Hillel Steiner, não se consideram autênticos “marxistas”, embora tenham tendências claramente à esquerda. Em todo caso, o grupo, conforme afirma G. A. Cohen (2000:105), foi tomado como exemplo por muitos marxistas, semimarxistas e ex-marxistas que procuraram enveredar e engajar-se em questões de filosofia moral e política que antes eram vistas com desdém pelo marxismo ortodoxo, mas que passaram a ser revalorizadas, sobretudo após a publicação de *Uma teoria da justiça*, de John Rawls, em 1971.

Sob o novo prisma adotado, a “luta de classes”, ou a sobreposição dos exploradores em relação aos explorados, pôde ser vista como uma imoralidade capaz de gerar sentimentos de reprovação, e a divisão do trabalho e dos encargos sociais, à luz do mundo dos fatos, revelou-se uma escandalosa injustiça.

Se se imaginar que a divisão da sociedade em “classes” (no sentido que lhe atribui o marxista analítico Erik Olin Wright<sup>10</sup>), no atual estágio do capitalismo, mantém-se como um dado universalmente constatável, inclusive ampliado com

a globalização da economia, então mesmo os eventos ligados ao 11/9 poderiam revelar uma reprodução da luta entre classes.

### **3. Osama Bin Laden e a *Al Qaeda*: explorados ou exploradores?**

Uma primeira pergunta que surge é se grupos terroristas como a Al Qaeda assumiram não apenas parte do discurso e da estratégia ofensiva dos grupos esquerdistas de atuação internacional, mas também alguma “consciência de classe”. Por certo, o “mundo islâmico” abrange tão distintos matizes políticos e sociais, favorecidos ou prejudicados pelo capitalismo global em diferentes medidas, que não espanta que as relações diplomáticas entre os Estados Unidos e países muçulmanos sejam tão variadas. Em diversas sociedades islâmicas, do Afeganistão aos Emirados Árabes, verifica-se, nas devidas proporções, o abismo social entre pessoas e grupos inseridos/excluídos nas benesses proporcionadas pelo mercado.

Disso são perceptíveis duas circunstâncias. A primeira é que pessoas socialmente excluídas formam em sua grande maioria o contingente dos cooptados pelas organizações terroristas de apelo religioso, que passam a professar a ideologia do grupo sem que desconsiderem os interesses (racionais) que individualmente possuem por inclusão, aceitação e respeito.

A segunda circunstância é que essas organizações parecem operar de modo consciente dentro do sistema capitalista, e certamente têm claro que precisam fazer investimentos para poderem financiar suas ofensivas, as quais, a propósito, costumam se lançar sobre polos que lhes são concorrentes no mercado global. Vale lembrar que a participação da família Bin Laden no *Carlyle Group*<sup>11</sup>, no qual também participava a família Bush, dá margem para que se diga que a “Guerra ao Terror” lançada por George W. Bush simbolizou um conflito entre *burguesias*.

Tratando especificamente dos interesses do lado oriental, pode-se afirmar que a autorrealização dos atores centrais da ofensiva islâmica dependeu, e depende, de um ataque direto contra a hegemonia político-econômico-militar dos Estados Unidos, enquanto polo central capitalista. Isso, apesar dos custos altíssimos, pode se revelar útil ou imprescindível para a reordenação das distribuições das forças produtivas no contexto internacional.

É mesmo plausível imaginar que os megainvestidores do mundo islâmico fomentam o ódio contra o Ocidente por razões mais econômicas do que políticas ou religiosas, mas ocultando capciosamente isso.

Uma estratégia tão ou mais importante do que contratar notáveis executivos, para eles, talvez seja arregimentar recursos humanos capazes de agir leal e coordenadamente objetivando desestabilizar o poder de potências concorrentes. Pode-se dizer que, para as oligarquias islâmicas, a esta altura já, parafraseando Marx, Hayek suplantou Maomé.

E se é verdade que há um segmento socialista mais “xiita” que poupa críticas ao fundamentalismo islâmico, por compactuar com ele acerca de um antiamericanismo contumaz, é necessário ressaltar, por razões de coerência, que o capitalismo islâmico não deveria ser considerado menos injusto e alienador que aquele praticado no Ocidente. Dado o seu apego à lógica analítica, o marxismo analítico aparece como uma solução possível para o problema do sectarismo.

#### **4. Os microfundamentos do 11/9 e a filosofia da história marxista**

Outro ponto da abordagem analítica que merece destaque é que, conforme Roberto Gargarella (2008), muitos autores do Grupo de Setembro, influenciados pelo individualismo metodológico, desenvolveram a preocupação de apresentar os microfundamentos observáveis em empreendimentos coletivos (de grupos), de modo a revelar as “motivações e crenças dos agentes que participam deles” (p. 108).

No caso da ação terrorista do 11/9, a consideração é tanto mais pertinente quando se lembra que, em *Uma introdução a Karl Marx*, Elster disse que é preciso “abrir a caixa preta” (*apud* GARGARELLA:idem:108).

Discordando um pouco de Habermas (*apud* BORRADORI:45), há de se supor que na caixa preta dos aviões que se lançaram sobre as torres gêmeas, haveria um conjunto de justificativas políticas, defensáveis do ponto de vista de Mohammed Atta e dos demais terroristas suicidas, que poderiam ser lidas como uma estratégia de competição. A escolha do alvo principal, um símbolo do capitalismo financeiro globalizado, provavelmente não foi acidental.

Há ainda um pano de fundo histórico que precisa ser compreendido a partir da própria filosofia da história marxista, lembrando-se que “uma das áreas em que mais se concentraram os marxistas analíticos foi a da filosofia marxista da história” (GARGARELLA:idem:115).

Gerald Cohen, recuperando o Marx do *Prefácio para uma contribuição à crítica da Economia política* (1859), afirma que “a história é, fundamentalmente, o desenvolvimento<sup>5</sup> das forças humanas produtivas, e as formas de ascensão e queda das sociedades de acordo como permitiram e promoveram, ou desanimaram, esse desenvolvimento” (*apud* ROEMER, 1986:11)<sup>12</sup>

Embora nesse ponto Cohen explicitamente siga os postulados do marxismo tradicional, de que as forças produtivas da sociedade são um fator explicativo de primeira importância e que essas forças se desenvolvem no decorrer da história, os demais marxistas analíticos em sua maioria acreditam que essa parte da teoria de Marx não deveria ser mantida. Para algumas situações, eles opõem objeções empíricas que mostram que, em muitos casos, o interesse de atores individuais seriam a causa, e não o efeito, do incremento das forças produtivas.

Isso levaria a refletir se os agentes que se destacaram individualmente no 11/9, com suas motivações, interesses e estratégias, estariam formatando um novo momento histórico (ainda que não se possa delimitá-lo), ao invés de serem apenas um subproduto da nova ordem.

Se o 11/9 foi um sucesso do ponto de vista estratégico, a teoria analítica da ação coletiva explicaria que não apenas os interesses do grupo que empreendeu com êxito a ação se mostraram comuns, mas também os agentes envolvidos ponderaram os custos e julgaram válidos os benefícios ao alcance.

Embora pareça difícil aceitar que esse cálculo de interesses tenha sido também realizado pelos pilotos suicidas, convém lembrar que o marxismo analítico, divergindo da teoria da escolha racional pura, não vê a razão subjetiva como livre de desejos e crenças. Ao contrário, autores como Jon Elster (em especial em *Uvas amargas: sobre a subversão da racionalidade*, de 1988) têm claro que as decisões do indivíduo são afetadas por valores e podem se modificar com o tempo. Assim que, um membro da Al Qaeda escolheria participar da organização e execução do ataque suicida às torres gêmeas, e sua escolha seria racional dentro da lógica do grupo.

Em todo caso, pode ter sido a prevalência da dimensão descritiva, interessada em perscrutar microfundamentos e revelar continuidades, sobre a dimensão normativa/moral, que tenha feito o marxismo analítico perder espaço na filosofia política recente.

## **5. O marxismo analítico fornece a lente apropriada?**

É possível afirmar que o marxismo analítico encontrou seu ocaso desde alguns anos, embora os pensadores ligados ao Grupo de Setembro continuem publicando trabalhos regularmente, em temas mais ou menos distantes da pauta marxista tradicional.

As análises continuam mirando criticamente o capitalismo e as sequelas que invariavelmente produz, mas não mais fazendo uso daquelas ferramentas da filosofia analítica. No campo da filosofia política e moral, a ausência de base normativa sustentável em nível deontológico fez com que a teoria da escolha racional fosse excomungada como uma “peste” (BORÓN:2004:137) no final do séc. XX. E isso obviamente fez com que o marxismo analítico tivesse a sua metodologia em certa medida desacreditada para o trato de questões morais.

Em todo caso, é evidente que algumas hipóteses que o marxismo analítico provavelmente lançaria para o exame do 11/9, como aqui se supôs, não restam inteiramente comprometidas. A percepção correta de que as motivações daquilo que foi considerado por muitos como o maior atentado terrorista da história recente foram mais político-econômicas do que culturais, pelo menos assegura ao “non-bullshit marxism” um reconhecimento maior do que se deva dar ao culturalismo.

Do tanto que se pesquisou, não se pôde encontrar qualquer trabalho de autores do Grupo de Setembro que tratasse diretamente do 11/9 e suas razões. Porém, o objetivo do artigo foi o de experimentar os postulados de uma teoria ousada sobre um fato histórico que volta e ocupar o espaço dos debates públicos.

## Referências

- ARNSPERGER, Christian; PARIJS, Philippe Van. **Ética econômica e social**. São Paulo; Loyola, 2003.
- BORON, A. A. **Justiça sem capitalismo, capitalismo sem justiça**. In: VITA, Álvaro de; BORON, A. A. (orgs.). Teoria e filosofia política: A recuperação dos clássicos no debate latino-americano. São Paulo: EDUSP, Buenos Aires: Clacso, 2004.
- BORRADORI, G. **Filosofia em tempos de terror**: diálogos com Jürgen Habermas e Jacques Derrida. Rio de Janeiro: Jorge Zahar Ed., 2004.
- COHEN, G. A. **If you're an egalitarian, how come you're so rich?** Harvard University Press: Cambridge MA and London, 2000.
- \_\_\_\_\_. **Karl Marx's theory of history**: a defence. Oxford University Press Inc. New York, 1978.
- GARGARELLA, Roberto. **As teorias da justiça depois de Rawls**: um breve manual de filosofia política. São Paulo: Martins Fontes, 2008.
- PARIJS, Philippe Van. **Marxism recycled**. Studies in Marxism and Social Theory. Cambridge: Cambridge University Press, 1993.
- PRZEWORSKI, Adam; CHEIBUB, José Antônio and LIMONGI, Fernando. **Democracia e cultura**: uma visão não culturalista. *Lua Nova* [online]. 2003, n.58, pp. 9-35. ISSN 0102-6445.
- ROEMER, John (Ed.). **Analytical marxism**. Studies in Marxism and Social Theorie. Cambridge: Cambridge University Press, 1986.
- SANTOS, José Alcides Figueiredo. **A teoria e a tipologia de classe neomarxista de Erik Olin Wright**. *Dados* [online]. 1998, vol.41, n.2, pp. 377-410. ISSN 0011-5258.
- SAYERS, Sean. Whatever happened to analytical Marxism? **Radical Philosophy**, nº 104, nov./dec. 2000, pp. 39-41.
- XAVIER, Fernando C. C. **Justiça social para além do liberalismo político**. 2006, 156 f. Dissertação (Mestrado em Direitos Fundamentais e Relações Sociais) – Universidade Federal do Pará, Belém, 2006.

## Notas

<sup>1</sup> Artigo apresentado em forma de comunicação no “VII Simpósio Internacional Principia – A filosofia de Nelson Goodman”, em 16 de agosto de 2011, na Universidade Federal de Santa Catarina, em Florianópolis (SC). Na oportunidade, agradeço a advogada Karen Lorene Gomes Beales pela solícita ajuda na revisão do texto.

<sup>2</sup> Giovanna Borradori qualifica o 11 de setembro como “um acontecimento histórico único”, um “*major event*”. Entrevistado por Borradori em relação ao episódio, Habermas referiu que “Também entre a esquerda existe uma consciência generalizada de que estamos vivendo um ponto crucial na história”. Cf. BORRADORI, G. *Filosofia em tempos de terror*: diálogos com Jürgen Habermas e Jacques Derrida. Rio de Janeiro: Jorge Zahar Ed., 2004, p. 9, 38 e 95.

<sup>3</sup> O presente artigo foi pensado e começou a ser escrito antes da morte de Bin Laden, ocorrida em 1º de maio de 2011, em uma polêmica ação militar empreendida pelos Estados Unidos no território do Paquistão.

<sup>4</sup> Jacques Derrida se refere ao 11 de setembro como sendo um “*fait date*”, como “algo [que] marca uma data”, mas reconhecendo, em 2004, em razão de suas preocupações linguísticas, que “não sabemos de fato o que estamos dizendo ou nomeando assim: 11 de setembro, *le 11 septembre*, o 11 de setembro [...] metonímia – um nome, um número – [que] destaca o inqualificável”. In: BORRADORI, op. cit., p. 95-96.

<sup>5</sup> Neste mesmo ano, ocorreu o primeiro atentado terrorista contra o World Trade Center patrocinado por um grupo islâmico fundamentalista. Cf. BORRADORI, op. cit., p. 76.

<sup>6</sup> Mazrui (1997:118) prefere considerar o islamismo como uma “civilização”, em lugar de “cultura”. Cf. MAZRUI, A. Islamic and Western Values. *Foreign Affairs*, vol. 76, n. 5, 1997, pp. 118-132.

<sup>7</sup> Após o 11 de setembro, Huntington, em companhia de Peter L. Berger, organizou o texto *Muitas globalizações. Diversidade cultural no mundo contemporâneo* (Nova York, Oxford University Press, 2002), aproveitando o sucesso editorial que *O choque das civilizações* (Nova York, Foreign Affairs, 1993) havia conquistado.

<sup>8</sup> Figuram como membros fundadores o saudoso Gerald Allan “Jerry” Cohen (Universidade de Oxford), os cientistas políticos Jon Elster (Universidade de Chicago), Robert J. van der Veen (Universidade de Amsterdam) e Adam Przeworski (Universidade de Nova Iorque), o economista e cientista político John Roemer (Universidade de Yale), o filósofo Allen W. Wood (Universidade Cornell), o sociólogo Erik Olin Wright (Universidade de Wisconsin), o historiador Robert Brenner (Universidade da Califórnia), os filósofos políticos Hillel Steiner (Universidade de Manchester) e Phillip Van Parijs (Universidade Católica de Louvain), e os economistas Samuel Bowles (Universidade de Massachusetts) e Pranab Bardhan (Universidade da Califórnia); posteriormente, foram admitidos o economista Thomas Piketty (*Paris School of Economics*) e o filósofo político Joshua Cohen (Universidade de Stanford).

<sup>9</sup> “Não faz muito tempo, Cohen era a figura eminente de um movimento com esse nome [marxismo analítico] que prometeu trazer os padrões autoproclamados da clareza e do rigor da filosofia analítica ao marxismo. Depois da publicação de seu *Teoria da História de Karl Marx* (1978) – o trabalho o mais importante a emergir desse movimento – um grupo de pensadores com ideias comuns rapidamente formou-se. A par de Cohen, outros membros proeminentes incluíam Jon Elster, John Roemer e Erik Olin Wright. Ficaram conhecidos como o ‘Grupo de Setembro’ porque se encontravam anualmente nesse mês, [apresentando-se] de modo alternativo como ‘Grupo do Marxismo ‘Sem-Bobagens’” (tradução livre). Cf. SAYERS, Sean. Whatever happened to analytical Marxism? *Radical Philosophy*, nº 104, nov./dec. 2000, p. 39. Embora o fundador Cohen tenha optado pelo nome jocoso “*Non-Bullshit Marxism*” – aqui traduzido por “Marxismo Sem-Bobagens” –, popularizou-se o nome “Grupo de Setembro”, inclusive pela autorreferência de outros marxistas analíticos como Philippe Van Parijs. Cf. VAN PARIJS, Philippe. *Marxism recycled*. *Studies in Marxism and Social Theory*. Cambridge: Cambridge University Press, 1993, p. xi, xiii, xiv, 4, 210, 245. O “marxismo analítico” recebeu

ainda os nomes alternativos de “*game-theoretic marxism*” e “*rationalchoice marxism*”, devido às influências por ele recebidas da “teoria dos jogos” e da “teoria da escolha racional”.

<sup>10</sup> “Formações coletivas, não-indivíduos atomizados, são os veículos característicos das lutas de classes, ou seja, das formas organizadas de práticas de classe *antagonísticas*” (Wright, 1997:381-382; 1985:145).

<sup>11</sup> Donald Nonini diz que “[...] entre as figuras periféricas da camarilha oligárquica, inclui-se elites individuais que não são compostas por cidadãos americanos, mas [inter allia] por membros da família real saudita [...] e até o 11 / 9, os membros da família Bin Laden”. NONINI, Donald M. Making the case for kleptocratic oligarchy: (as the dominant form of rule in the United States). In: KAPFERER, Bruce (ed.). Oligarchs and oligopolies: new formations of global power. *Social Analysis* 49, 1, 177-189, 2005.

<sup>12</sup> Na esteira, o autor ainda respalda a definição de história das sociedades do Marx do *Manifesto Comunista*, a saber, que a história é, em termos gerais, a própria “luta de classes” (idem:19). Embora Cohen transcreva passagem em que Marx se refere ao “desenvolvimento de suas forças [dos homens] produtivas materiais” (em inglês: *development of their material productive forces*), ele (Cohen) não utiliza *development* (desenvolvimento), e sim *growth*, que ainda assim se optou por traduzir como “desenvolvimento”, inclusive porque de acordo com a terminologia alemã (*Entwicklung*).

# **A QUESTÃO DOS DEVERES INDIRETOS PARA COM OS ANIMAIS NÃO-HUMANOS SEGUNDO A FILOSOFIA MORAL KANTIANA**

GABRIEL GARMENDIA DA TRINDADE<sup>1</sup>

*Mestrando pelo Programa de Pós-Graduação em Filosofia – PPGF da Universidade Federal de Santa Maria (UFSM-RS). Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).*

garmendia\_gabriel@hotmail.com

LAUREN DE LACERDA NUNES<sup>2</sup>

*Doutoranda pelo Programa de Pós-Graduação em Filosofia – PPGF da Universidade Federal de Santa Maria (UFSM-RS). Professora assistente na área de humanidades da Universidade Federal do Pampa (Unipampa-RS), campus São Borja.*

laurenlacerdanunes@gmail.com

Pode-se afirmar que a questão sobre qual deve ser o tratamento dispensado aos animais possui caráter secundário no que tange à filosofia moral de Kant. A referida temática é examinada em apenas poucos e esparsos parágrafos nas obras de cunho ético do filósofo de Königsberg. Em observância a isso, o presente trabalho, através de um estudo histórico-bibliográfico, objetiva, primeiramente, localizar e detalhar certas passagens, especialmente na obra *Metafísica dos Costumes* de 1797, nas quais são abordadas questões concernentes ao relacionamento moral entre humanos e animais. Em segundo lugar, almeja-se analisar e problematizar a respectiva temática tendo por base as colocações de comentadores e autores kantianos contemporâneos.

A justificativa para o desenvolvimento e consecução dessa pesquisa, baseia-se principalmente na escassez – para não falar na nulidade – de textos abordando a visão kantiana sobre o estatuto ético-filosófico outorgado aos animais. Embora exista uma enorme bibliografia versando sobre as relações de índole moral e jurídica estabelecidas entre homens e animais, poucos artigos e ensaios em língua inglesa ou espanhola tratam desse assunto no ínterim da filosofia moral kantiana. Mais complicada ainda é a situação das pesquisas em português, haja vista que até o momento não foram encontrados trabalhos que façam uma análise crítica aprofundada da situação dos animais não-humanos dentro da filosofia moral formulada por Kant. As poucas obras que citam a concepção kantiana acerca dessa temática o fazem quase que exclusivamente em nível de curiosidade histórico-filosófica. Nesse sentido, o presente texto visa suprir, ainda que superficialmente, essa visível lacuna teórica referente ao exame apropriado em nível holístico do pensamento moral kantiano.

## Considerações iniciais

Antes de se adentrar no estudo acerca das relações entre humanos e animais, é preciso esclarecer alguns aspectos basilares da abordagem moral elaborada por Kant (2008). Nesse sentido, primeiramente, há de se ressaltar que Kant (2008) não construiu um sistema ético que se preocupasse com as consequências da ação praticada, e sim com os princípios e normas que regeriam todo agir humano e a sua fundamentação. Ao adotar esse ponto de partida, Kant (2008) teria buscado fundar a moralidade baseando-a na racionalidade e liberdade humanas. Postulou que a origem da moral encontrar-se-ia na autonomia e na autolegislação da vontade. Para ele, o homem seria um ser livre e dotado de racionalidade, e sobre estes dois pressupostos deveriam estar baseados todos os princípios de uma moralidade que aspirasse a validade objetiva.

Tendo isso em vista, poder-se-ia afirmar que o dever seria o princípio supremo da moralidade. Kant (2008) pretende encontrar a justificação objetiva e universal do dever. Para tanto, inicia sua análise realizando a fundamentação de sua ética através de um exame crítico da razão prática. Necessário ser dito que, ao adentrar neste campo de estudo, Kant não desvinculou a razão prática da teórica, explorada em sua *Crítica da Razão Pura*. Tratar-se-ia apenas, de um novo âmbito de sua aplicação, o campo prático, onde a razão seria chamada a desempenhar seu papel na avaliação e justificação dos juízos práticos, da mesma forma que ela teria sido convocada em seu uso teórico a buscar fundamentação objetiva para o conhecimento empírico.

A faculdade da razão prática seria responsável pela capacidade do ser racional de escolher suas ações independentemente de fundamentos determinantes sensíveis. Sendo assim, ele agiria puramente por dever. A razão prática é identificada com o próprio dever. E, além disso, pelo motivo próprio de ser uma faculdade capaz de tornar o ser racional livre das inclinações sensíveis, formularia as próprias leis a que se submetesse. Após esta breve exposição acerca da ética kantiana, no sentido de elucidar os seus principais pontos e o quanto a mesma é focada nos conceitos de dever e de racionalidade, fica claro porque pode ser difícil entabular uma discussão acerca do estatuto moral outorgado aos animais considerando-a como pano de fundo. Por isso, uma busca por referências contemporâneas<sup>1</sup> e releituras de Kant foi inevitável.

Nesse sentido, foi feita a leitura e análise do artigo *Fellow creatures: Kantian Ethics and Our Duties to Animals* (2004). Nesse, a filósofa norte-americana Christine Korsgaard examina as diferentes facetas do pensamento kantiano no que se refere à noção de racionalidade, bem como outros tópicos pertinentes ao debate concernente status moral dos animais. Depois de feita a leitura do referido ensaio, evidenciou-se que, para Kant, a racionalidade “é a capacidade de autogoverno normativo. Ela torna-nos capazes de aferir e julgar os princípios que regem nossas crenças e ações, e regulamentar as mesmas de acordo com esses juízos” (KORSGAARD, 2004, p. 11).

Em síntese, a racionalidade, no que diz respeito à filosofia kantiana prática, não é a capacidade de escrever, falar ou resolver problemas, mas sim é concebida como uma faculdade humana de determinar a si mesmo a agir de acordo com a representação de certas leis. Não obstante, segundo Kant (2008), tal característica se encontra presente apenas em seres racionais. De fato, como faz questão de salientar Korsgaard, “embora um ser racional possa ser descrito como agindo de acordo com certos princípios em suas crenças e ações, tais princípios não são objeto de sua atenção” (KORSGAARD, 2004, p. 9-10). Por outro lado, no que concerne aos animais racionais, esses “*avaliam* os princípios que governam suas crenças e ações” (KORSGAARD, 2004, p.10).

### **Os animais na *Metafísica dos Costumes* (1797)**

A principal obra de Kant a ser tratada no presente artigo é a *Metafísica dos Costumes*, a qual foi originalmente publicada em 1797, sendo dividida em duas partes distintas, porém complementares: a *Rechtslehre* (Doutrina do Direito) e a *Tugendlehre* (Doutrina da Virtude). Em observância a isso, salienta-se que a análise a seguir versa exclusivamente sobre duas passagens específicas da Doutrina da Virtude. Todavia, é importante frisar que outras questões referentes à Doutrina do Direito e ao estatuto ético-jurídico outorgado aos animais dentro do sistema kantiano possivelmente serão examinadas em pesquisas futuras pelos autores.

Não obstante, alguns esclarecimentos cabem ser feitos acerca da referida obra de Kant, para que se compreenda o motivo pelo qual ela foi escolhida para o presente trabalho, bem como o seu papel em comparação às outras obras de Kant. A *Metafísica dos Costumes* deve ser compreendida como sendo o resultado aplicável da fundamentação adequada de uma ciência da moral iniciada nos escritos de 1785. Nela Kant (2008) dedicar-se-á a apresentar as diretrizes teóricas e práticas de sua proposta ético-filosófica pautada pela razão. Ressaltam-se duas passagens centrais para a compreensão e o detalhamento da temática sugerida por esse trabalho. A primeira diz respeito à distinção entre *humanidade* e *animalidade*, a qual está diretamente ligada à racionalidade e ao sentimento moral:

A faculdade de fixar-se um fim – qualquer fim que seja – é o que caracteriza a humanidade (enquanto distinta da animalidade). Por conseguinte, está vinculada também ao fim da humanidade pela cultura em geral, no sentido de procurar obter ou estimular a faculdade de realizar todas as modalidades de fins possíveis, na medida em que isso é para ser encontrado num ser humano ele mesmo. Em outras palavras, o ser humano tem um dever de cultivar as rudes inclinações de sua natureza, através do que o animal é, num primeiro momento, promovido a ser humano. (KANT, 2008, p. 235-236)

Para Kant (2008), graças à capacidade de distinção da razão, um ser racional possuiria duas perspectivas: enquanto pertencente ao mundo sensível sob o

comando de leis naturais que o ordenariam heteronomamente, e enquanto pertencente ao mundo inteligível sob o comando de leis independentes da natureza, não empíricas, que se fundamentariam apenas na sua própria razão, portanto, ordenando-o autonomamente. Assim, é possível dizer que, ao adotar o ponto de vista de um ser racional enquanto pertencente ao mundo inteligível, o homem não teria como intuir a causalidade de sua própria vontade senão sob a idéia de liberdade, pois esta significaria a independência das inclinações determinantes do mundo sensível, e a razão deveria ser sempre capaz de considerar a si mesma como independente de tais causas externas a si.

Por outro lado, como pontua Korsgaard, “um animal inteligente, porém não-racional, pode ser compelido a crer ou esperar algo, quando percebe outra coisa, pois aprendeu a fazer certas conexões ou associações entre coisas distintas anteriormente” (KORSGAARD, 2004, p. 9). Ainda, a filósofa argumenta também que “nós, como animais racionais, estamos cientes de que somos inclinados a tomar uma coisa como evidência para outra e, portanto, podemos nos questionar se devemos fazê-lo”. (KORSGAARD, 2004, p. 9)

Ademais, poder-se-ia compreender a noção de animalidade distinta de humanidade na medida em que na primeira há uma ausência da faculdade de distinguir e determinar fins para suas ações. Além disso, faz-se necessário ressaltar que a noção de humanidade se relaciona com a concepção kantiana de sentimento moral. Haja vista que a consciência da obrigação, segundo Kant (2008), é dependente do sentimento moral, na medida em que este está ligado ao próprio constrangimento referente ao pensamento do dever. Com efeito, a ausência do sentimento moral resultaria na dissolução daquilo que torna alguém um ser humano.

Faz-se necessário ressaltar, todavia, que a diferenciação entre homens e animais, bem como a fundamentação de uma filosofia moral pautada pela razão não é uma abordagem imune a críticas. Priscilla Cohn, por exemplo, em seu texto *Kant y el problema de los derechos de los animales* (1987), apresenta algumas das dificuldades existentes na sustentação teórica do pensamento kantiano. Segundo Cohn (1987), ao afirmar que é a racionalidade o passaporte para a entrada na comunidade moral, Kant conseqüentemente precisa comprometer-se com a ideia de que “visto que todos os seres humanos são agentes morais, todos os seres humanos são racionais” (COHN, 1987, p. 202) – sob pena de ser inconsistente com a sua proposta se não o fizer.

De fato, ao se analisar as colocações de Cohn (1987), se é a capacidade de determinar-se a agir de acordo com certas leis a característica que define *per se* o agente moral na obra kantiana, então seria possível concluir que nem todos os seres humanos são agentes morais, pois muitos são incapazes de demonstrar tal faculdade. Crianças muito pequenas, por exemplo, não apresentam esse fator tão essencial para a filosofia moral kantiana. Com efeito, poder-se-ia extrair dessa constatação uma possível exclusão delas da comunidade moral humana? Muitos autores afirmam que não, na medida em que essas crianças são tidas como sendo *potencialmente racionais*. Todavia, como Cohn (1987) faz questão de frisar, tal

resposta não é satisfatória, haja vista que não resolve o problema em nenhum sentido significativo. Pois, “Kant mantém que o valor básico de um indivíduo está radicado em sua liberdade, ou em sua capacidade de regulamentar a si mesmo a partir da lei moral, a lei dada pela razão” (COHN, 1987, p. 202). Por conseguinte, Cohn assevera que “um ser apenas potencialmente racional carece desta capacidade e, portanto, carece a si mesmo de valor moral”. (COHN, 1987, p. 202).

Os apontamentos levantados por Cohn (1987) possibilitam introduzir apropriadamente um dos aspectos centrais da segunda passagem da *Metafísica dos Costumes* a ser aqui trabalhada. Nesse sentido, o referido excerto será abordado sob dois prismas diferentes. O primeiro concerne à questão da crueldade no trato com os animais, enquanto o segundo diz respeito à argumentação kantiana referente ao caso dos deveres indiretos para com os animais. Assim, segundo Kant:

No que toca à parte animada, mas destituída de razão, da criação, o tratamento violento e cruel dos animais é muitíssimo mais estreitamente oposto ao dever de um ser humano para consigo mesmo e ele tem o dever de abster-se de tal prática, pois esta embota seu sentimento compartilhado do sofrimento deles, de modo a enfraquecer e gradualmente desarraigar uma predisposição natural que é muito útil à moralidade nas nossas relações com outros seres humanos. O ser humano está autorizado a matar animais rapidamente (sem produzir sofrimento) e submetê-los a um trabalho que não os force além de suas forças (trabalho ao qual ele mesmo deve submeter-se). Mas experimentos físicos que sejam dolorosos aos animais a serviço da mera especulação, quando o objetivo almejado poderia também ser atingido os dispensando, se apresentam como abomináveis. A gratidão ao longo serviço prestado por um velho cavalo ou um velho cão (tal como se fossem membros da comunidade doméstica) diz respeito *indiretamente* ao dever de um ser humano *em sua consideração* a esses animais; do prisma de um dever direto, todavia, é sempre somente um dever do ser humano *para consigo mesmo*. (KANT, 2008, p. 285-286)

No tocante à primeira questão a ser discutida, a respectiva passagem visivelmente demonstra a preocupação de Kant (2008) referente às implicações para outros seres racionais de um tratamento inadequado para com os animais. Ou seja, o agir violento para com os animais é, em primeiro lugar, nocivo aos interesses éticos dos homens na medida em que tal prática embotaria os sentimentos empáticos naturais do ser humano, sentimentos os quais são significativos à moralidade. Nesse sentido, é possível notar que as obrigações que os seres humanos possuem para com os membros de outras espécies são apenas indiretas. Pois, os desdobramentos negativos da crueldade no manejo dos animais dizem respeito somente à possibilidade de uma atuação cruel para com outros seres racionais.

Em outras palavras, Kant (2008) condena a atuação cruel no trato para com os animais devido ao fato de que isso resultaria no embotamento de um possível sentimento de compaixão relativo ao sofrimento desses seres. Nesse contexto, tal “dessensibilização” concernente às moléstias e mazelas dos animais poderia

acabar refletindo no tratamento moral devido a outros seres humanos, agentes morais racionais. Se um indivíduo atua violentamente em sua interação com os animais, tal comportamento pode ser estendido aos homens, o que seria contrário aos seus deveres. Com efeito, torna-se claro que os animais são alvos apenas de obrigações indiretas por parte dos seres humanos. Nesse sentido, Cohn (1987) afirma que, de acordo com a perspectiva sustentada por Kant, afirmar que os deveres para com os animais são indiretos “equivale a dizer que os animais devem ser excluídos do reino moral devido a sua falta de racionalidade<sup>2</sup>” (COHN, 1987, p. 202). Por conseguinte, fica patente uma hierarquização de deveres no que tange à filosofia moral kantiana. Acerca disso, Cohn salienta:

Nosso dever primário é para com os seres racionais que são fins em si mesmos, porém também temos deveres indiretos para com os animais e deveres de menor importância no que diz respeito a seres inanimados ou coisas. [...] Não há dificuldade em se aceitar a atitude de Kant com respeito aos objetos inanimados, sejam fabricados ou naturais. Sem embargo, fica claro que Kant não nos diz que nossos deveres indiretos para com os objetos inanimados sejam idênticos a nossos deveres para com os animados, pois ainda que Kant às vezes chame de “coisas” ou “objetos” os animais, ele sempre distingue entre animais e objetos inanimados, seja verbalmente, ou dando exemplos das classes de coisas a que está se referindo. (COHN, 1987, p. 198)

Antes de se adentrar em um exame mais aprofundado sobre o caráter dos deveres indiretos para com os animais, é preciso fazer um breve esclarecimento acerca da natureza, bem como a distinção entre deveres no âmago do sistema ético kantiano. Tendo isso em vista, o texto *Kant on duties regarding nonrational nature* (1998) de A. Wood e O. O’Neill foi analisado por ser bastante esclarecedor no que tange o exame e detalhamento desse assunto. Os autores colocam que, a primeira divisão dos deveres relativa à filosofia moral proposta por Kant é entre “deveres de direito, ou deveres judiciais e deveres de virtude, ou deveres éticos” (WOOD & O’NEILL, 1998, p. 192). Mais especificamente, “deveres de direito são aqueles que podem ser coercivamente obedecidos por intermédio da lei e do estado. Deveres de virtude são aqueles onde o agente moral é constrangido somente internamente pela razão” (WOOD, & O’NEILL, 1998, p. 192).

Os deveres de direito não serão alvo da análise desse trabalho, embora seja necessário salientar que não foi encontrado nenhum estudo específico pertinente a qualquer tentativa de abarcar os animais dentro de alguma forma de legislação tendo por base o sistema jurídico kantiano<sup>5</sup>. Por outro lado, no que se refere aos deveres de virtude, uma diferenciação basilar deve ser feita, a entre *deveres para consigo mesmo* e *deveres para com os outros*. Acerca disso WOOD & O’NEILL esclarecem:

A divisão kantiana de deveres éticos em deveres para conosco e deveres para com os outros pode ser considerada corolário do princípio da personificação. O dever *d* é um dever *com respeito* a S, se e somente se S é um ser racional (ou mais de um), e a condição moral para cumprir *d* está

fundada na condição moral de respeitar a humanidade *na pessoa de S.*  
(WOOD & O'NEILL, 1998, p. 193)

Em outras palavras, os *deveres para conosco* concernem ao respeito devido à própria humanidade do agente moral em sua própria pessoa<sup>4</sup> enquanto *deveres para com os outros* se referem ao respeito devido a outros seres racionais, que são considerados como fins em si, algo expresso por uma das formulações do imperativo categórico, a fórmula da humanidade<sup>5</sup>: “Age de tal maneira que uses a humanidade, tanto na tua pessoa como na pessoa de qualquer outro, sempre e simultaneamente como fim e nunca simplesmente como meio” (KANT, 2008, p.69).

Tendo em vista o exposto, é possível dar continuidade ao detalhamento sobre os deveres indiretos para com os animais, assim como localizar o seu espaço dentro da esfera moral construída por Kant (2008). Assim, na medida em que as obrigações possuídas pelos seres humanos com respeito aos animais são apenas indiretas, poder-se-ia questionar em que sentido elas realmente seriam capazes de resguardar esses animais da imposição de sofrimentos que possam ser caracterizados como desnecessários ou inaceitáveis. Em outras palavras, devido ao fato de que os animais se encontram em um degrau abaixo na hierarquia moral edificada por Kant (2008), de que maneira esses seres poderiam realmente ser protegidos eficientemente?

De fato, Kant (2008) defende que certas práticas relativas ao tratamento dos animais se apresentam como abomináveis à moralidade. Contudo, Kant (2008) não estabelece um padrão ou mesmo oferece um critério que possa ser utilizado para distinguir entre, por exemplo, experimentos e procedimentos aceitáveis ou inaceitáveis. Nesse sentido, fica evidente que qualquer interesse que os membros de outras espécies possam ter em não sofrer permanece abaixo dos desejos e preferências humanas, o que implica no fato de que os animais se encontram permanentemente a mercê dos homens. Em observância a isso, Korsgaard (2004) ressalta um importante aspecto da filosofia moral kantiana que poderia ser tomado como base para solucionar esse dilema ao menos preliminarmente. Dessa forma, segundo Korsgaard:

Se possuímos obrigações concernentes aos animais, elas poderiam resultar apenas de normas que legislamos para nós mesmos. A única fonte possível da lei e da obrigação é a vontade racional, e *nesse sentido*, um animal não-racional não poderia ser fonte de obrigações. Porém, disso não se segue que outros animais não possam ser fins-em-si-mesmos no primeiro sentido – a origem de reivindicações normativas legítimas – por isso, não se segue que não existiria algum sentido no qual eles poderiam nos obrigar. Segundo Kant, leis são pela sua própria natureza universais, e uma lei universal pode estender sua proteção a alguém que não participa, e não poderia ter participado de sua legislação. [...] O fato de que os animais não-humanos não podem participar na legislação moral é insuficiente para se estabelecer que eles não possam nos obrigar nesse último sentido. Então, a questão é se nós, seres humanos, tomaríamos isso como algo necessário, sob a base de uma reflexão racional, e decidiríamos por ter

leis cuja proteção se estenderia a outros animais. (KORSGAARD, 2004, p. 21-22)

Em última instância, evidencia-se aqui uma peculiaridade muito relevante para a discussão levantada. Embora os animais não se apresentem como seres racionais, isto é, não demonstrem a faculdade que serve de sustentáculo primeiro para a filosofia moral kantiana, isso não quer dizer que eles não possam vir a ser adequadamente salvaguardados. Ou seja, poder-se-ia formular leis universais de tratamento aos animais<sup>6</sup> mesmo em um sistema ético-filosófico onde eles não são tidos como sujeitos morais.

## Considerações finais

A fim de concluir esta análise, é necessário retomar alguns pontos. O primeiro deles é notar que, para Kant (2008), os deveres que os seres humanos têm para com os animais são apenas indiretos. Todavia, Kant (2008) também salienta que a crueldade no tratamento dos animais precisara ser evitada, pois se mostraria contrária ao dever. Isso se dá porque as obrigações e deveres dos agentes morais, segundo a perspectiva kantiana, estão intrinsecamente ligados a relação que esses possuem com outros seres racionais.

Por um lado, existiriam os deveres para consigo mesmo, ou seja, referentes ao respeito à humanidade do próprio agente moral, ao passo que, por outro lado, haveria os deveres para com os outros, referentes à consideração devida à humanidade alheia. Ademais, devido a sua falta de racionalidade, isto é, a capacidade de entender a concessão da lei em si mesma, os animais estariam fora da comunidade moral, ainda que permaneçam sendo alvo de obrigações humanas de maneira indireta. Contudo, eles ainda assim poderiam ser protegidos através de leis morais específicas, se isso for matéria de consideração da vontade geral. Evidentemente, a análise apresentada nessa pesquisa é preliminar, haja vista que seu intento primeiro é localizar o lugar no qual se encontram os animais dentro da filosofia moral kantiana. Por conseguinte, outras questões relativas aos tópicos aqui debatidos deveriam ser problematizadas devido à valiosa contribuição que um resgate do pensamento kantiano acerca desta temática traria à ética contemporânea.

## Referências

- COHN, P. Kant y el problema de los derechos de los animales. In: GUISÁN, E. **Esplendor y miseria de la ética kantiana**. Barcelona: Editorial Anthropos, 1988.
- KANT, I. **Começo conjectural da história humana**. São Paulo: UNESP, 2010.

KANT, I. **Fundamentação da metafísica dos costumes**. Lisboa: Edições 70, 2008.

KANT, I. **Metafísica dos costumes**. 2ª ed. rev. São Paulo: Edipro, 2008.

KORSGAARD, C. M. Fellow Creatures: Kantian ethics and our duties to animals. In: **The Tanner Lectures on Human Values**. University of Michigan, 2004.

WOOD, A. W. & O'NEILL, O. Kant on duties regarding nonrational nature. **Proceedings of the Aristotelian Society**. . Vol. 72. 1998. p. 189-228.

## Notas

<sup>1</sup> As traduções presentes no corpo do texto foram realizadas livremente pelos autores.

<sup>2</sup> Ainda sobre esse tópico, Cohn argumenta: O que podemos concluir, então, em termos do pensamento kantiano, se aceitarmos que os animais não podem ter uma concepção de “lei”? Para Kant, se segue disso que os animais não podem ter deveres e, portanto, que não podem cumpri-los. Esta conclusão está de acordo com o senso comum, já que não censuramos um leão por matar e comer um antílope. O leão não pode fazer outra coisa além de matar sua presa se almeja sobreviver; caso contrário, morreria de fome. De todo modo, não nos ocorre pensar que o leão se comporta de maneira imoral, porque ele não possui noção nenhuma do que vem a ser a morte, o sofrimento, etc. (COHN, 1987, p. 203).

<sup>3</sup> Sobre essa questão, WOOD & O'NEILL (1998) asseveram que nós não temos deveres de direito para com a natureza não-racional, haja vista que apenas seres racionais finitos (em nossa experiência, seres humanos) possuem *direitos* legais (MS 6:241). Quaisquer deveres judiciais que possamos ter envolvendo o tratamento da natureza não-racional devem ser resultado dos direitos dos seres humanos e das leis criadas pela vontade geral do estado – por exemplo, seu direito sobre sua propriedade de coisas não-racionais e de leis que promovam o bem comum ou a consecução de seus deveres morais coletivos (tais como os deveres de caridade para com os pobres) (MS 6:325–328). Kant não menciona nenhum dever jurídico específico acerca dos animais ou do ambiente natural sob esse título, porém é digno de nota mencionar que haveria lugar para eles. Na teoria de Kant, o fato de que seres não-racionais não possuem direitos não implica que a vontade geral do estado não possa legislar restrições sobre como eles poderiam ser usados ou tratados.

<sup>4</sup> A noção de *pessoa* dentro do sistema moral kantiano é descrita mais detalhadamente em sua obra *Fundamentação da Metafísica dos Costumes*, datada de 1785. Nessa, Kant argumenta: Os seres cuja existência depende, não em verdade da nossa vontade, têm contudo, se são seres irracionais, apenas um valor relativo como meio e por isso se chamam *coisas*, ao passo que os seres racionais se chamam *peças*, porque a sua natureza os distingue já como fins em si mesmos, quer dizer como algo que não pode ser empregado como simples meio e que, por conseguinte, limita nessa medida todo o arbítrio (e é um objecto do respeito) (KANT, 2008, p. 70).

<sup>5</sup> Em seu artigo de 1786 denominado *Começo conjectural da história humana*, Kant vale-se da história do Jardim do Éden como uma alegoria para descrever o processo que levaria à origem da autoconsciência e o desenvolvimento da moralidade em quatro passos. No quarto e último passo, Kant apresenta o momento em que o homem tomou consciência de seu lugar no mundo em relação aos outros animais: “O quarto e último passo da razão, que eleva o homem acima da sociedade com os animais, foi fazê-lo compreender (ainda de maneira obscura) que era propriamente *o fim da natureza*, e nada do que vive sobre a terra poderia fazê-lo renunciar a isso. Na primeira vez que disse à ovelha: “A pele que portas, não te deu a natureza para ti, mas para mim”, arrancando-a e com ela vestindo-se (Gênesis, 3: 21), descobriu um privilégio que tinha, em razão de sua natureza, sobre todos os animais, já não considerados companheiros na criação, mas sim meios e instrumentos colocados

à disposição de sua vontade para o sucesso de seus propósitos.” (KANT, 2010. p. 22). Sobre essa passagem em especial, Korsgaard afirma: “Nossa compreensão de que somos fins em nós mesmos está aqui firmemente ligada com o momento em que deixamos de considerar os outros animais como companheiros de criação, e passamos a considerá-los como meros meios em vez disso.” (KORSGAARD, 2004. p. 13).

<sup>6</sup> Tais leis serão tópico central em um futuro ensaio sobre essa mesma temática.

# RESPOSTAS AO DILEMA MALTHUSIANO NAS CONFERÊNCIAS SOBRE EVOLUÇÃO E ÉTICA DE HUXLEY E DEWEY

JOSÉ CLÁUDIO MORELLI MATOS

UDESC

doutortodd@gmail.com

*Acostuma-te à lama que te espera!  
O homem, que nesta terra detestável  
Habita entre feras, sente, inevitável  
Necessidade de também ser fera.  
(Augusto dos Anjos)*

**Resumo:** A relação entre os princípios naturais e os princípios éticos é objeto de inúmeras discussões. Uma das mais polêmicas é o *Ensaio Sobre a População* de Malthus, que serviu de inspiração para Darwin. A conferência de Huxley sobre “Evolução e Ética” pode ser entendida como uma resposta ao argumento de Malthus sobre a ética estar subordinada ao poder de princípios naturais. Mas a estratégia de Huxley é criticada por Dewey que, ao escrever sobre o tema, oferece uma consideração sobre a relação entre evolução e ética, em que ambas são componentes de um mesmo processo. Para Dewey, não há oposição entre elas, mas sim, o desenvolvimento de um princípio evolutivo, que atua do mesmo modo sobre os valores e linhas de conduta, assim como sobre os seres vivos na natureza.

## Introdução

Este trabalho tem o objetivo de examinar a relação entre os princípios naturais e os princípios éticos, a partir do pensamento de Thomas Malthus, Thomas Huxley e John Dewey. Em termos da história intelectual, é interessante lembrar que Malthus, como um dos precursores das ciências humanas na passagem do século XVIII para o XIX, empregou o método experimental inspirado no trabalho de Newton, a fim de tratar o tema da população. Charles Darwin, quando buscava um mecanismo com o qual explicar o enorme volume de observações que indicavam uma ancestralidade comum das espécies, julgou encontrar no princípio da luta pela existência, formulado por Malthus, o mecanismo que buscava. Como adepto do darwinismo, Thomas Huxley parece, ainda assim, manter uma concepção do ser humano que deposita confiança nas escolhas e valores morais, e resguardar a virtude ética contra a árida lei natural da luta pela

existência. Huxley supõe que os princípios éticos seguem em direção oposta ao processo de evolução por seleção natural. Dewey, finalmente, ao comentar as idéias de Huxley, reveste de novo significado os termos “adaptado”, “seleção natural” e “luta pela existência” a fim de entender a ética e a natureza humana sob um mesmo princípio, regente de toda a mudança evolutiva. A possibilidade aberta pelo ponto de vista de Dewey é, portanto, a de superar qualquer oposição entre os princípios da natureza física e biológica e os da ética. O resultado é uma concepção naturalista e evolutiva da ética, que leva a novos fundamentos para seu estudo e reflexão.

Do ponto de vista conceitual, o elemento que une as noções destes autores é a preocupação com o efeito dos princípios naturais sobre a dimensão ética. Parte-se da leitura de um ancestral remoto do pensamento evolucionista – Thomas Malthus – e de uma interpretação viável de seu pensamento, que o aproxima das indagações éticas.

O alvo final do presente esforço é compreender a contribuição deweyana a esta discussão. Tal contribuição, conforme se espera mostrar, delega um papel central à evolução, assim como reconhece a necessidade incontornável de que o processo ético esteja pautado pelas condições individuais e sociais de crescimento e de reconstrução dos valores e das condutas.

## **Delimitação da questão**

O que se chama aqui de “dilema malthusiano” não se refere especificamente ao tema da população, muito menos a aspectos da economia política. Trata-se, antes, daquilo a que George G. Simpson (1949) chamou de “ética das presas e garras”: A noção de que a luta pela sobrevivência imposta pela condição do homem na natureza determina a perspectiva ética das sociedades humanas. Após a popularização da teoria da seleção natural, a relação entre a natureza humana e a ética, assim como a possibilidade de avanço no campo dos valores passa a ser objeto de acalorada discussão. Entre os antecedentes do pensamento darwiniano, o autor cuja obra manifesta mais nitidamente a “ética das presas e garras” é Thomas Malthus no *Ensaio Sobre a População* (1798).

O ponto aqui discutido é se o desenvolvimento ético está ou não submetido, de alguma forma, às regularidades e condições naturais em que os agentes se encontram. O ponto do presente trabalho, então, não é o tema da população e nem a questão ética relacionada com o controle de natalidade, ou com as medidas de auxílio e proteção aos mais pobres. Estes assuntos são considerados da maior relevância teórica e social, mas estão fora do recorte principal deste artigo, pela óbvia razão de evitar-se abrir demais o foco da discussão e perder de vista, por isso, o próprio foco.

A título de ilustração e informação histórica, contudo, é possível fazer aqui referência ao artigo de H. S. Spalding intitulado “Ethics and the Neo-

malthusianism”. É um trabalho publicado em 1917, quando a polêmica em torno das conseqüências éticas da teoria evolutiva fazia parte da ordem do dia entre intelectuais e leigos. O artigo não é muito mais do que um manifesto contra o controle de natalidade, escrito por um militante católico, e dirigido contra um ponto de vista que ele denomina de Neo-malthusianismo. Segundo ele, no século XX, a doutrina original de Malthus sobre a população já caiu no descrédito.

Mas o ensinamento de Malthus apareceu novamente: o neo-malthusianismo. De novo ele advoga a limitação da taxa de nascimentos, não por que a Terra não pode suportar uma raça humana mais numerosa, mas por que ele almeja uma raça mais saudável, mais perfeita e mais adaptada (Spalding, 1917, p. 610).

O autor deste artigo está, de fato, pretendendo discutir os termos éticos de uma política de controle de natalidade. Seria inoportuno fazer referência a este trabalho, em virtude da diferença na abordagem, não fosse o fato de que ele concebe o neo-malthusianismo como correlato ao pragmatismo – de que Dewey é um representante – num sentido muito especial do termo. Ele afirma:

Nós discordamos inteiramente destes pragmatistas que defendem não haver algo intrinsecamente mal, que a moralidade e a bondade são questões de costume, que aquilo que é errado hoje pode ser certo daqui a cinquenta anos. O auto-abuso e a prostituição sempre foram errados, são errados e sempre serão errados. Eles são errados por sua própria natureza. Por que? Pela mesma razão que a gula e a bebedeira são errados. Elas são contra a natureza (Spalding, 1917, p. 611).

O problema identificado nesta concepção, que toca mais de perto na temática aqui discutida, é que ela ignora o processo evolutivo, tratando “natureza” como um quadro fixo. Assim como considera os contextos sociais onde os valores funcionam como formas de dirigir a atividade das pessoas como um quadro fixo, do qual se pode falar por meio de conceitos imutáveis e absolutos. O pragmatismo criticado por Spalding pode ser reconhecido na atitude de Dewey, conforme será mostrado adiante, segundo a qual é preciso examinar as questões éticas a partir do método evolutivo, justamente porque a mudança nas condições de vida implica, em resposta, uma mudança nas condutas e nos valores.

Ao tratar, logo abaixo da passagem acima citada, de “sentimentos e paixões que a natureza implantou na espécie humana para um propósito bom e sagrado” (Spalding, 1917, p. 611), seu autor está manifestando um injustificado fixismo, contrário a todo o conhecimento disponível sobre a natureza e a sociedade. Mas está, também, referindo-se a um desígnio intencional operando na natureza. A revolução darwiniana ataca frontalmente esta suposição de desígnio, e a substitui por regularidades impessoais e sem intenção proposital, como é o caso do mecanismo de seleção natural. Todo o propósito de que se tem evidência é o propósito dos seres humanos que, em circunstâncias determinadas no meio natural e social, arregimentam os meios disponíveis a fim de satisfazer sua necessidade de adaptação.

## O poder da lei natural segundo Malthus

Thomas Malthus, no *Ensaio Sobre a População*, desenvolve um argumento para mostrar que a melhora nas condições da vida nas sociedades humanas é constantemente limitado em razão de uma incontornável tendência natural da população ao crescimento. Declaradamente, o autor escreve em resposta ao argumento da perfectibilidade do ser humano, defendido por autores como Condorcet e Godwin.

Em seu livro intitulado *Malthus*, Jean-Marie Poursin e Gabriel Dupuy procuram desenvolver um comentário acerca do *Ensaio Sobre a População*, tentando compreender as teses de seu autor, e ao mesmo tempo desmistificar as inúmeras atribuições e rótulos que ele tem recebido da posteridade. A imagem de um pensador preocupado com os exageros do humanismo iluminista, e com a falta de fundamentos para inferir o progresso constante da sociedade emerge desta tentativa. É assim que o *Ensaio Sobre a População* pode ser descrito como um “Panfleto político dirigido contra certos utopistas inspirados pela Revolução Francesa” (Poursin & Dupuy, 1972, p. 9). Mas para os objetivos da presente discussão, o que é mais relevante é a consideração feita por Malthus da situação do homem no mundo natural. Malthus concebe que o ser humano é um ser vivo na natureza, constituído e sujeito a seus princípios e regularidades. A sociedade, os valores éticos, as relações econômicas e todas as realizações culturais humanas permanecem sob o alcance inevitável dos princípios naturais. Esta consideração é bem percebida por Poursin e Dupuy, que afirmam sobre isso:

Essa reintegração do homem no seio da grande comunidade dos seres vivos está de acordo com o esforço precedente do movimento científico, que nega ao homem uma posição privilegiada e contesta que ele tenha um lugar à parte entre todos os organismos (p. 31).

Como consequência desta “reintegração”, inspirada numa visão científica da condição humana, observa-se o quadro de luta pela existência, marcada por um tênue equilíbrio entre o impulso para a reprodução de um lado, e a escassez de meios de subsistência de outro. Este equilíbrio só é mantido à custa da impossibilidade de elevar os padrões éticos das sociedades. Por isso se afirma acima que o dilema malthusiano conduz à suposição de uma ética “das presas e garras”, e a uma rejeição da possibilidade de aperfeiçoamento e desenvolvimento moral. Em suas palavras:

Tem sido dito que a grande questão está hoje em debate: se doravante o homem se lançará para a frente, com velocidade acelerada, em direção a um aperfeiçoamento ilimitado e até agora inimaginável, ou se será condenado a uma permanente oscilação entre a prosperidade e a miséria e, depois de todo o esforço, ainda permanecerá a uma incomensurável distância do objetivo desejado (Malthus, 1986, p. 279).

Segundo uma concepção mantida na Modernidade por muitos pensadores, seria possível inferir nas sociedades humanas - através da observação e do recurso

a princípios especulativos – um curso de aperfeiçoamento constante e elevação gradativa da felicidade, da liberdade e da virtude na sociedade humana.

A objeção malthusiana ao argumento da perfectibilidade do homem possui dois aspectos. O primeiro é metodológico, e funda-se em uma defesa do método experimental de raciocínio aplicado a assuntos morais, a fim de mostrar que a inferência pela perfectibilidade não está suficientemente justificada. Foi este aspecto que interessou Charles Darwin. A forma do argumento malthusiano - e a decisão metodológica sobre o que deve contar como uma legítima explicação - inspirou o autor da *Origem das Espécies* a aplicar este modelo na consideração do tema da evolução. O resultado foi o desenvolvimento do argumento da “variação cega e retenção seletiva” que Darwin usou para reunir e dar sentido ao enorme conjunto de dados e informações sobre a origem e observável diversidade dos seres vivos.

Agora, se há de fato um dilema, é um dilema entre o que e o que? O princípio da população poderia não constituir dilema, mas ser apenas um obstáculo, um impedimento ao avanço do processo ético. Se o termo empregado no título deste artigo está adequado, quais são as duas alternativas? A primeira é a virtude, que leva ao aumento da população, pela elevação das condições gerais de vida somadas ao impulso natural para a reprodução. A segunda é o vício, que devasta a população e impede o desenvolvimento social e o aperfeiçoamento da humanidade. O dilema se impõe sobre as populações humanas, não importando o seu modelo de sociedade, pois a origem do dilema é a própria condição do ser humano na natureza.

Acerca da atitude de seus interlocutores de inferir um progresso e um aperfeiçoamento nas sociedades humanas, Malthus afirma que:

O atual furor pela especulação ampla e ilimitada parece ser uma espécie de intoxicação mental, decorrente talvez das grandes e inesperadas descobertas que foram feitas nos últimos anos, em todos os ramos da ciência. Para os homens entusiasmados e aturdidos com tais sucessos, todo acontecimento parece estar dentro do domínio dos poderes humanos (Malthus, 1986, p. 322, nota).

Embora, neste sentido, o longo argumento de Malthus represente uma interessante amostra da apropriação dos princípios experimentais derivados das ciências da natureza para o estudo dos assuntos relacionados à conduta e à sociedade humana, não é este aspecto de seu pensamento que se elege aqui como tema principal da discussão.

O segundo aspecto – que é o foco central deste trabalho – oferece uma objeção direta ao conceito de aperfeiçoamento do homem e da sociedade. Segundo Malthus, esta perfectibilidade é inviabilizada pelo fato de que os meios de subsistência crescem em uma razão inferior ao crescimento da população, tendo como conseqüência a queda nos padrões de vida e, portanto, uma queda dos padrões morais segundo os quais os homens orientam suas condutas. Malthus pretendia mostrar que a possibilidade de progresso está submetida, não

exatamente às capacidades morais e intelectuais do ser humano, mas a uma lei da natureza, segundo a qual os seres humanos estariam determinados, por sua condição ambiental, a uma luta pelos meios de subsistência. Ele afirma: “Impedir o retorno da miséria está – infelizmente – além do poder do homem” (Malthus, 1986, p. 303). O progresso econômico e social – e por consequência o progresso moral – estão submetidos a regularidades naturais, tanto no que toca à natureza humana, com seus instintos e tendências herdadas, como no que toca às forças que atuam sobre a população situada em um determinado ambiente.

As leis naturais, para Malthus, determinam o que ocorre com as populações humanas a tal ponto que a única conclusão autorizada por ele é a de que

essa oscilação necessária, essa causa mantenedora constante da miséria cíclica existiu sempre desde que nós tivemos uma história da humanidade; existe no presente e continuará a existir para sempre, a menos que alguma mudança radical ocorra na constituição física de nossa natureza (Malthus, 1986, p. 319).

Isto tudo conduz a reflexão ao tema da liberdade de decisão, em oposição à determinação por causas externas, no tocante ao comportamento humano. Este recurso a evidências de ordem empírica de que Malthus se vale, é um caso da adoção do método experimental, comum entre os pensadores de sua época. Entretanto, mesmo seu emprego do método experimental pode ser questionado, mediante um exame mais acurado. Em seu livro intitulado *A Rational Animal* (1978), Anthony Flew discute alguns aspectos da teoria de Malthus, dentre os quais a sua influência no pensamento de Darwin. Mas a respeito da aplicação do método experimental, interessa notar que Malthus, segundo a análise de Flew “não reconhece nossa fundamental distinção entre os dois sentidos do termo ‘poder’” (Flew, 1978, p. 45). Por que isso importa? Por que um dos significados usuais do termo refere-se a “poder” no “sentido em que a palavra pode ser aplicada a objetos inanimados” (Flew, 1978, p. 45), enquanto no “sentido em que a palavra é tipicamente aplicada a pessoas, e talvez apenas a pessoas, um poder é uma habilidade da vontade de fazer ou abster-se de fazer o que quer que seja” (Flew, 1978, p. 44). Assim, ao referir-se aos poderes que agem sobre a população, Malthus negligencia a capacidade humana de escolha, e por isso, submete toda a esfera da ação livre, de onde a ética final emana, à esfera das regularidades e princípios da natureza física. Reabilitar o sentido e o efeito do poder de escolha humana é eixo em torno do qual se organizam as estratégias dos interessados em objetar as conclusões que conduzem ao dilema de Malthus.

A questão implicada na visão malthusiana é que nenhum princípio, nenhum conjunto de escolhas ou de hábitos poderia interferir na atuação implacável do princípio da população, tal como ele o descreve. Em suas palavras:

Que a população cresce invariavelmente onde existem os meios de subsistência, é provado fartamente pela história de cada povo que tenha existido em qualquer época.

E que o poder superior de crescimento da população não pode ser reprimido sem gerar a miséria ou o vício, é persuasivamente comprovado

pela maior parte dos elementos tão amargos que compõem a sorte da vida humana e a continuidade das causas físicas que parecem tê-los produzido (Malthus, 1986, p. 287).

Não haverá crescimento, nem progresso, nem desenvolvimento no terreno da ética, não haverá aumento da virtude. O vício moral tem como efeito a mortalidade, a despovoação, a interrupção do ciclo reprodutivo em uma determinada população. Vício é aquilo que impede a vida de crescer. E neste sentido, qualquer conduta humana que manifeste este efeito, pode ser considerada um vício. Enquanto o crescimento da população continuar a produzir a miséria e o vício, a ética permanecerá submetida ao princípio da luta pela vida. Esta desoladora conclusão faz da situação das pessoas em sociedade uma figura sombria, impiedosa, onde os valores mais elevados parecem não ser mais do que pálidos anseios, frente à eficácia e à adequação mais imediatas, requeridos de cada um pela imperativa demanda de sobreviver e de se reproduzir.

Há diversos equívocos cometidos por este autor quanto a previsões e cálculos em que pretende apoiar suas teses; equívocos já vastamente apontados e respondidos pela posteridade. Mas, à parte isso, em seu aspecto conceitual, encontramos em Malthus uma concepção fixista da natureza humana, ainda presa a uma noção de essência ou forma imutável que é, ao longo do século XIX, substituída por uma visão mais transformista, caracterizada pela consideração das condições históricas e ambientais com as quais a natureza humana interage. Um dos responsáveis por esta revolução no pensamento teórico, conforme se reconhece, é Charles Darwin. Ao apontar o efeito cumulativo das variações transmitidas e selecionadas por pressões ambientais, Darwin provoca uma mudança fundamental na concepção da relação entre o ser humano e a natureza.

Esta mudança de concepção abre inúmeras questões, entre as quais está a da relação entre o processo de evolução e os conceitos e princípios da ética. O próprio Dewey dedica um artigo a discutir mais pormenorizadamente os efeitos da obra de Darwin sobre o pensamento filosófico, e se expressa acerca disso nos seguintes termos: “A influência de Darwin sobre a filosofia reside em sua conquista do fenômeno da vida para o princípio de transição, e então em libertar a nova lógica para sua aplicação à mente, à moral e à vida” (Dewey, 1910, p. 9). Entre suas conclusões, Dewey enfatiza a continuidade entre o mundo natural e o mundo humano, submetidos ambos ao abrangente princípio de mudança e evolução constante, que Darwin pela primeira vez formulou e fundamentou suficientemente em sua obra.

## **Huxley e a separação entre natureza e ética**

Thomas Huxley, como adepto e defensor das idéias darwinianas, discute os efeitos do modelo de seleção natural para a ética, em um artigo intitulado “Evolução e Ética”. Este artigo resultou de uma conferência apresentada por Huxley na famosa

série de conferências denominada *Romanes Lectures*, e foi publicado originalmente em 1893. Acompanha este texto outro artigo, de 1894, na forma de *Prolegomena* ao artigo inicial. Nestes escritos, ao referir-se ao efeito do crescimento populacional na manutenção do padrão de moralidade, parece manifestar uma compreensão tão malthusiana do princípio da população, como Darwin manifestou ao referir-se ao mundo natural como um ambiente onde está em ação a lei da “sobrevivência do mais adaptado”. Huxley expõe a noção de que um conjunto de leis naturais origina o que ele denomina de “processo cósmico”, enquanto um conjunto de leis morais origina, nas sociedades humanas, o que ele denomina de “processo ético” (Huxley, 2001). Mas, quanto à submissão da moralidade a uma inescapável tendência natural, Huxley argumenta que o processo ético se opõe ao processo cósmico e que quanto mais cultivados os valores éticos, mais eles superam o que Huxley chama de “auto-asserção”, atitude característica dos seres na luta pela vida.

O processo ético substitui a lei natural por um modo de conduta baseado em valores solidários, que tornam o aperfeiçoamento humano possível como uma oposição, como uma superação da tendência natural. Huxley afirma:

O progresso social significa uma restrição do processo cósmico a cada passo e a substituição dele por outro, que pode ser chamado de processo ético; o fim do qual não é a sobrevivência daqueles que pode acontecer serem os mais adaptados a respeito da totalidade das condições que obtêm, mas daqueles que são eticamente os melhores (Huxley, 2001, p. 81).

Assim, quanto mais as relações do homem em sociedade avançam em complexidade, menos o homem está submetido à lei natural da sobrevivência do mais adaptado. Os valores da vida social, para Huxley, distanciam o ser humano da submissão às leis naturais. O propósito humano, a finalidade de seus esforços na direção da virtude ética, parece a Huxley ser justamente este: “A história da civilização detalha os passos pelos quais o homem obteve sucesso em construir um mundo artificial dentro do cosmos” (Huxley, 2001, p. 83). A própria moralidade é vista como uma invenção do engenho humano para resistir ao princípio limitador da natureza, corporificado em uma de suas versões mais populares, na forma do princípio da população. Huxley, desta forma, desacredita das conseqüências sombrias do argumento malthusiano ao afirmar que “o progresso ético da sociedade depende, não de imitar o processo cósmico, muito menos de fugir dele, mas de combatê-lo” (Huxley, 2001, p. 83). Finalmente, Huxley revela-se um adepto da possibilidade de constante aperfeiçoamento da humanidade, aberta aqui pela separação e oposição da atitude ética contra a regularidade natural da luta pela sobrevivência.

Confrontando diretamente qualquer conclusão de que a natureza condiciona, e menos ainda impede o avanço da moralidade, ele afirma ao fim de seu artigo que não vê “limites para a extensão em que a inteligência e a vontade, guiadas por princípios salutareos de investigação, e organizadas num esforço comum, possam modificar as condições de existência” (Huxley, 2001, p. 85). Huxley, a partir do horizonte darwiniano, é capaz de achar uma saída para a lei

limitadora da natureza, que condena a sociedade à miséria e ao vício moral, propondo em seu lugar o processo ético. Mas o faz sob o pressuposto de que a lei moral segue um princípio e uma regularidade contrária à da lei natural da qual, em seu pensamento, resulta o quadro da luta indiscriminada pela existência.

## Dewey e a ética evolutiva

Em seu artigo, também intitulado “Evolução e Ética” e publicado inicialmente em 1898, John Dewey responde a Huxley que o processo ético e o cósmico não são opostos, mas contínuos. Em sua compreensão da herança darwiniana, Dewey entende que os avanços no campo ético são resultantes de uma pressão seletiva do ambiente cultural e social, onde há a reprodução e variabilidade não apenas dos organismos, mas das condutas, valores e crenças humanas. O resultado deste processo é uma modificação, não só dos indivíduos, mas do próprio ambiente social. A conduta moral sofre uma pressão seletiva, que reflete-se em modificações na condição ambiental. A posição de Dewey acerca da ética configura-se - segundo se supõe - uma estratégia distinta de resposta à desoladora conclusão de Malthus acerca da impossibilidade de contínuo aperfeiçoamento moral e social da humanidade.

Em um trabalho intitulado “*Evolution and Ethics: the Huxley/Dewey Exchange*”, John Teehan comenta: “É curioso notar que a despeito da eminência de Dewey na filosofia do século XX, e de sua adoção do pensamento evolutivo, ele tem sido ignorado no debate corrente sobre evolução e ética” (Teehan, 2002, p. 226). A fim de dar um adequado tratamento a esta parte do pensamento de Dewey, é preciso fazer uma consideração do lugar ocupado por estas reflexões no conjunto de seu pensamento e no debate acerca da relação entre evolução e ética. No artigo de Dewey, encontramos menção direta ao pensamento de Malthus na seguinte passagem: “Seria ingrato em qualquer discussão deste assunto não referir-se à clássica ilustração de Malthus do festim espalhado pela natureza - que não é grande o suficiente para os convidados” (p. 51). Se o filósofo de inclinação naturalista pretende manter a integridade intelectual de sua abordagem sem remeter-se a um processo distinto ou contrário ao processo natural, tem que posicionar-se diante das implicações éticas que têm sido extraídas de concepções como a de Malthus, quanto à relação entre princípios naturais e princípios morais.

Enquanto Huxley entende a moralidade como uma dimensão distinta, e mesmo oposta aos processos da natureza, Dewey prefere conceber a moralidade como um processo evolutivo, de complexidade crescente, que se desenvolve na relação dos indivíduos com as condições existentes no ambiente social. Sem referir-se especificamente ao princípio da população, estabelece um ponto de vista que será mantido ao longo de toda a sua longa carreira filosófica. Quase três décadas depois da redação deste artigo, Dewey continua a manter um ponto de vista naturalista evolutivo ao referir-se aos temas da ética.

Na introdução de seu livro de 1922, *Natureza Humana e Conduta*, ele afirma: “Estas páginas são a discussão de algumas fases da mudança ética envolvida no respeito positivo pela natureza humana quando a última é associada ao conhecimento científico” (Dewey, 2002, p. 4). Para ele, é imprescindível estudar o processo evolutivo dos valores, como acompanhamento adaptativo da mudança dos ambientes sociais em que estes valores são incorporados. Num universo em constante mudança, os agentes sempre estão submetidos à necessidade de adaptar-se às condições ambientais existentes. A previsão e a intercomunicação pelas quais os interesses socialmente compartilhados se tornam comuns deixam margem para a escolha, que reveste os valores de significado. Esta margem de escolha abre espaço para a mudança na conduta ética. O crescimento adaptativo dos valores, assim ocasionado, acompanha o crescimento dos ambientes sociais, permitindo a contínua reconstrução das condutas humanas.

Segundo Teehan, mencionado acima: “Se formos consistentemente naturalistas, temos que aceitar que as condições da sociedade são partes das condições ambientais tanto quanto as rochas, o solo e a água” (Teehan, 2002, p. 232). Manter a unidade entre o processo ético e o processo cósmico não é somente uma abordagem metodológica e conceitual da ética. É antes disso a manifestação do princípio unificador de toda a filosofia de Dewey: o princípio de continuidade.

A partir de toda a discussão que foi conduzida até aqui, pode-se extrair a seguinte apreciação do assunto: Segundo o parecer de Malthus, a lei natural submete a lei moral, de modo irrevogável. Esta lei natural produz como resultado a miséria e o vício, por causa da imposição de uma luta pela sobrevivência. Para poder reagir a esta concepção, Huxley propõe que a lei moral se opõe à lei natural. Conforme seu modo de ver, o processo ético substitui o processo cósmico de luta pela vida. O processo ético substitui, num trajeto acumulativo de avanço da civilização, a luta pela vida pela luta pela felicidade por meio da virtude moral. A lei natural deixaria de ter jurisdição nas sociedades em que o processo ético atinge um grau avançado e estável de desenvolvimento. A ética, estabelecida pelo propósito e pela inteligência humana, invalida o princípio malthusiano de invariável retorno da sociedade à miséria e ao vício moral.

Dewey não concorda com Huxley sobre a oposição entre o processo ético e o cósmico. Para ele a lei natural e a lei moral são contínuas, são expressões da mesma grande regularidade observável e reconhecível por seus efeitos, e tornada acessível ao investigador pela aplicação do método científico. Na dimensão moral, os hábitos são selecionados por seu sucesso adaptativo. Mais relevante ainda, o que Malthus não foi capaz de ver é que os ambientes sociais estão sujeitos à mudança e ao crescimento de complexidade. “O ambiente é agora distintamente um ambiente social, e o conteúdo do termo ‘adaptado’ deve ser construído com referência à adaptação social” (Dewey, 1972, p. 41). A seleção de hábitos e de condutas toma o lugar da seleção dos organismos individuais, que são seus portadores. Na reconstrução do sentido do termo “seleção” que Dewey propõe, cabe aplicar o mesmo mecanismo também aos produtos dos organismos, contados como estratégias adaptativas, tais como comportamentos, crenças, hábitos. E, em um ambiente mais complexo aos valores, conceitos e significados.

A posição de Dewey é bastante ousada, e envolve a adoção de um princípio de mudança e crescimento constante, que é distinto - mais que isso, que é contrário - à cosmologia das leis e dos estados fixos, que ainda estava subjacente no pensamento teórico da época de Malthus. Nos termos em que Dewey apresenta sua consideração, fica clara sua adesão a uma grande lei da contínua mudança, e da adaptação que os organismos buscam atingir em resposta a ela, na forma do crescimento de complexidade, por meio da seleção cumulativa de características:

Não há apenas a triagem pela morte, há a triagem pelo sucesso ou falha de atos especiais - a contraparte, eu suponho, da assim chamada seleção fisiológica. Nós não precisamos entrar na aborrecida questão da herança de caracteres adquiridos. Nós sabemos que através do que chamamos de opinião pública e educação certas formas de ação são constantemente estimuladas e encorajadas, enquanto outros tipos são constantemente objetados, reprimidos e punidos. Que diferença existe em princípio entre esta mediação dos atos do indivíduo pela sociedade e o que é ordinariamente chamado de seleção natural, eu sou incapaz de ver (Dewey, 1972, p. 51).

Acerca do mecanismo de variação e seleção adaptativa atuar sobre replicadores culturais, tais como palavras, crenças, hábitos de conduta, o trabalho de Dawkins (2007) e Dennett (1998) têm dado origem, nos últimos anos, a um intenso e interessante debate. O que parece é que Dewey faz algumas insinuações nesta direção, quando defende que a luta pela sobrevivência é travada entre artefatos, entre projetos culturalmente e socialmente reproduzidos, de modos de agir e de significar.

## **Considerações Finais**

O assunto, finalmente, poderia ser posto sob a forma assumida pela seguinte pergunta: Onde termina a natureza e começa a moralidade? Uma versão mais ampla desta questão vem sendo alvo de grandes discussões acerca dos fundamentos das ciências humanas, quando se pergunta o que se deve atribuir à natureza e o que se deve atribuir à cultura no caso do ser humano e de suas produções. Claro que os termos da pergunta estão postos de forma pouco rigorosa, mas no mínimo ilustram satisfatoriamente a dificuldade.

Malthus considera a moralidade como subjugada à lei natural. Por isso, em sua forma de pôr o dilema, o ser humano e suas produções não podem ascender a um grau mais elevado de sofisticação, do ponto de vista dos valores e princípios da ética, pois na prática, na conduta efetivamente manifesta pelos indivíduos e instituições, está em operação o princípio inescapável da população, que mantém os padrões da moralidade próximos demais da miséria e do vício.

Huxley observa que o aparecimento dos processos éticos tem o efeito de impedir o quadro de luta indiscriminada pela existência, por meio da

consideração do bem estar alheio, por meio do cultivo de valores e hábitos morais. Mas este processo ético é um rompimento, uma oposição ao curso das regularidades naturais. Onde ele está em operação, a natureza e suas regularidades, consubstanciadas no modelo selecionista, não têm ocasião de operar.

Dewey, como se tem tentado afirmar, desenvolve uma concepção ainda mais refinada desta situação. Afirma que a própria moralidade vem gradativamente emergindo a partir de instintos e impulsos que são modificados pela pressão seletiva, exercida por ambientes gradativamente modificados. O destino da sociedade em geral, e em particular da ética não pode ser previsto, como pretendiam os defensores da perfectibilidade humana criticados por Malthus. Contudo, ao propor o método evolutivo como instrumento de estudo da moralidade, Dewey faz uma indicação importante: Por meio deste método é possível reunir o aspecto ideal e o aspecto material dos valores, a fim de desenvolver comportamentos que possam levar a sociedade a um maior sucesso adaptativo, na manutenção e no aprimoramento de seu horizonte moral. Efeitos da aplicação deste método observam-se fartamente na obra posterior de Dewey sobre a Educação, por exemplo. Deste modo, a própria ética é uma estratégia evolutiva. Sua origem mais remota são os traços comportamentais instintivos que tornaram possível a sociabilidade de nossos ancestrais. Seu progresso recente é efeito da seleção cumulativa de variações transmitidas pela comunicação, pela educação, sob pressão seletiva do ambiente social moderno.

A antecipação, a previsão inteligente das conseqüências, resultante do pensamento e da comunicação, representa um obstáculo à ação perniciosa da miséria e do vício, por que leva a uma conduta baseada na experimentação e aperfeiçoamento dos interesses socialmente compartilhados. Pode-se encerrar esta reflexão com a conclusão de que o darwinismo em Dewey – sua adoção de um princípio selecionista, evolutivo, naturalista – é mais amplo e mais profundo do que a versão mantida por Huxley, anunciando caminhos aos pensadores darwinistas da segunda metade do século XX. Além disso, sua postura leva a explicar de modo mais compreensível como a cultura que é reproduzida na vida social pode dar origem a um ambiente adaptativo que cria e renova o mundo dos valores, segundo os quais se orienta a conduta humana.

## Referências

DAWKINS, Richard. (1976). *O Gene Egoísta*. São Paulo: Cia das Letras. 2007.

DENNETT, Daniel. (1995). *A Perigosa Idéia de Darwin*. Rio de Janeiro: Rocco. 1998.

DEWEY, John. (1898). "Evolution and Ethics". In: *The Early Works of John Dewey 1882-1898, Vol. 5 (1895-1898)*. Jo An Boydston (Ed.). Carbondale & Edwardsville: Southern Illinois University Press. 1972.

- \_\_\_\_\_. The Influence of Darwinism on Philosophy. In: *The Influence of Darwin on Philosophy, and Other Essays in Contemporary Thought*. New York: Holt, 1910; London: Bell, 1910.
- \_\_\_\_\_. (1922). *Human Nature and Conduct*. New York: Prometheus Books. 2002.
- FLEW, Anthony. *A Rational Animal – and other philosophical essays on the nature of man*. Oxford: Clarendon Press. 1978.
- HUXLEY, Thomas Henry. “Evolution and Ethics” (1893). In: *Collected Essays. Vol. 9. Evolution and Ethics and other Essays*. London: Elibron Classics. 2001.
- MALTHUS, Thomas. *Ensaio Sobre a População* (1798). São Paulo: Nova cultural. 1986.
- POURSIN, Jean-Marie & DUPUY, Gabriel (1972). *Malthus*. São Paulo: Cultrix & Editora da USP. 1975.
- SPALDING, H. S. “Ethics and the Neo-Malthusianism”. In: *The American Journal of Sociology*. The University of Chicago Press. Vol. 22, n. 5 (pp. 609-615). 1917.
- TEEHAN, John. “Evolution and Ethics: The Huxley/Dewey exchange”. In: *The journal of speculative philosophy*. Vol. 16, No. 3. p. 225-238. 2002.

# O SUPREMO TRIBUNAL FEDERAL E A ANENCEFALIA: UMA REFLEXÃO SOBRE A LEGITIMIDADE DEMOCRÁTICA DO JUDICIÁRIO À LUZ DE RAWLS, HABERMAS E NINO

MARIA EUGENIA BUNCHAFT<sup>1</sup>

*Unifoa – Centro Universitário de Volta Redonda*

mbunchaft@ig.com.br

**Resumo:** O debate entre Habermas e Rawls representa uma contribuição fundamental para a compreensão das questões sobre reconhecimento, multiculturalismo e do pós-secularismo, introduzindo diferentes concepções filosóficas que podem contribuir sobre a temática acerca do uso público da razão, a fim de elucidar as diferentes percepções teóricas capazes de atender aos desafios propostos pelas sociedades pluralistas. Carlos Santiago Nino, por sua vez, estabeleceu uma estratégia teórica denominada “construtivismo epistemológico”, delineada a partir do debate Habermas-Rawls. Por conseguinte, pretendemos articular os fundamentos filosóficos atinentes à concepção de razão pública delineada por Rawls e Habermas e ao construtivismo epistemológico de Carlos Santiago Nino com a análise da legitimidade democrática do STF na resolução de uma questão moral controversa que é objeto de um desacordo moral razoável: o aborto de fetos anencéfalos.

**Palavras-chave:** Habermas; Rawls; Nino; razão pública; ativismo judicial.

## 1. Introdução

O debate entre Habermas e Rawls representa uma contribuição fundamental para a compreensão do tema do reconhecimento, do multiculturalismo e do pós-secularismo, introduzindo diferentes concepções filosóficas que podem enriquecer em muito a temática acerca do uso público da razão, a fim de elucidar as diferentes percepções teóricas capazes de atender aos desafios propostos pelas sociedades pluralistas. Carlos Santiago Nino, por sua vez, estabeleceu uma estratégia teórica denominada “Construtivismo epistemológico”, delineada a partir do debate Habermas-Rawls. Por conseguinte, pretendemos articular os fundamentos filosóficos atinentes à concepção de razão pública delineada por Rawls e Habermas e ao construtivismo epistemológico de Carlos Santiago Nino com a análise da legitimidade democrática do STF na resolução de uma questão moral controversa que é objeto de um desacordo moral razoável: o aborto de fetos anencéfalos.

Nesse sentido, a possibilidade de articulação dos fundamentos filosóficos da teoria de Rawls, Nino e Habermas, assim como a integração do conceito de razão pública à cultura política brasileira, especialmente no que se refere à análise das consequências para a legitimidade democrática do STF, são fundamentais para o debate sobre a judicialização em temas delicados como anencefalia, transexualismo ou uniões homoafetivas. Para Rawls, a atuação do poder político em relação a questões constitucionais essenciais e de justiça básica somente ocorre por meio da fundamentação das decisões com base em princípios políticos que invocam os valores da razão pública. Nas palavras de Rawls, “nosso exercício do poder político é próprio e, por isso, justificável somente quando é exercido de acordo com uma Constituição cujos elementos essenciais se pode razoavelmente esperar que todos os cidadãos endossem, à luz de princípios e ideais aceitáveis para eles, enquanto razoáveis e racionais.”<sup>2</sup> Assim, Rawls desenvolve uma abordagem peculiar acerca do exercício do poder político, cuja legitimidade pressupõe o dever moral de justificação das decisões sobre questões constitucionais essenciais e de justiça básica à luz dos valores políticos da razão pública.

Mas, a questão fundamental é: quais as implicações democráticas, sob o ponto de vista sociológico, da integração das concepções de razão pública delineadas por Habermas e Rawls ou do construtivismo epistemológico desenvolvido por Nino em relação a questões constitucionais controvertidas a serem julgadas pelo Judiciário tais como o aborto de fetos anencéfalos? O presente trabalho pretende investigar em que medida os referenciais teóricos dos referidos autores atendem aos desafios propostos por uma estratégia de legitimação do Supremo Tribunal Federal na decisão sobre a possibilidade de descriminalização da antecipação terapêutica da gestação de fetos anencéfalos ventilada na Arguição de Descumprimento de Preceito Fundamental n. 54/DF.

## **2. O debate entre Habermas e Rawls**

De início, é premente lecionar que a tradição liberal enfatiza uma dimensão deontológica que estabelece princípios neutros de justiça como pressuposto para que os indivíduos escolham suas concepções de vida boa. Os autores comunitaristas, por sua vez, partem de uma crítica a filosofias universalistas, que estabelecem a prioridade do justo sobre aspectos relativos ao bem, ocultando singularidades culturais específicas em uma perspectiva homogeneizante. Tanto Habermas como Rawls rompem com a tradição comunitarista quando sustentam a inviabilidade de se adotar uma concepção de justiça comprometida com objetivos coletivos substantivos.

Diante dessa estrutura conceitual, o embate teórico entre Habermas e Rawls é um dos temas mais relevantes da filosofia política contemporânea, contemplando uma crítica de Habermas, uma réplica do segundo e um novo

ensaio crítico do filósofo alemão. O debate surge em um número da revista *The Journal of Philosophy*, tendo como ponto de partida o ensaio crítico de Habermas, “Reconciliation through the public use of reason: remarks on John Rawls’s Political Liberalism”. Em verdade, tanto Habermas quanto Rawls distanciam-se de concepções comunitárias, pois delineiam uma concepção de justiça desvinculada de uma perspectiva apoiada em objetivos etnoculturais.

Nesse ponto, Habermas configura uma teoria da justiça que tem como pressuposto o estabelecimento de procedimentos por meio dos quais as reivindicações identitárias são tematizadas no espaço público, de forma a atender às demandas de uma sociedade plural. Diante dessa estrutura filosófica, a prioridade do justo sobre o bem traduz-se em uma concepção procedimental na qual o sistema de direitos institucionaliza as condições que garantem as autonomias pública e privada. Tal concepção é extremamente sofisticada, pois o autor procura desenvolver um ideal universalista com uma sensibilidade inclusiva em relação às diferenças culturais, uma vez que os princípios constitucionais passam a ser interpretados à luz de singularidades culturais específicas.

Nesse quadro teórico, Rawls, em *O Liberalismo Político*, estabelece um instrumental teórico destinado a analisar o fato do pluralismo razoável, ou seja, a existência de uma pluralidade de doutrinas abrangentes opostas e incompatíveis, porém razoáveis. Rawls delineia uma concepção política de justiça, pressupondo que os cidadãos que integram uma sociedade bem-ordenada, embora possuam concepções morais e filosóficas, são caracterizados pelos atributos da razoabilidade e da racionalidade. A ideia de posição original, delineada em *Teoria da Justiça*, é vinculada a uma filosofia contratualista desenvolvida com base em um acordo hipotético em um contexto não histórico por meio do qual os cidadãos estabelecem uma concepção política de justiça, sendo desprovidos de suas concepções particulares de mundo. Os indivíduos não apenas ignoram o seu lugar no mundo, como também as contingências históricas e sociais.

Outrossim, com o intuito de estabelecer uma concepção política de justiça que irá pautar-se por princípios endossados por cidadãos razoáveis, Rawls recorre, em um segundo momento, ao conceito de consenso sobreposto. O consenso sobreposto irá inspirar a adesão por cidadãos racionais e razoáveis que, a despeito de suas doutrinas abrangentes, recorrem ao ideal da razão pública. A convergência de doutrinas abrangentes, porém razoáveis, é assegurada por meio do consenso sobreposto inerente a uma sociedade liberal. Por meio do consenso sobreposto, os cidadãos razoáveis endossam uma concepção política de justiça, recorrendo ao ideal da razão pública, que reflete valores de uma cultura política democrática. Essa concepção de justiça, que configura a melhor forma de concepção política, satisfaz os interesses de cidadãos livres e iguais.

Como se sabe, a ideia de consenso sobreposto pressupõe que as concepções políticas inerentes a doutrinas abrangentes deverão atender ao requisito da razoabilidade, articulando-se à concepção política de justiça. Nas palavras de Rawls, “as pessoas são razoáveis em um aspecto básico quando, entre iguais, por exemplo, estão dispostas a propor princípios e critérios como termos equitativos

de cooperação e a submeter-se voluntariamente a eles, dada a garantia de que os outros farão o mesmo.”<sup>3</sup> É mister elucidar que, para Rawls, a ideia da razão pública, portanto, funciona como um substrato legitimador das discussões fundamentais por meio das quais os cidadãos irão delinear uma concepção política de justiça.

Com o intuito de elucidar a concepção de razão pública, assume especial relevância a distinção estabelecida por Rawls entre o ponto de vista público de perspectivas não públicas. De início, o autor pressupõe a existência de uma única razão pública e muitas razões não públicas, incluindo as autocompreensões morais das igrejas, universidades, sociedades científicas. Estas integram o horizonte de uma cultura de fundo de uma sociedade, incluindo as razões da sociedade civil, em contraposição à cultura política pública. Desse modo, os cidadãos poderiam fazer uso exclusivo de suas doutrinas abrangentes religiosas quando participam da “cultura de fundo”, mas, devem submeter-se aos limites da razão pública quando participam da defesa política do espaço público. A discussão pública sobre questões políticas fundamentais deve atender ao requisito da razoabilidade, sendo esta necessária para a cooperação democrática em uma sociedade pluralista.

Com efeito, para Rawls, o fundamento de legitimidade das decisões sobre questões políticas fundamentais deve basear-se em razões que sejam publicamente aceitas por todos os cidadãos razoáveis. Rawls distingue a forma pela qual o ideal da razão pública aplica-se aos cidadãos e às autoridades estatais. Nas palavras de Rawls, “esse ideal aplica-se aos fóruns oficiais e, por isso, aos legisladores, quando falam no recinto do parlamento, e ao executivo, em seus atos e pronunciamentos públicos.”<sup>4</sup> E conclui: “aplica-se também, de uma forma especial, ao judiciário e, sobretudo, ao Supremo Tribunal numa democracia constitucional com controle da constitucionalidade das leis (revisão judicial)”.<sup>5</sup>

De fato, para o autor, o Tribunal Constitucional é um caso exemplar de razão pública, uma vez que os juízes devem explicar e justificar suas decisões de acordo com a melhor compreensão da Constituição, dos estatutos e precedentes. Para Rawls, “ao aplicar a razão pública, o tribunal deve evitar que a lei seja corroída pela legislação de maiorias transitórias ou, mas provavelmente, por interesses estreitos, organizados e bem-posicionados, muito hábeis na obtenção do que querem.”<sup>6</sup> A Suprema Corte deve sempre satisfazer ao ideal da razão pública, mesmo quando não estejam em jogo elementos constitucionais essenciais ou questões de justiça básica, diferentemente dos cidadãos e legisladores, que podem votar de acordo com suas doutrinas abrangentes nas questões que não envolvam os referidos elementos.

É por meio de uma dimensão restritiva imposta pela ideia de razão pública, que Rawls, no ensaio “The Idea of Public Reason Revisited”<sup>7</sup>, irá justificar o requisito da tradutibilidade, a fim de ajustar o exercício da manifestação das doutrinas religiosas na esfera pública a uma linguagem inerente aos pressupostos de valores políticos fundamentais. Ou seja, os cidadãos, ao participarem do fórum político público, com o intuito de deliberarem sobre questões constitucionais

essenciais e de justiça básica, não possuem a obrigação de se desvincularem de doutrinas seculares ou religiosas, sendo possível a manifestação de doutrinas abrangentes na esfera pública. Não obstante, a ideia de razão pública, com o objetivo de administrar a pluralidade de doutrinas abrangentes que conformam uma sociedade pluralista, assume uma dimensão restritiva por meio da exigência da tradutibilidade, ou seja, requer a satisfação do pressuposto de traduzir argumentos não-públicos para uma linguagem que seja publicamente aceitável consoante valores políticos.

Não obstante, Habermas leciona que a teoria rawlsiana de justiça revela algumas insuficiências teóricas. Uma das críticas do filósofo alemão centra-se em torno da ideia de consenso sobreposto, uma vez que, para Rawls, este será o instrumento capaz de viabilizar a estabilidade social, tendo em vista a ideia de razoabilidade. Para Habermas, a teoria do autor americano pressupõe erroneamente que o teste da aceitabilidade de seus princípios conduziria à auto-estabilização da sociedade. O filósofo argumenta que a teoria rawlsiana apenas poderia ser utilizada como um mecanismo de aceitação, mas não de aceitabilidade. Habermas estabelece uma diferenciação entre procedimentos democráticos que garantem a estabilidade da ordem racional, tendo em vista a dimensão da aceitabilidade, das questões relativas à aceitação. O autor alemão alega ainda que a posição original contempla uma dimensão substantiva. Nessa perspectiva, se os interesses das partes são compreendidos como bens, ainda que estes se rearticulem e passem a ser compreendidos na forma de direitos, isso não anula o conteúdo substantivo dos princípios de justiça. Uma vez que os direitos básicos estão equiparados a bens, trata-se de uma ética substancialista que minimiza a sua pretensão deontológica. Ademais, Habermas critica a ideia de posição original de acordo com a qual, as partes desprovidas de suas contingências e particularidades estabelecem princípios de justiça, pois a ideia de privação de informações enfraquece o potencial de imparcialidade do procedimento de razão prática, impedindo processos de aprendizagem.

Em face dessa leitura, nas palavras de Habermas, “pela limitação da informação, Rawls impõe uma perspectiva comum aos participantes da posição original e neutraliza assim de antemão, mediante um artifício, a multiplicidade das perspectivas particulares de interpretação.”<sup>8</sup> E conclui: “a ética do discurso, pelo contrário, vê o ponto de vista moral como encarnado no procedimento de uma argumentação levada a efeito intersubjetivamente, que exorta os participantes a erguerem as barreiras de suas perspectivas de interpretação.”<sup>9</sup> O autor recorre aos postulados fundamentais da ética discursiva para contrapor seu modelo de comunidade ideal de fala à ideia de posição original.

Em suma, a posição original não seria suficientemente deliberativa, inviabilizando que os indivíduos, por meio de processos de aprendizagem, transcendam a si mesmos a partir de suas próprias concepções de bem. De um lado, a posição original não satisfaz plenamente a imparcialidade exigida pelos pressupostos deontológicos. De outro lado, ao atribuir primazia aos direitos liberais, em contraposição ao princípio democrático, Rawls não atinge seu

objetivo de conciliar a liberdade dos modernos com a dos antigos, uma vez que sua teoria subordina o segundo princípio de justiça (igualdade) ao primeiro (liberdade). Habermas também demonstra-se cético em relação à renúncia rawlsiana à pretensão de verdade, uma vez que a sua própria concepção de razoabilidade é insuscetível de optar teoricamente entre aceitabilidade racional, visando a justiça e aceitação social voltada para a estabilidade. Nesse sentido, a concepção rawlsiana de razão pública não trabalha com os conceitos de verdade, pois resgata a ideia do politicamente razoável, tendo em vista o dever de civilidade e a ideia de cidadania democrática, pressupondo os princípios da tolerância e da liberdade de consciência.

No ensejo, Rawls inicia a sua réplica estabelecendo as diferenciações entre a sua teoria e a ética discursiva, tendo em vista os diferentes mecanismos de representação: a situação ideal de fala e a posição original.<sup>10</sup> Aduz que a teoria da ação comunicativa, ao basear-se na ideia de situação ideal de fala, pretende dar conta da construção intersubjetiva da verdade e validade dos juízos morais. Nesse aspecto, Rawls argumenta que a ética discursiva e sua teoria da ação comunicativa pressupõem uma doutrina compreensiva, questionando o seu caráter procedimental. Diferentemente, a justiça como equidade, ao desvincular-se de doutrinas compreensivas, mantém intacta a sua dimensão procedimental e a prioridade do justo sobre o bem. O autor recorre a um equilíbrio reflexivo amplo por meio do qual a concepção de justiça, delineada na posição original, adequa-se às convicções e aos juízos mais profundos. Trata-se de um processo de ajustamento mútuo no qual os juízos e as convicções particulares ajustam-se à concepção de justiça e esta aos primeiros. De fato, a posição original representa um artifício por meio do qual as partes são desvinculadas de suas “doutrinas compreensivas” e convicções particulares de natureza moral, religiosa ou filosófica. Em Habermas, não é necessário esse artifício, porquanto a própria estrutura linguística permite um procedimento argumentativo sobre as concepções de bem.

Sob essa ótica, Rawls contra-argumenta no sentido de que o seu procedimentalismo não se confunde com aquele delineado por Habermas, tendo em vista sua distinção entre questões de justificação e de aplicabilidade. Uma justificação pública inerente à concepção política de justiça é sedimentada na multiplicidade de doutrinas abrangentes razoáveis, sendo intrínseca à cultura política das democracias contemporâneas. Todas aquelas doutrinas abrangentes que não satisfazem ao requisito da razoabilidade são insuscetíveis de formar a concepção política de justiça. Nesse sentido, sua concepção de estabilidade inerente ao consenso sobreposto seria um fator de justificação de concepções políticas e não apenas de equilíbrio entre doutrinas abrangentes conflitantes. Ademais, nega que sua teoria estabeleça uma concorrência entre as autonomias pública e privada, sustentando que a justiça como equidade, se corretamente interpretada, respeita a relevância e equilíbrio entre ambas as autonomias .

Diante dessa estrutura conceitual, outro ponto polêmico diz respeito à ideia de tradução. Nesse ponto, para Rawls, enquanto o requisito da tradução aplica-

se de forma ampla em relação em relação à esfera pública formal, pode ser minimizado na esfera pública informal, com o intuito de permitir as manifestações religiosas. As doutrinas religiosas podem ser introduzidas na esfera pública informal, desde que atendam ao requisito da razoabilidade e da tradutibilidade. Entretanto, surge a questão: como encontrar justificativas seculares independentes de crenças particulares sem comprometer a integridade da identidade dos cidadãos enquanto crentes religiosos?

Assim, Habermas, no ensaio “Religião na Esfera Pública - Pressuposições Cognitivas para o Uso Público da Razão de Cidadãos Seculares e Religiosos”, pondera que a ideia de razão pública delineada por Rawls é suscetível a críticas à medida que impõe uma sobrecarga cognitiva sobre cidadãos religiosos, porquanto muitos destes não poderiam empreender uma divisão artificial dentro de suas próprias mentes. Habermas assinala que a concepção rawlsiana de razão pública parte de uma perspectiva psicológica que, se não impede a introdução de doutrinas religiosas na esfera pública, sujeita-se a críticas relativas à exigência da tradutibilidade, que impõe aos cidadãos a tarefa de, quando se manifestarem no espaço público informal, empreenderem uma divisão artificial em suas mentes entre valores religiosos e políticos. Nas palavras de Habermas, “muitos cidadãos que se posicionam quanto a questões políticas, assumindo uma perspectiva religiosa, não têm nem conhecimentos suficientes para encontrar fundamentações seculares independentes de suas convicções autênticas”. E conclui: “a pessoa piedosa encara sua existência a partir da fé. E fé verdadeira não é apenas doutrina, conteúdo no qual se crê, mas também fonte de energia da qual se alimenta a vida inteira do crente”.<sup>11</sup> Disso se infere, a nosso ver, que a conexão entre identidade religiosa e processos de aprendizagem, na ética discursiva, configura uma forma de preencher a lacuna teórica decorrente do modelo de razão pública delineado por John Rawls.

Nesse sentido, há uma tensão entre o reconhecimento da inclusão dos discursos religiosos na esfera pública e a exigência de que esta seja regulada pela razão natural. Habermas utiliza-se de duas estratégias para resolver esta tensão. De um lado, através da separação institucional entre religião e Estado; de outro lado, através da distinção entre esfera pública informal e a estrutura institucional do Estado, que inclui os parlamentos, ministérios e o Judiciário. Nesse ponto, a introdução de argumentos religiosos na esfera pública informal pressupõe processos de aprendizagem por meio dos quais tanto cidadãos religiosos como cidadãos seculares devem se engajar em uma autorreflexão crítica sobre suas crenças e concepções de bem. Tais processos de aprendizagem, inerentes a uma moral pós-convencional, permitem aos indivíduos confrontarem as autocompreensões de outros cidadãos com suas próprias doutrinas religiosas particulares, viabilizando a tradução de argumentos religiosos para argumentos publicamente aceitáveis na esfera pública, tendo em vista *insights* críticos e reflexivos.

Outrossim, não é possível saber de antemão se a proibição de manifestações religiosas na esfera pública informal estaria privando a sociedade de “recursos

importantes para a criação de sentido”, ou seja, de intuições morais possivelmente relevantes. E o trabalho de tradução pressupõe um esforço cooperativo do qual participam crentes e cidadãos não-religiosos, buscando-se um ponto de vista moral imparcial por meio da assunção recíproca de perspectivas. Portanto, o cidadão não religioso deve revelar-se aberto aos processos de aprendizagem estabelecidos por meio do diálogo com o concidadão religioso. No entanto, consideramos que tal alegação habermasiana, ao aplicar-se ao espaço público brasileiro, deve ser interpretada com uma certa cautela, no que se refere à discussão de direitos fundamentais de grupos minoritários que envolvam aspectos exclusivamente privados, ou seja, questões de moralidade autorreferente, tal como descreve Carlos Santiago Nino. Nesse quadro teórico, passamos a examinar o construtivismo epistemológico de Nino.

### **3. O Construtivismo Epistemológico de Carlos Santiago Nino**

De início, é premente lecionar que o construtivismo epistemológico, na obra de Nino, situa-se como uma posição intermediária entre as perspectivas de Rawls e Habermas, afirmando a centralidade de princípios morais, compreendidos como o substrato legitimador dos direitos humanos. Nino procura dialogar com ambas as vertentes filosóficas, ressaltando o caráter dialógico na formulação dos juízos morais, mas legitimando também a reflexão individual. Assim, em passagem elucidativa, sublinha que:

(...)...quanto ao conhecimento da verdade moral, Rawls parece assumir que só é acessível através da reflexão individual, enquanto Habermas sustenta que somente a discussão coletiva é um método adequado para acessar princípios morais válidos. Acredito que a posição correta sobre ambas as questões é intermediária entre sustentadas por ambos os filósofos.  
(...) <sup>12</sup>

Com efeito, tanto Habermas como Nino ressaltam valor epistêmico da deliberação democrática para a resolução de desacordos morais, uma vez respeitadas as condições procedimentais do processo discursivo, suscitando um diálogo aberto e inclusivo no qual os interessados podem livremente confrontar seus argumentos. Resgatando ambas as matrizes epistemológicas, Nino leciona que, não obstante o processo deliberativo seja o método mais confiável para atingir princípios morais, tal compreensão não anula o fato de que a reflexão individual pode igualmente se configurar como um meio válido para alcançar princípios moralmente corretos sempre que a deliberação estiver desprovida de valor epistêmico. Nesse sentido, para Nino, em tese, o processo democrático, por seu valor epistêmico intrínseco, é o método mais confiável para se atingir princípios morais, em contraposição à reflexão individual monológica. Nas palavras de Nino, “Isto, sem embargo, não exclui a possibilidade de que, por meio da reflexão individual, se possa ter acesso ao conhecimento de soluções corretas, ainda que

se admita ser este método muito menos confiável que o coletivo, devido à dificuldade de permanecer fiel à representação dos interesses dos outros e ser imparcial. “<sup>13</sup>

É mister elucidar que, para o autor argentino, quando a deliberação estiver desprovida de valor epistêmico, a reflexão individual revelar-se-ia como um meio válido para resolução de questões morais controvertidas. Habermas, por sua vez, aposta no potencial racionalizador do debate a partir das contribuições dos próprios afetados, que tornam-se autores e destinatários dos seus direitos, em contraposição a uma produção monológica de normas. Mas, em países periféricos de modernidade tardia, não podemos ignorar que, muitas vezes, há um déficit de representação de grupos minoritários no processo político majoritário, o que legitima a atuação do Poder Judiciário - identificado pelo autor como a expressão da reflexão individual - quando o processo deliberativo estiver desprovido de condições que configurem seu valor epistêmico.

Nesse particular, para Nino, as três dimensões fundamentais do constitucionalismo são: Democracia, reconhecimento de direitos e Constituição Histórica. Se o aspecto real do constitucionalismo constitui a Constituição Histórica, a Constituição ideal expressa-se nas ideias de Democracia e na dimensão dos direitos. Ademais, a legitimidade da “Constituição Histórica” tem como conteúdo o reconhecimento de princípios morais. Desse modo, o valor da autonomia moral é o elemento fundamental de qualquer processo discursivo e, com base nesta ideia, o autor refuta “qualquer política que imponha ideais pessoais sobre os indivíduos.”<sup>14</sup> Nino elucidada que, em contraposição ao modelo liberal, “a visão perfeccionista recupera a adoção de um modelo de virtude pessoal definido por ideais cívicos e comunitários.”<sup>15</sup>

Nessa perspectiva, o autor estabelece três princípios básicos de uma concepção liberal de sociedade: o “princípio da inviolabilidade da pessoa humana”, o “princípio da autonomia da pessoa humana” e o “princípio da dignidade da pessoa humana”. O princípio da autonomia da pessoa relaciona-se ao valor da autonomia moral, sendo inadmissível qualquer forma de interferência externa sobre a escolha de ideais de excelência pessoal e virtude. O Estado deve manter-se neutro em relação aos projetos individuais de vida e ideais de excelência humana, contrapondo-se ao perfeccionismo.

Como se sabe, para o republicanismo, uma vez estimulada a participação política dos cidadãos, estes serão menos propensos a perseguir interesses privados, conectando democracia e o exercício das virtudes cívicas. Na perspectiva republicana, “o princípio da autonomia pessoal seria honrado se certas virtudes de caráter forem promovidas como um fim em si mesmas e não por seus méritos intrínsecos, mas como uma forma de obter ações que beneficiem ou evitem danos a outras pessoas.”<sup>16</sup> O republicanismo parte do pressuposto segundo o qual a adoção de ideais de excelência articula-se a uma perspectiva comunitária, e não em uma esfera eminentemente privada. Nas palavras de Nino, “em contraste com o dogma do liberalismo, a visão perfeccionista recupera a

adoção de um modelo de virtude pessoal definido por ideais cívicos e comunitários.”<sup>17</sup>

Não obstante, segundo o autor, embora autores republicanos pressuponham uma conexão entre democracia e virtudes cívicas, não podemos ignorar que determinadas ações individuais são insuscetíveis de ensejar efeitos negativos na esfera alheia, refutando teorias perfeccionistas que interfiram no princípio da autonomia pessoal. Nino parte de uma distinção entre uma moralidade pública, intersubjetiva, e uma moralidade privada, autorreferente ou pessoal. Esta consiste “naqueles ideais de excelência pessoal ou virtude, que valoram as ações em relação a seus efeitos sobre a qualidade de vida ou o caráter dos mesmos agentes.”<sup>18</sup> Se o princípio da autonomia da pessoa atinge princípios intersubjetivos por meio dos quais julgamos as condutas dos indivíduos que afetam interesses alheios, é necessário limitar a autonomia de uns para preservar a de outros. No entanto, se o valor da autonomia contempla princípios morais autorreferentes, não há motivo para limitar uma esfera de autonomia pessoal, restringindo a eleição de ideais de excelência humana ou virtude pessoal, porque tais ideais não afetam a autonomia das demais pessoas.

A seu turno, o segundo princípio, denominado “princípio da inviolabilidade da pessoa”, é um mecanismo capaz de resguardar interesses individuais contra demandas coletivas. Na sua primeira formulação, “o princípio da inviolabilidade da pessoa proíbe a diminuição da autonomia de uma pessoa para alcançar como único propósito o incremento da autonomia que gozam outros indivíduos.”<sup>19</sup> O autor pretende, indubitavelmente, contrapor-se a concepções holísticas e coletivistas baseadas na “existência de uma entidade coletiva que constitui uma pessoa moral independente com interesses irredutíveis.”<sup>20</sup> Ou seja, o referido princípio veda a imposição de encargos e sacrifícios a certos indivíduos com o intuito de beneficiar a maioria da população. Trata-se de uma contraposição ao utilitarismo, que, para Rawls e Nino, desconhece a independência e a separabilidade das pessoas. Para o autor argentino, a articulação dos princípios da autonomia e da inviolabilidade da pessoa, entretanto, não constitui um referencial teórico suficiente para delinear uma concepção liberal de sociedade.

Em face desta leitura, refletindo sobre os limites teóricos dos referidos princípios, destaca que “o princípio da autonomia pessoal implica, paradoxalmente, uma supervisão permanente dos indivíduos para desqualificar qualquer decisão pessoal que restrinja sua própria autonomia, ainda quando o resultado obtido fora o incremento da autonomia de outras pessoas.”<sup>21</sup> Pretendendo preencher tal lacuna teórica, o autor delinea o “princípio da dignidade da pessoa humana”. Este princípio pressupõe as pessoas devem ser julgadas e tratadas exclusivamente com base em suas ações voluntárias e não segundo certos atributos e propriedades como raça, sexo, particularidades físicas ou classe social a que pertençam. Analisando o tema, Nino tece um dos mais lúcidos comentários a respeito do princípio da dignidade da pessoa, mencionando que:

(...) O limite que o princípio da dignidade da pessoa estabelece em relação ao princípio da inviolabilidade da pessoa, pode anular o limite que o último estabelece em relação ao princípio da autonomia da pessoa. Então, quando se aplica o princípio da dignidade da pessoa (já que a pessoa afetada consente em uma relação normativa resultante em uma perda de autonomia), a vedação à autonomia de um indivíduo, para incrementar a de outros, pode ser ignorada. Isto se manifesta no direito de que dispõem as instituições jurídicas de estabelecer obrigações e responsabilidade, dependendo do consentimento dos afetados - como ocorre nos casos de um contrato, o matrimônio e as leis penais - deveria estar justificado sobre a base de promover a autonomia da sociedade em geral. (...) <sup>22</sup>

Disso se infere que, para Nino, a deliberação democrática nem sempre constitui o método mais confiável para adotar soluções moralmente corretas que interfiram nos ideais de excelência pessoal e projetos pessoais de vida. Nas palavras do autor, “os juízes não têm razões para subordinar seus juízos morais a uma lei democrática baseada em ideais pessoais de virtude ou excelência. Não existe nenhuma base epistêmica que justifique essa decisão. Neste ponto, somente o juízo dos indivíduos é relevante.”<sup>23</sup> Em suma, a discussão democrática não possui valor epistêmico para impor ideais pessoais que violem o princípio da autonomia moral e, em consequência, qualquer juiz estaria legitimado para afastar a constitucionalidade de uma lei perfeccionista que imponha determinadas concepções de bem.

No ensejo, analisando o caso *Bowers v. Hardwick*<sup>24</sup>, no qual a Suprema Corte declarou a constitucionalidade de uma lei do Estado da Geórgia que proibia a conduta homossexual, o autor considera criticamente o posicionamento da mesma, tendo em vista a afirmação de argumentos perfeccionistas. Em síntese, alega ser papel do Judiciário declarar a inconstitucionalidade de uma legislação perfeccionista que imponha ideais de excelência. O autor estabelece uma distinção das questões de moralidade intersubjetiva, como, por exemplo, leis criminalizam o uso de droga para uso pessoal, das questões de moralidade autorreferente. Uma vez que lei protege dois bens jurídicos, objetivando também resguardar direitos de terceiras pessoas, seria legítimo que tal temática seja resolvida pelo processo político majoritário, pois não se trata de um ideal de excelência pessoal.

Sob esse aspecto, a ideia é investigar os motivos originários que deram ensejo à elaboração de determinadas leis, com o intuito de legitimar ou não o controle de constitucionalidade. Nas palavras de Nino, “se a proibição implica este último, ela só pode ser determinada pelo processo político ou corrigida por ele. Nesta situação, a proibição não deveria ser realizada pelo poder judicial, inclusive quando for incorreta.”<sup>25</sup> Indubitavelmente, é fundamental refletir sobre se a questão da criminalização da homofobia, do casamento gay, do transexualismo e do aborto de feto anencéfalos envolve um ideal de excelência pessoal ou um padrão moral intersubjetivo.

Diante do exposto, depreende-se que o direito ao casamento gay ou a pretensão de transexuais em realizar a adequação do prenome no registro civil,

configuram uma moralidade autorreferente vinculada a ideais de excelência humana ou virtude pessoal, sem repercussão na autonomia das demais pessoas. Em contraposição, a questão atinente à criminalização da homofobia representa uma moralidade intersubjetiva, pois o valor da autonomia pessoal relativa à não-discriminação por orientação sexual afeta interesses alheios relativos à liberdade religiosa e de expressão de entidades religiosas, devendo ser regulada por um processo político democrático sensível às pretensões normativas que são racionalizadas no espaço público.

Sob essa ótica, a questão é relevante para a problemática da anencefalia, pois, se partirmos do pressuposto segundo o qual não há bem jurídico vida a ser tutelado, trata-se, portanto, de uma moralidade autorreferente. Feitas essas considerações, passo à análise da questão da anencefalia e da legitimidade democrática do STF.

## 5. A questão da anencefalia no STF

A anencefalia caracteriza-se como uma malformação congênita, estando os hemisférios cerebrais ausentes ou representados por massas pequenas que repousam na base. É uma alteração na formação cerebral, resultante de falha no mecanismo de fechamento do tubo neural. A questão da anencefalia chegou pela primeira vez ao STF em 26 de fevereiro de 2004, no *Habeas Corpus* 84025-6, impetrado por Gabriela de Oliveira Cordeiro, uma jovem de 19 anos que pretendia autorização judicial, objetivando realizar a antecipação terapêutica de parto de feto anencéfalo. Gabriela e seu marido procuraram a Defensoria Pública de Teresópolis e, em novembro de 2003, solicitaram a autorização, sendo esta indeferida na comarca de Teresópolis.

Sob esse prisma, em 7 de novembro, o recurso foi interposto ao Tribunal de Justiça, obtendo a autorização em 9 de novembro. Contra essa decisão, foi interposto um recurso por dois advogados, mas a relatora do recurso, Desembargadora Gizelda Leitão Teixeira, confirmou o seu entendimento anterior, concedendo a autorização para a antecipação; todavia, nesse momento, o Movimento Pró-vida da Igreja Católica já havia impetrado o remédio de *Habeas Corpus* no Supremo Tribunal de Justiça. A Ministra Laurita Vaz cassou liminarmente a autorização judicial que havia sido concedida anteriormente.

Nessa trajetória jurídica, em 26 de fevereiro de 2004, entretanto, um pedido de *Habeas Corpus*, em favor de Gabriela, foi impetrado no Supremo Tribunal Federal pelas Organizações não-governamentais Anis (Instituto de Bioética, Direitos Humanos e Gênero) e Themis (Assessoria Jurídica e Estudos de Gênero), com base no direito à saúde, à liberdade e à dignidade. Em 4 de março, foi iniciada a votação, tendo como relator o Min. Joaquim Barbosa. Não obstante, o pedido foi considerado prejudicado, pois, em sessão de julgamento, o Tribunal foi informado do fato de que o parto da gestante fora realizado, dando à luz a uma

criança que veio a falecer alguns minutos após seu nascimento. Em decorrência, a ANIS e a THEMIS passaram a defender a tese de que o STF deveria solucionar o tema, razão pela qual foi interposta a ADPF n° 54. Mas, o voto do Min. J. Barbosa foi amplamente divulgado na mídia, sendo uma nítida expressão da sensibilidade do STF à temática da autonomia reprodutiva da mulher. Nesse quadro teórico, elucidativo o seguinte trecho do voto do Min. J. Barbosa:

(...) Nesse ponto, portanto, cumpre ressaltar que a procriação, a gestação, enfim, os direitos reprodutivos, são componentes indissociáveis do direito fundamental à liberdade e do princípio da autodeterminação pessoal, particularmente da mulher, razão por que, no presente caso, ainda com maior acerto, cumpre a esta Corte garantir seu legítimo exercício, nos limites ora esposados.

Lembro, invariavelmente, que essa concepção fundada no princípio da autonomia ou liberdade individual da mulher é a que tem prevalecido nas Cortes Constitucionais e Supremas que já se debruçaram sobre o tema.... (...)<sup>26</sup>

Nessa linha de raciocínio, tornar-se-à necessário postular que a problemática fundamental do debate teórico envolvido na questão da ADPF n° 54 consiste, nas palavras do advogado e constitucionalista, Luis Roberto Barroso, “se ao declarar a não incidência do Código Penal a uma determinada situação, estaria o STF interpretando a Constituição – que é o seu papel – ou criando uma nova hipótese de não punibilidade do aborto, em invasão da competência do legislador.”<sup>27</sup> E conclui: “Como se sabe, o Tribunal, por maioria, conheceu da ação, reconhecendo tratar-se de uma questão de interpretação constitucional e não de criação de direito novo.”<sup>28</sup> Em suma, de acordo com a doutrina, a interpretação conforme a Constituição pode implicar: a) uma delimitação do conteúdo da norma; b) sua não aplicação a um determinado contexto fático; c) a inconstitucionalidade de uma das normas decorrentes do texto, sendo que não há necessidade de uma declaração de inconstitucionalidade da norma, que permanece íntegra em ambas as situações.

Assim, o aborto provocado pela gestante ou por terceiro é tipificado nos artigos 124 a 126 do Código Penal. O aborto necessário, entretanto, indispensável para salvar a vida da gestante, não é punido, assim como o aborto no caso de gravidez resultante de estupro. Inexiste previsão, no artigo 128 do Código Penal, da possibilidade de antecipação terapêutica de parto decorrente da anencefalia, em razão do simples fato de que no momento histórico em que foi editado o Código Penal, os recursos tecnológicos eram escassos, inviabilizando o diagnóstico de anomalias fetais, tal como a anencefalia. Como sublinha o constitucionalista, “impor à mulher o dever de carregar por nove meses um feto que sabe, com plenitude de certeza, não sobreviverá, causando-lhe dor, angústia e frustração, importa violação de ambas as vertentes de sua dignidade humana”.<sup>29</sup>

Nesse sentido, a denominada técnica de “interpretação conforme a Constituição”, que tem sido desenvolvida pelo STF, ocorre quando, diante de uma variedade de métodos interpretativos, o Tribunal opta por uma linha de

interpretação possível a respeito de uma norma legal, adequando-a a um determinado contexto fático.

É premente postular que, no caso da ADPF n. 54/DF, a tese proposta por Luis Roberto Barroso foi a de que o STF procedesse à interpretação conforme a Constituição de dispositivos do Código Penal relativos ao aborto, determinando a sua não incidência ao contexto fático em que a gestante de feto anencefálico opte voluntariamente pela interrupção terapêutica da gravidez. A ADPF nº 54/DF foi ajuizada pela Confederação Nacional dos Trabalhadores na Saúde, afirmando em nota prévia serem distintas as situações de antecipação terapêutica e a do aborto, tendo em vista que este pressupõe potencialidade de vida extrauterina do feto. Articula o envolvimento de preceitos fundamentais relativos à dignidade da pessoa humana, da legalidade, da liberdade e da autonomia da vontade, bem como o do direito à saúde. Sobre a inexistência de outro meio eficaz para viabilizar a antecipação terapêutica do parto, evoca o fato retratado no HC 84025-6, declarado prejudicado pelo Plenário.

Em vista disso, o autor requereu a suspensão do andamento de processos ou dos efeitos de decisões judiciais que tenham por objeto a aplicação de dispositivos do Código Penal na hipótese de antecipação terapêutica de parto de fetos anencéfalos, estabelecendo o direito constitucional da gestante de se submeter a procedimento que leve à interrupção da gravidez, desde que atestada a ocorrência da anomalia. Assim, o pedido final constitui a declaração de inconstitucionalidade, com eficácia abrangente e efeito vinculante da interpretação dos artigos 124, 126, e 128, I e II do Código Penal, como impeditiva da interpretação terapêutica de parto. Em 1º de julho de 2004, o Min. Marco Aurélio de Mello concedeu o pedido liminar, reconhecendo o direito da gestante em optar pela antecipação terapêutica do parto, uma vez atestada a anencefalia por laudo médico. A liminar somente produziu efeitos por mais três meses, pois foi tornada sem efeito pelo Plenário do STF em 27 de abril de 2005, por sete votos a quatro; entretanto, igual votação admitiu que a ADPF fosse processada para que seu mérito fosse apreciado. Nesse momento, os Ministros Carlos Ayres de Britto, Gilmar Mendes, Sepúlveda Pertence e Nelson Jobim manifestaram-se no sentido de rejeitar a alegação de que seria incabível a propositura da ADPF, sendo vencidos os Ministros Eros Grau, Cesar Peluso e Ellen Gracie, que não conheceram da ação, tendo em vista que a mesma ofenderia a separação de poderes, criando uma nova hipótese de não punibilidade do aborto.

Diante dessa estrutura conceitual, trata-se de uma questão moralmente controversa que envolve direitos de minoria (no caso, a gestante), mas a matéria não foi inteiramente regulada pelo processo político, legitimando uma inevitável expansão da atuação judicial. Ademais, a problemática jurídica enfrentada distingue-se da questão constitucional enfrentada em *Roe vs. Wade*.<sup>30</sup> Nessa configuração normativa, analisando o tema, Eduardo Appio menciona que, no caso de aborto de feto anencefálico, a questão fundamental não é a proteção do direito à privacidade em confronto com o direito à vida, uma vez que, como alega a parte autora, inexistente qualquer possibilidade de vida da criança após o

nascimento. Nesse ponto, é oportuno transcrever as explicações teóricas de Eduardo Appio:

(...) Não existem, contudo, pontos de semelhança entre os casos *Roe v. Wade* e a ADPF 54, na medida em que o objeto central do debate em *Roe* foi a proteção da privacidade da mulher, assegurando-lhe um direito de escolha condicionado ao período de gestação. No caso do aborto de feto anencefálico, esta liberdade de escolha não está em jogo. Muito pelo contrário, os autores da ação alegam que em cem por cento dos casos não há possibilidade de sobrevivência do feto fora do útero materno após encerrado o período da gestação. Portanto, enquanto em *Roe v. Wade* se buscava proteger um “direito de escolha” da mulher gestante, na ADPF 54 esse direito não existe, porque sequer se coloca a questão sobre a futura condição da mãe da mulher envolvida. Sabe-se, de antemão, que, não havendo sobrevivência, não existe expectativa futura no tocante a esta condição e, portanto, não haverá limitação das opções futuras da mulher. (...) <sup>31</sup>

Indubitavelmente, a questão fundamental que é objeto de discussão na ADPF n.º 54, corresponde à proteção constitucional à saúde da gestante a ser tutelada pelo Estado, assim como sua integridade moral, tendo em vista a inexistência de uma estrutura cerebral própria que legitime a proteção estatal ao direito constitucional à vida do feto. Como assinala Eduardo Appio, “o que se encontra verdadeiramente em jogo é a integridade física e moral da gestante e, nesse sentido, o Estado tem o dever de assegurar um tratamento igual a todas as mulheres.” <sup>32</sup> Assim como o autor, compreendemos que, se a questão fundamental a ser analisada é a integridade física e moral da gestante, inexistindo proteção constitucional ao direito à vida. Nesse ponto, o Supremo Tribunal Federal tem legitimidade para tutelar o direito fundamental à integridade física e moral da mulher, ainda que se adote uma perspectiva minimalista, não sendo necessário definir o momento em que a vida tem início, que é objeto de um desacordo moral razoável em uma sociedade pluralista.

Nessa abordagem constitucional, não há, portanto, necessidade de uma postura maximalista, com implicações futuras na discussão sobre o direito ao aborto no Brasil, podendo o STF ater-se aos limites específicos da demanda, resguardando a integridade física e moral de uma minoria sem invasão à competência do legislador. Se inexistir previsão, no artigo 128 do Código Penal, da possibilidade de antecipação terapêutica de parto decorrente da anencefalia, isso não exime o Estado de assegurar o direito à integridade física e moral da gestante de fetos anencefálicos, sendo legítimo ao Judiciário declarar incidentalmente a inconstitucionalidade da referida lei, que estabelece tratamento diferenciado entre duas situações idênticas, em ofensa à isonomia. Compreendemos que, em um Estado pluralista, os fundamentos jurídicos utilizados para a análise da questão relativa ao aborto de fetos anencefálicos, devem estar pautados em argumentos de “razões públicas”, e não por autocompreensões de mundo particulares, religiosas ou morais, insuscetíveis de legitimar a

imposição estatal. A cada indivíduo deve ser assegurada uma esfera de autonomia moral que não pode ser instrumentalizada a um projeto político majoritário.

Nessa perspectiva, entretanto, não pretendemos legitimar, por exemplo, a atuação jurisdicional na proteção de uma esfera privada de gestantes de fetos anencéfalos e outras minorias como reflexo de teorias liberais que enfatizam o indivíduo atomizado, mas apenas de ressaltar que, quando maiorias parlamentares impõem doutrinas abrangentes, frustrando o debate, o Judiciário dispõe de legitimidade para proteger direitos fundamentais de minorias em uma dimensão substancialista, corrigindo os desvios do procedimento. Um exemplo da ausência de valor epistêmico do processo deliberativo na regulamentação de direitos de minorias são os diversos projetos de lei que tentavam tratar da antecipação terapêutica da gestação de fetos anencéfalos. Os Projetos de Lei n° 4.403/04, que inseria um inciso ao artigo 128 do Código Penal, e n° 4360/04 tentavam regulamentar a matéria, mas foram arquivados. Em 14 de Agosto de 2007, o deputado Pinotti apresentou requerimento de desarquivamento do Projeto de Lei n° 4.360/04, mas foi indeferido em 21 de Agosto. Em relação ao projeto n° 4.403/04, foi solicitado o desarquivamento em 2007, tendo a tramitação sido retomada na Comissão de Constituição e Justiça e de Cidadania da Câmara. Há o Projeto de Lei n° 183 de 2004, que é de autoria do Senador Duciomar Costa, pretendendo modificar o artigo 128 do Código Penal, para inserir o aborto de feto anencéfalo, sendo objeto de análise na Comissão de Constituição e Justiça do Senado no momento de elaboração do presente artigo.

Nesse quadro teórico, tais exemplos demonstram que, em uma sociedade como a brasileira, de população majoritariamente católica, a idéia habermasiana de um debate público franco e aberto – capaz de inspirar processos de aprendizagem entre cidadãos religiosos e concidadãos que professam doutrinas seculares – encontra-se distante da realidade fática, devendo ser analisada com a devida cautela em questões de moralidade autorreferente. Ademais, é patente também que o processo político majoritário não regulamenta adequadamente a questão em virtude de uma ausência de valor epistêmico em razão da qual as condições de abertura e participação dos afetados não estão presentes. Na prática, os sistemas partidários formam *lobbies* que atuam segundo acordos estratégicos no exercício do poder político. Esta descaracterização do uso público da razão na instância deliberativa legitima a reflexão individual exercida pelo Poder Judiciário nas questões de moralidade autorreferente. Defendemos que o STF – enquanto instância de representação de minorias – também pode funcionar como um fórum da razão pública capaz de resguardar uma esfera de moralidade autorreferente, sempre que as instâncias deliberativas não forem sensíveis a expectativas normativas relativas a determinadas minorias, tais como gestantes de fetos anencéfalos.

## 6. Conclusão

O conceito de razão pública assume especial relevância nas sociedades contemporâneas, que têm como elemento incontornável a existência do denominado desacordo moral razoável acerca de concepções de vida digna, tendo em vista a inexistência de um consenso substantivo sobre valores. Em um contexto marcado pelo pluralismo, o conceito de razão pública tem como aspecto fundamental a necessidade de justificação das decisões políticas, tanto na esfera legislativa, como na judicial, com base em valores que sejam publicamente aceitos por cidadãos.

Indubitavelmente, o voto do Min. Joaquim Barbosa foi capaz de demonstrar a independência do direito, revelando como arena constitucional - enquanto cenário simbólico de lutas pelo reconhecimento - tem potencialidade em conectar autonomia judicial e legitimidade democrática, com base em argumentos fundamentados em “razões públicas”. Tal decisão demonstra como juízes podem usar princípios constitucionais de abertura argumentativa para mediar o conflito e o diálogo público sobre pretensões normativas vinculadas a direitos de minorias. Nesse ponto, se a deliberação democrática falhar, no que se refere, ao ideal de neutralidade, tendo em vista a possibilidade fática de que majorias parlamentares imponham concepções abrangentes que frustram o debate, descaracteriza-se o uso público da razão e o valor epistêmico do processo democrático, legitimando-se a expansão da atividade judicial.

Outrossim, percebemos, com base em Nino, que a reflexão judicial também representa um meio moralmente válido de resolução do desacordo moral razoável, sempre que, descaracterizado o uso público da razão e o valor epistêmico do processo democrático, seja pela imposição de doutrinas religiosas abrangentes que violam direitos de minorias, seja pela ausência de abertura e participação dos afetados no processo político. Diante do exposto, depreende-se que, no caso da anencefalia, a questão fundamental constitui a integridade física e moral da gestante, sem implicações na esfera jurídica alheia, devendo ser considerada não como uma moralidade intersubjetiva, mas como uma moralidade autorreferente, tendo em vista a ausência de viabilidade de vida. A questão fundamental a ser considerada não é o direito à privacidade da gestante em face do direito à vida do feto, mas o direito à saúde e à integridade moral da gestante, buscando alcançar um tratamento isonômico na aplicação da lei penal.

## Referências

APPIO, Eduardo (2009). *Direito das Minorias*. São Paulo: Revista dos Tribunais.

AUDI, R. (1993): “The Place of Religious Argument in a Free and Democratic Society”. *San Diego Law Review*, v. 30, n. 4, 677-702.

- (2000): *Religious Commitment and Secular Reason*. Cambridge: Cambridge University Press.
- BARRY, B. (2001): *Culture and Equality: An Egalitarian Critique of Multiculturalism*. Cambridge (Mass.): Harvard University Press.
- BARROSO, Luis Roberto (2006): "Gestação de Fetos Anencefálicos e Pesquisas com Células-tronco: dois temas acerca da vida e da dignidade na Constituição". In: SARMENTO, Daniel e GALDINO, Flavio (orgs. ). *Direitos Fundamentais – Estudos em Homenagem ao Professor Ricardo Lobo Torres*. Rio de Janeiro: Renovar.
- BAYNES, K. (1992): *The Normative Grounds of Social Criticism: Kant, Rawls and Habermas*. Albany, State University of New York Press.
- BAYNES, K. and VON SCHOMBERG, R. (eds.) (2002): *Discourse and Democracy: Essays on Habermas's Between Facts and Norms*. Albany, State University of New York Press.
- BENHABIB, S. (2002): *The Claims of Culture: Equality and Diversity in the Global Era*. Princeton: Princeton University Press.
- BENHABIB, S. (ed.) (1996): *Democracy and Difference: Contesting the Boundaries of the Political*. Princeton: Princeton University Press.
- BERGER, P. (ed.) (1999): *The Desecularization of the World: Resurgent Religion and World Politics*. Washington: Ethics and Public Policy Center.
- BROWNING, D. and SCHÜSSLER FIORENZA, F. (eds.) (1992): *Habermas, Modernity, and Public Theology*. New York: Crossroad Publishers.
- COOKE, M. (1997): "Authenticity and Autonomy: Taylor, Habermas, and the Politics of Recognition". *Political Theory*, v. 25, n. 2, 258-288.
- (2000): "Five Arguments for Deliberative Democracy", *Political Studies*, v. 48, n. 5, 947-969.
- (2006): "Salvaging and Secularizing the Semantic Contents of Religion: The Limitations of Habermas's Postmetaphysical Proposal," *International Journal of the Philosophy of Religion*, v. 60, ns. 1-3, 187-207.
- (2007): "A Secular State for a Postsecular Society? Postmetaphysical Political Theory and the Place of Religion". *Constellations*, v. 14, n. 2, 224-238.
- CUNEO, T. (ed.) (2004): *Religion in the Liberal Polity*. Notre Dame: University of Notre Dame Press.
- DOMBROWSKI, D. (2001): *Rawls and Religion: The Case for Political Liberalism*. Albany: State University of New York Press.
- DWORKIN, R. (2000): *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge (Mass.): Harvard University Press.
- ESTLUND, D. (1998): "The Insularity of the Reasonable: Why Political Liberalism Must Admit the Truth". *Ethics*, v. 108, 252-275.
- FACHIN, Luiz Edson (1997): "Aspectos Jurídicos da União de Pessoas do Mesmo Sexo". In: *A nova família: problemas e perspectivas*. Rio de Janeiro: Renovar.
- FRASER, N. (1997): *Justice Interruptus: Critical Reflections on the Postsocialist Condition*. London: Routledge.

- FRASER, N. and HONNETH, A. (2003): *Redistribution or Recognition? A Political-Philosophical Exchange*. London: Verso.
- FREEMAN, S. (2006): *Justice and the Social Contract: Essays on Rawlsian Political Philosophy*. Oxford: Oxford University Press.
- (2007): *Rawls* (Routledge Philosophers). London: Routledge.
- FREEMAN, S. (ed.) (2002): *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press.
- GUTMANN, A. (ed.) (1994): *Multiculturalism: Examining the Politics of Recognition*. Princeton: Princeton University Press.
- (2004): *Identity in Democracy*. Princeton: Princeton University Press.
- HABERMAS, J. (1997): *Direito e Democracia: entre facticidade e validade*. Rio de Janeiro: Tempo Brasileiro [*Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Frankfurt: Suhrkamp, 1992].
- (1998a): *The Inclusion of the Other: Studies in Political Theory*. Cambridge (Mass.): MIT Press [*Die Einbeziehung des Anderen. Studien zur politischen Theorie*. Frankfurt: Suhrkamp, 1996].
- (1998b) HABERMAS, Jürgen. “Reconciliación mediante el uso público de la razón”. In: HABERMAS, Jürgen; RAWLS, John. *Debate sobre el Liberalismo Político*. Barcelona: Paidós.
- (2001a): *A Constelação Pós-Nacional: Ensaio Político*. São Paulo: Littera Mundi, 2001 [*Die postnationale Konstellation. Politische Essays*. Frankfurt: Suhrkamp, 1998].
- (2001b): *Kommunikatives Handeln und detranszendentalisierte Vernunft*. Stuttgart, Reclam Verlag [in: 2007].
- (2002a): *Religion and Rationality: Essays on Reason, God, and Modernity* (edited by Eduardo Mendieta). Cambridge (Mass.): MIT Press.
- (2002b): *Glauben und Wissen*. Frankfurt: Suhrkamp [in: 2004b].
- (2003): *Era das Transições*. Rio de Janeiro: Tempo Brasileiro [*Zeit der Übergänge. Kleine politische Schriften IX*. Frankfurt: Suhrkamp, 2001].
- (2004a): *Verdade e Justificação: Ensaio Filosófico*. São Paulo: Edições Loyola [*Wahrheit und Rechtfertigung. Philosophische Aufsätze*. Frankfurt: Suhrkamp, 1999].
- (2004b): *O Futuro da Natureza Humana: a caminho de uma eugenia liberal?* São Paulo: Martins Fontes [*Die Zukunft der menschlichen Natur. Auf dem Weg zu einer liberalen Eugenik?* Frankfurt: Suhrkamp, 2001].
- (2006a): *O Ocidente Dividido*. Rio de Janeiro: Tempo Brasileiro [*Der gespaltene Westen. Kleine politische Schriften X*. Frankfurt: Suhrkamp, 2004].
- (2006b): *The Dialectics of Secularization: On Reason and Religion* (with Joseph Ratzinger; edited by Florian Schuller). San Francisco: Ignatius Press, 2006 [*Dialektik der Säkularisierung. Über Vernunft und Religion*. Vienna: Herder Verlag, 2005; in: 2007].
- (2007): *Entre Naturalismo e Religião: Estudos Filosóficos*. Rio de Janeiro: Tempo Brasileiro [*Zwischen Naturalismus und Religion. Philosophische Aufsätze*. Frankfurt: Suhrkamp, 2005].

- HARRINGTON, A. (2007): "Habermas and the Post-Secular Society". *European Journal of Social Theory*, v. 10, n. 4, 543-560.
- (2007): "Habermas's Theological Turn?" *Journal for the Theory of Social Behaviour*, v. 37, n. 1, 45-61.
- HELD, D. (1996): *Models of Democracy*. Stanford: Stanford University Press.
- HEYD, D. (1996): *Toleration: An Elusive Virtue*. Princeton: Princeton University Press.
- LAFONT, C. (2003): Procedural justice? Implications of the Rawls–Habermas Debate for Discourse Ethics". *Philosophy and Social Criticism*, v. 29, n. 2, 163-181.
- (2007): "Religion in the Public Sphere: Remarks on Habermas's Conception of Public Deliberation in Postsecular Societies". *Constellations*, v. 14, n. 2, 239-259.
- MACEDO, S. (ed.) (1999): *Deliberative Politics: Essays on Democracy and Disagreement*. Oxford: Oxford University Press.
- McKINNON, C. and CASTIGLIONE, D. (eds.) (2003): *The Culture of Toleration in Diverse Societies: Reasonable Tolerance*. Manchester: Manchester University Press.
- NINO, Santiago. (1989): *El Constructivismo Ético*. Madrid: Centro de Estudios Constitucionales.
- (1992): *Los Fundamentos Del Derecho Constitucional*. Buenos Aires: Astraa.
- (1997): *La Constitución de la democracia deliberativa* Barcelona: Gedisa.
- NORRIS, P. and INGLEHART, R. (2004): *Sacred and Secular: Religion and Politics Worldwide*. Cambridge: Cambridge University Press.
- PERES, A. P. (2001): *Transexualismo - O Direito a uma Identidade Sexual*. Rio de Janeiro: Renovar.
- PERRY, M. (1999): *Religion in Politics: Constitutional and Moral Perspectives*. Oxford: Oxford University Press.
- (2003): *Under God? Religious Faith and Liberal Democracy*. Cambridge: Cambridge University Press.
- RAWLS, J. (1996): *Political Liberalism* (paperback edition with a new introduction). New York: Columbia University Press.
- (1998): "Réplica a Habermas". In: HABERMAS, Jürgen; RAWLS, John. *Debate sobre el Liberalismo Político*. Barcelona: Paidós.
- (1999a): *The Law of Peoples; with "The Idea of Public Reason Revisited"*. Cambridge (Mass.): Harvard University Press.
- (1999b): *Collected Papers* (edited by Samuel Freeman). Cambridge (Mass.): Harvard University Press.
- (2000a): *Lectures on the history of moral philosophy* (edited by Barbara Herman). Cambridge (Mass.): Harvard University Press.
- (2000b): *O Liberalismo Político*. São Paulo: Ática.

— (2001): *Justice as Fairness: A Restatement* (edited by Erin Kelly). Cambridge (Mass.): Harvard University Press.

— (2007): *Lectures on the history of political philosophy* (edited by Samuel Freeman). Cambridge (Mass.): Harvard University Press.

RAZ, J. (1994): *Ethics in the Public Domain*. Oxford: Clarendon Press.

— (1999): *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press.

## Notas

<sup>1</sup> Doutora e Mestre em Teoria do Estado e Direito Constitucional pela PUC-Rio. Autora do livro, "O Patriotismo Constitucional na perspectiva de Jürgen Habermas", Ed. Lumen Juris, 2010. Professora e Pesquisadora do Unifoa-Centro Universitário de Volta Redonda. Pós-doutoranda em Filosofia pela UFSC com bolsa UFSC-REUNI.

<sup>2</sup> RAWLS, John. *O Liberalismo Político*. São Paulo: Ática, 2000, p. 266.

<sup>3</sup> *Ibidem*, p. 93.

<sup>4</sup> *Ibidem*, p. 264.

<sup>5</sup> *Ibidem*, p. 264.

<sup>6</sup> *Ibidem*, p. 264.

<sup>7</sup> RAWLS, John . "The Idea of Public Reason Revisited". In: *Collected Papers*. Cambridge: Harvard University Press, 1999. Neste ensaio, Rawls irá estabelecer algumas modificações na ideia de razão pública delineada na sexta conferência da edição de 1993 *Political Liberalism*

<sup>8</sup> HABERMAS, Jürgen. "Reconciliación mediante el uso público de la razón". In: HABERMAS, Jürgen; RAWLS, John. *Debate sobre el Liberalismo Político*. Barcelona: Paidós, 1998, p. 53.

<sup>9</sup> *Ibidem*, p. 53.

<sup>10</sup> RAWLS, John. "Réplica a Habermas". In: HABERMAS, Jürgen; RAWLS, John. *Debate sobre el Liberalismo Político*. Barcelona: Paidós, 1998.

<sup>11</sup> HABERMAS, Jürgen. "Religião na Esfera Pública. Pressuposições Cognitivas para o Uso Público da Razão de Cidadãos Seculares e Religiosos". In: HABERMAS, Jürgen. *Entre Naturalismo e Religião – Estudos Filosóficos*. Rio de Janeiro: Tempo Brasileiro, 2007, p. 144.

<sup>12</sup> NINO, Santiago. *El Constructivismo Ético*. Madrid: Centro de Estudios Constitucionales, 1989, p. 14.

<sup>13</sup> NINO, Santiago. *La Constitución de la democracia deliberativa*. Barcelona: Gedisa, 1997, p. 161.

<sup>14</sup> NINO, Santiago. *La Constitución de la democracia deliberativa*. Barcelona: Gedisa, 1997, p. 278.

<sup>15</sup> *Ibidem*, p. 140.

<sup>16</sup> *Ibidem*, p. 141.

<sup>17</sup> *Ibidem*, p. 140.

<sup>18</sup> *Ibidem*, p. 278.

<sup>19</sup> *Ibidem*, p. 79.

<sup>20</sup> *Ibidem*, p. 79.

<sup>21</sup> *Ibidem*, p. 80.

<sup>22</sup> *Ibidem*, p. 80-81.

<sup>23</sup> *Ibidem*, p. 278.

<sup>24</sup> *Bowers v. Hardwick*. 478 U. S 186 (1986).

<sup>25</sup> *Ibidem*, p. 279-280.

<sup>26</sup> STF, HC n. 84025-6. Voto do Min. Joaquim Barbosa.

<sup>27</sup> BARROSO, Luis Roberto. "Gestação de Fetos Anencefálicos e Pesquisas com Células-tronco: dois temas acerca da vida e da dignidade na Constituição". In: SARMENTO, Daniel e GALDINO, Flavio (orgs. ). *Direitos Fundamentais - Estudos em Homenagem ao Professor Ricardo Lobo Torres*. Rio de Janeiro: Renovar, 2006, p. 700.

<sup>28</sup> *Ibidem*, p. 701.

<sup>29</sup> *Ibidem*, p. 684.

<sup>30</sup> Roe v. Wade. 410 U.S. 113 (1973).

<sup>31</sup> Roe v. Wade. 410 U.S. 113 (1973).

<sup>32</sup> APPIO, Eduardo. *Direito das Minorias*. São Paulo: Revista dos Tribunais, 2009, p. 358.

# A “FÓRMULA DO FIM EM SI” E A OBJEÇÃO DE FORMALISMO À FILOSOFIA MORAL DE KANT.

MILENE CONSENSO TONETTO

UFSC

mitonetto@yahoo.com.br

## Considerações iniciais

A filosofia moral de Kant tem recebido várias críticas e dentre elas pode-se destacar a de ser meramente formal. Hegel, por exemplo, no parágrafo 135 de *Princípios da Filosofia do Direito*, afirma que a ética kantiana do “dever pelo dever” leva a um vão formalismo (HEGEL, 1997)<sup>1</sup>. Neste trabalho, pretende-se investigar se alguns elementos do projeto de uma “metafísica dos costumes” de Kant podem dar uma resposta a essa objeção. Dois pontos principais são explorados. Primeiro, analisa-se de que modo o princípio da dignidade da humanidade, elaborado a partir da “Fórmula do Fim em Si” do Imperativo Categórico<sup>2</sup>, fornece conteúdo à ética de Kant. O estabelecimento dessa fórmula está fortemente relacionado com a teoria dos fins que Kant defende de modo mais sistemático na *Metafísica dos Costumes* (1797). A partir da ideia de que há fins que são deveres confirmados pela razão, Kant defende que o Imperativo Categórico possui um conteúdo. Num segundo momento, examina-se qual o papel da “Fórmula do Fim em Si” na derivação dos deveres de virtude, mais especificamente dos deveres para consigo próprio. Assim, pretende-se mostrar que outros aspectos da teoria de Kant são importantes para a compreensão de sua filosofia moral e que ela não se resume a um mero formalismo.

## 1. A “Fórmula do Fim em Si”

O principal objetivo de Kant na *Fundamentação da Metafísica dos Costumes* é apresentar o princípio supremo da moralidade, a saber, o Imperativo Categórico. De acordo com ele, a máxima da minha ação deve ser submetida a um teste a fim de torná-la uma regra universal. É importante mencionar que para se identificar o valor moral da ação, segundo Kant, são necessárias três etapas: 1) identificar as máximas, ou seja, os princípios subjetivos das ações; 2) fazer um

“teste” dessas máximas pelo Imperativo Categórico e 3) cumprir as leis práticas que passaram pelo teste de um modo específico, a saber, pelo puro respeito ao dever. A crítica de Hegel, mencionada no início deste trabalho, parece se referir somente a esta terceira etapa e assim negligenciar outros aspectos importantes na determinação dos deveres morais. A seguir, isso será averiguado.

A fórmula *geral* do Imperativo Categórico é apresentada nos seguintes termos: “Age apenas segundo uma máxima tal que possas ao mesmo tempo querer que ela se torne lei universal” (GMS, AA 04: 420)<sup>3</sup>. Como se pode perceber, ela consiste num teste completamente formal. Mas isso não pode ser sustentado em relação à segunda formulação do Imperativo Categórico, que para Kant é outra maneira de representar este mesmo princípio da moralidade (GMS, AA 04: 421, 436)<sup>4</sup>. Nela, Kant introduz a noção de humanidade da seguinte maneira: “Age de tal modo que uses a humanidade, tanto na tua pessoa como na pessoa de qualquer outro, sempre e simultaneamente como fim e nunca simplesmente como meio” (GMS, AA 04: 429). Para entender essa formulação, devemos considerar dois conceitos, a saber, o de vontade e o de fim. Segundo Kant, tudo na natureza age segundo leis. Mas, “só um ser racional tem a capacidade de agir *segundo a representação* das leis, isto é, segundo princípios, ou: só ele tem uma *vontade*. Como para derivar as ações das leis é necessária a razão, a vontade não é outra coisa senão razão prática” (GMS, AA 04: 413). Quer dizer, a vontade é a faculdade existente em nós de derivar resultados práticos (isto é, ações) das exigências da razão. E, mais adiante, ele explica porque a vontade deve ser determinada por um fim:

A vontade é concebida como a faculdade de se determinar a si mesmo a agir em conformidade com a representação de certas leis. E uma tal faculdade só se pode encontrar em seres racionais. Ora, aquilo que serve à vontade de princípio objetivo da sua autodeterminação é o fim (*Zweck*), e este, se é dado só pela razão, tem de ser válido igualmente para todos os seres racionais. (GMS, AA 04: 427).

A vontade de um ser racional está sempre direcionada a um fim que ele se coloca.

O conceito de fim ou matéria aparece na *Fundamentação*, mas é explicado com maior clareza na *Metafísica dos Costumes*. Na introdução da *Doutrina da Virtude*, Kant afirma: “a ética oferece uma matéria (um objeto do arbítrio livre), um fim da razão pura, que ao mesmo tempo se apresenta como um fim objetivamente necessário, isto é, um dever para o homem” (TL, AA 06: 380). O fim é, portanto, um objeto do livre arbítrio (*Willkür*), cuja representação o determina para uma ação. Toda ação tem um fim e isso não é um efeito da natureza, mas um ato da liberdade do agente. Assim, para Kant, “tem, pois, que haver um tal fim e um imperativo categórico que lhe seja correspondente” (TL, AA 06: 385). Quer dizer, os fins que são ao mesmo tempo deveres, a saber, a perfeição própria e a felicidade alheia<sup>5</sup>, não são estabelecidos a partir dos impulsos sensíveis da natureza humana, mas a partir do seu livre arbítrio. A razão pura prática “é uma

faculdade dos fins em geral; portanto, ser indiferente em relação a eles, isto é, não ter qualquer interesse neles [nos fins], é uma contradição; porque, então, a razão pura prática tão pouco determinaria as máximas relativas às ações (como contendo sempre estas últimas um fim), não sendo, nessa conformidade, razão prática” (TL, AA 06: 392). Sendo assim, Kant defende que o Imperativo Categórico tem uma forma e também possui uma matéria, isto é, um fim.

Tendo esclarecido a relação da segunda fórmula com os fins da razão pura, pode-se agora analisar propriamente o conceito de dignidade. Para Kant, todos os seres racionais estão submetidos à lei que ordena que cada um deles jamais se trate a si mesmo ou aos outros “*simplesmente como meios*, mas sempre *simultaneamente como fins em si*”. Disso resulta “uma ligação sistemática de seres racionais por meio de leis objetivas comuns, isto é, um reino que, exatamente porque estas leis têm em vista a relação destes seres uns com os outros como fins e meios, se pode chamar um reino dos fins” (GMS, AA 04: 433).

No reino dos fins tudo tem um preço ou uma dignidade. “Quando uma coisa tem um preço, pode-se pôr em vez dela qualquer outra como *equivalente*; mas quando uma coisa está acima de todo o preço e, portanto, não permite equivalente, então tem ela dignidade”. (GMS AA 04: 434). Kant sustenta que a condição para algo ser considerado um fim em si mesmo é a de possuir dignidade: “Aquilo (...) que constitui a condição só graças à qual qualquer coisa pode ser um fim em si mesma, não tem somente um valor relativo, isto é, um preço, mas um valor íntimo, isto é, *dignidade*” (GMS, AA 04: 435). Desse modo, deve-se tratar a humanidade enquanto fim em si mesma porque ela possui dignidade.

É importante ter claro o que Kant entende por humanidade. Na *Fundamentação*, esse conceito é definido como capacidade de estabelecer fins: “a natureza racional distingue-se das restantes por se pôr a si mesma um fim” (GMS, AA 04: 437). Do mesmo modo, na *Doutrina da Virtude*, Kant afirma que “a capacidade de em geral se propor um fim, qualquer que ele seja, é o que constitui o elemento característico da humanidade (ao invés da animalidade)” (TL, AA 06: 392). Ao fim da humanidade na nossa própria pessoa está associada também a vontade racional e, por conseguinte, “o dever de se tornar digna da humanidade mediante a cultura em geral, o dever de buscar ou de promover a capacidade de realizar quaisquer fins possíveis, na medida em que esta faculdade só no homem é suscetível de ser encontrada” (TL, AA 06: 392).

Para Allen Wood, a ideia da dignidade humana que fundamenta a “Fórmula do Fim em Si” é “o princípio kantiano que tem talvez a maior ressonância na consciência moral de nossa cultura. Essa ideia tem também o maior apelo universal, pois ela parece fundamentar os direitos humanos (...)” (WOOD, 1998, p. 165). De fato, na *Fundamentação*, Kant menciona que tratar um indivíduo enquanto mero meio, ou seja, não respeitar a humanidade de uma pessoa como um fim em si mesma, constitui uma violação de um direito humano. Nas palavras dele,

[m]ais claramente ainda dá na vista esta colisão com o princípio da humanidade em outros homens quando tomamos para exemplos ataques

à liberdade ou à propriedade alheias. Porque então é evidente que o violador dos direitos dos homens tenciona servir-se das pessoas dos outros simplesmente como meios, sem considerar que eles, como seres racionais, devem ser sempre tratados ao mesmo tempo como fins (...). (GMS AA 04: 430)

O ser humano não é uma coisa e, por isso, não pode ser utilizado arbitrariamente pela vontade dos outros. Nesse sentido, não se pode dispor do homem na minha pessoa e na dos outros para mutilá-lo, degradá-lo ou matá-lo.

É nítido que, na *Metafísica dos Costumes*, Kant recorre à “Fórmula do Fim em Si” para derivar deveres jurídicos. Pode-se confirmar isso já na divisão geral dos deveres de direito, onde Kant comenta o princípio “*honeste vive*” da seguinte maneira: “*Sê um homem honesto (honeste vive)*. A *honestidade jurídica (honestas iuridica)* consiste em manter, na relação com outros, seu valor como o de um homem, dever que é expresso pela proposição: ‘Não faças de ti um mero meio para os outros, mas sê ao mesmo tempo fim para eles’ “ (RL, AA 06: 236). Kant também recorre a esse princípio para fundamentar o único direito inato, a saber, a liberdade que, segundo ele, “pertence a cada homem por força de sua humanidade” (RL, AA 06: 237). Muitos dos deveres éticos citados por Kant na *Doutrina da Virtude*, por exemplo, os deveres para consigo próprio, estão explicitamente baseados na “Fórmula do Fim em Si”. A seguir, esses deveres são analisados.

## 2. A “Fórmula do Fim em Si” e os deveres para consigo

O primeiro dever para consigo estabelecido por Kant na *Doutrina da Virtude* é o de não cometer suicídio. O homem está, pela sua qualidade de pessoa, “obrigado a conservar sua vida e se tem nisso que reconhecer um dever [...] para consigo próprio” (TL, AA 06: 422). Segundo ele, o homem não pode alienar a sua personalidade na medida em que é um sujeito de deveres, por conseguinte, enquanto ele vive. “É uma contradição que [o homem] tenha a faculdade de se subtrair a toda a obrigação, isto é, a agir livremente, como se não necessitasse de qualquer autorização para esta ação” (TL, AA 06: 422). Desse modo, o suicídio é um ato que, por destruir a própria existência do indivíduo, destrói também a possibilidade da liberdade do ser humano. Destruir o sujeito da moralidade na sua própria pessoa equivale a extirpar do mundo a moralidade na sua própria existência, na medida em que esta depende dele. “Por conseguinte, dispor de si próprio como de um simples meio para um fim qualquer é equivalente a desvirtuar a humanidade na sua própria pessoa (*homo noumenon*), à qual, no entanto, foi confiada a conservação do homem (*homo phaenomenon*)” (TL, AA 06: 423)<sup>6</sup>.

Nos deveres perfeitos para consigo ainda pode ser incluída a proibição da desonra de si próprio pela voluptuosidade ou prazer carnal (TL, AA 06: 424).

Kant menciona o uso antinatural, portanto, abusivo das faculdades sexuais do indivíduo, pois viola o dever para consigo próprio e, assim, contraria a moralidade. O fundamento da prova para mostrar que o uso abusivo das faculdades sexuais é inadmissível consiste “em que o homem renuncie com isso (desdenhosamente) à sua personalidade, ao usar-se meramente como meio para a satisfação dos impulsos animais” (TL, AA 06: 425). Ao usar-se como mero meio o homem renuncia à sua liberdade. Segundo Kant, tal vício representa um elevado grau de violação da humanidade na própria pessoa. “Aquele que se entrega por completo à inclinação animal converte o homem numa coisa que se pode usufruir, mas, com isso, também numa coisa contrária à natureza, quer dizer, num objeto que suscita repulsa” (TL, AA 06: 425).

Outra violação de um dever para consigo analisado por Kant é o “do aturdimento de si próprio pela imoderação no uso da bebida ou da comida” (TL, AA 06: 427), isto é, o alcoolismo e a gula. Para Kant, colocar-se num estado de embriaguez constitui uma violação de um dever para consigo, pois nesse estado a pessoa “é como um mero animal e não se trata como um ser humano” (TL, AA 06: 427). Por causa do excesso de comidas e bebidas, o ser humano fica inibido durante algum período para realizar ações que requerem agilidade e reflexão no uso de suas forças. Os seres humanos, assim como os outros animais, possuem corpos que precisam ser nutridos. Sem se alimentar eles não podem sobreviver muito menos agir. Desse modo, os seres humanos devem fazer uma boa alimentação para preservar e promover a habilidade de agir racionalmente e livremente e não para diminuir essa capacidade. Porções moderadas de álcool podem trazer benefícios à saúde e alimentos variados são indispensáveis à vida para se ter a habilidade de agir livre e racionalmente.

Em seguida, Kant lista os deveres perfeitos para consigo como um ser moral. Ao considerar o homem como um ser moral, ele leva em conta unicamente a humanidade da pessoa. Desse modo, reconhece deveres que visam preservar o caráter e a dignidade e não só a existência corporal. Esses deveres opõem-se aos seguintes vícios: à mentira, à avareza e à falsa humildade (servilismo).

A maior violação do dever do homem para consigo próprio, enquanto ser moral, é o contrário da veracidade: a mentira. “A mentira é rejeição e, por assim dizer, destruição da própria dignidade do homem” (TL, AA 06: 429). Ao mentir a pessoa renuncia a faculdade de comunicar seus pensamentos a outras pessoas. Desse modo, ela destrói uma capacidade natural de comunicação que alguém pode usar no exercício de sua liberdade.

Por avareza (sovinice), Kant entende “a restrição da sua própria fruição dos meios para viver bem abaixo da medida da sua verdadeira necessidade” (TL, AA 06: 432). Isso, segundo Kant, opõe-se ao dever para consigo próprio. A ideia de Kant é que ao transformar o que devia ser um mero meio para alcançar um fim escolhido livremente em algo que vai ser acumulado e nunca usado, renuncia-se a capacidade de agir livremente.

Quanto ao servilismo, Kant entende que o homem não pode causar prejuízo à consciência de sua dignidade como homem racional e, desse modo, não deve

renunciar à auto-estima moral. Isto é, não deve procurar alcançar o seu fim, que é em si mesmo um dever, rebaixando-se de modo servil, como se buscasse um favor. O homem como animal racional (*homo phaenomenon*) é um ser de escassa importância e tem com os outros animais um valor comum.

Somente o homem, considerado como pessoa, isto é, como sujeito de uma razão prático-moral, está acima de todo o preço; pois que, como tal (como *homo noumenon*), não pode valorar-se apenas como meio para fins alheios, mas sim como fim em si mesmo, isto é, possui uma dignidade (um valor intrínseco absoluto) mediante a qual obriga todos os demais seres racionais do mundo a guardar-lhe respeito, podendo medir-se com qualquer outro desta espécie e valorar-se em pé de igualdade. (TL, AA 06: 434-5)

O homem não deve, portanto, renunciar a sua dignidade, mas sim estar sempre com a consciência do caráter sublime da sua disposição moral. Esta auto-estima é um dever do homem para consigo próprio.

Dando continuidade à exposição dos deveres, Kant trata do dever do homem para consigo próprio, enquanto juiz inato de si mesmo. Para ele, além do conhecimento da lei moral, cada ser humano também tem uma consciência moral. “A consciência de um tribunal interior do homem (perante o qual os seus pensamentos se acusam ou desculpam entre si) é a consciência moral” (TL, AA 06: 438). Todo homem tem consciência moral e encontra-se em observação por um juiz interior, que o ameaça e o mantém em respeito. E este respeito que se mantém nele pela observância das leis não é algo que elabora para si próprio, mas algo que está incorporado no seu ser. A partir disso, Kant anuncia o primeiro mandamento de todos os deveres para consigo próprio, dessa forma: “conhece-te a ti mesmo (examina-te, sonda-te) não quanto à tua perfeição física [...] mas quanto à perfeição moral em relação com o teu dever” (TL, AA 06: 441). O autoconhecimento moral requer que se penetre até as profundezas do coração é o começo de toda prudência humana. Ele exige do homem o afastamento dos obstáculos interiores e depois o desenvolvimento nele de uma vontade boa. A imparcialidade no julgamento que se faz de si próprio em comparação com a lei e a sinceridade na confissão pessoal do seu próprio valor moral interno ou a falta dele são deveres para consigo próprio, que decorrem diretamente deste primeiro mandamento de conhecer a si mesmo.

Para finalizar a lista dos deveres para consigo próprio, Kant defende o dever de desenvolver e aumentar a perfeição natural com um propósito pragmático. O homem não deve deixar de aproveitar a disposição natural e as faculdades que a sua razão pode fazer uso. Como ser capaz de se propor fins, não tem que agradecer simplesmente o uso das faculdades ao instinto natural pela vantagem que isso possa proporcionar. Assim, é um mandado da razão prático-moral cultivar as faculdades pessoais e ser um homem adaptado ao fim da própria existência. Nas palavras de Kant, “é dever do homem para consigo próprio ser um elemento útil para o mundo, porque isto também faz parte do valor da

humanidade na sua própria pessoa, valor que ele não deve, pois, degradar” (TL, AA 06: 446).

Como se pode perceber, os deveres de virtude incluem os deveres de preservar a existência própria, as capacidades morais, físicas e mentais, além disso, desenvolver as capacidades e talentos e preservar a dignidade da própria pessoa. Assim, pode-se defender que os deveres de virtude para consigo surgem da necessidade dos seres humanos preservar a dignidade e, conseqüentemente, da necessidade de não serem tratados como meros meios, de se preservar a existência de seres livres etc. Os deveres de virtude acima analisados visam preservar a possibilidade dos indivíduos exercitarem sua liberdade através do desenvolvimento dos talentos e habilidades que são usados para alcançar seus fins livremente estabelecidos. Na medida em que esses fins são alcançados, os deveres de virtude promovem a liberdade dos indivíduos e, além disso, contribuem com a sua felicidade.

## Considerações finais

Diante dessas observações, pode-se afirmar que a filosofia moral de Kant não pode ser acusada de formalista. A filosofia moral de Kant *não* requer apenas o cumprimento do “dever pelo dever”. Como foi visto, a partir do princípio da dignidade da humanidade, ligado à “Fórmula do Fim em Si” do Imperativo Categórico, Kant deriva os deveres de virtude para consigo na *Metafísica dos Costumes*. Sendo assim, a partir da segunda formulação do Imperativo Categórico, pode-se obter um guia concreto para a conduta humana. Desse modo, a segunda formulação do Imperativo Categórico e os deveres de virtude mostram que a filosofia moral de Kant possui conteúdo.

### Sistema de abreviações:

Citações das obras de Kant de acordo com a *Academie-Ausgabe* [AA]:

- GMS* Grundlegung zur Metaphysik der Sitten (AA 04)  
(Cito tradução de Paulo Quintela: *Fundamentação da Metafísica dos Costumes*. Lisboa: Edições 70, 1986.)
- MS* Die Metaphysik der Sitten (AA 06)
- RL* Metaphysische Anfangsgründe der Rechtstlehre (AA 06)  
(Cito tradução de Joãozinho Beckenkamp. *Princípios Metafísicos da Doutrina do Direito*. Obra não publicada.)
- TL* Metaphysische Anfangsgründe der Tugendlehre (AA 06)  
(Cito tradução de José Lamego: *Princípios Metafísicos da Doutrina da Virtude*. Lisboa: Fundação Calouste Gulbenkian, 2005.)

## Referências

- KANT, I. *Fundamentação da metafísica dos costumes*. Lisboa: Edições 70, 1986.
- KANT, I. *A metafísica dos costumes*. Lisboa: Fundação Calouste Gulbenkian, 2005.
- GUYER, P. (Ed). *The Cambridge Companion to Kant*. Cambridge: Cambridge University Press, 1999.
- GUYER, P. *Kant's System of Nature and Freedom: Selected Essays*. Oxford: Clarendon Press, 2005.
- HEGEL, F. *Princípios da Filosofia do Direito*. São Paulo, Martins Fontes, 1997.
- HILL, T. Jr. *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca, NY: Cornell University Press, 1992.
- WOOD, A. Humanity As End in Itself. In: GUYER, P. *Kant's Groundwork of the Metaphysics of Morals. Critical Essays*. Lanham (MA): Rowman & Littlefield, 1998.

## Notas

<sup>1</sup> No mesmo parágrafo, Hegel faz outras críticas à filosofia moral de Kant: “Poder-se-á decerto recorrer a uma matéria exterior e assim chegar a deveres particulares, mas desta definição do dever como ausência de contradição ou como acordo formal consigo [...] não se pode passar à definição dos deveres particulares, e quando um conteúdo particular de comportamento chega a ser considerado, aquele princípio não oferece o critério para saber se se trata ou não de um dever. Pelo contrário, permite ele justificar todo o comportamento injusto ou imoral. A mais rigorosa fórmula kantiana, a da capacidade de uma ação ser representada como máxima universal, introduz decerto a representação mais concreta de uma situação de fato mas não tem para si nenhum princípio novo, outro que não seja aquela ausência de contradição e a identidade formal” (HEGEL, 1997, pp119-120).

<sup>2</sup> Neste trabalho, o princípio supremo da moralidade chamado por Kant de “Imperativo Categórico” será escrito com iniciais em letra maiúscula. Por outro lado, os princípios testados pelo Imperativo Categórico devem ser chamados de imperativos categóricos com iniciais em letra minúscula.

<sup>3</sup> Ver sistema de abreviações das obras de Kant no final do texto.

<sup>4</sup> Segundo Kant, “as três maneiras indicadas de apresentar o princípio da moralidade são no fundo apenas outras tantas fórmulas da mesma lei, cada uma das quais reúne em si, por si mesma, as outras duas. Há, contudo, entre elas uma diferença, que na verdade é mais subjetiva do que objetivamente prática, para aproximar a idéia da razão mais e mais da intuição (segundo uma certa analogia) e assim do sentimento. Todas as máximas têm: 1. uma forma, que consiste na universalidade, e sob esse ponto de vista a fórmula do imperativo moral exprime-se de maneira que as máximas têm de ser escolhidas como se devessem valer como leis universais da natureza; 2. uma matéria, isto é, um fim, e então a fórmula diz: o ser racional, como fim segundo a sua natureza, portanto, como fim em si mesmo, tem de servir a toda a máxima de condição restritiva de todos os fins meramente relativos e arbitrários; 3. uma determinação completa de todas as máximas por meio daquela fórmula, a saber, que todas as máximas por legislação própria, devem concordar com a idéia de um reino possível dos fins como um reino da natureza” (GMS, AA 04:436).

<sup>5</sup> Para Kant, os fins que também são deveres "são a *própria perfeição do indivíduo e a felicidade dos outros*" (TL, AA 06: 385). Kant cita somente a própria perfeição da pessoa e não a felicidade dela mesma porque a felicidade própria é desejada naturalmente, não enquanto dever. Do mesmo modo, ele menciona só a felicidade dos outros e não a perfeição dos outros porque a perfeição dos outros consiste na perfeição da moralidade e isso é algo que cada um deve aperfeiçoar por conta própria.

<sup>6</sup> Outras considerações precisam ser feitas para interpretar de maneira correta esse dever. Kant defende que mutilar uma pessoa tendo em vista a venda de órgãos é proibido pela restrição geral de se autodestruir. Mas, a amputação de um membro comprometido do corpo que visa salvar a vida de uma pessoa não pode ser proibida (TL, AA 06: 423). O primeiro caso é proibido porque alguém tem que sacrificar parte do próprio corpo e, conseqüentemente, a sua liberdade de ação futura. Cometer suicídio meramente para escapar da dor e do sofrimento é proibido. Isso não parece ser o caso de alguém que intencionalmente prejudica sua própria vida em defesa de sua pátria. Para Kant, é possível que alguém desista de sua existência e de sua liberdade a fim de salvar a liberdade de outras pessoas. Sobre isso, conferir GUYER, 2005, pp. 243-274.

# SOCIABILIDADE E DIREITO NA OBRA A IDÉIA DE UMA HISTÓRIA UNIVERSAL DE UM PONTO DE VISTA COSMOPOLITA

REJANE SCHAEFER KALSING

*IFC - Campus Sombrio*

rejane.kalsing@yahoo.com.br

## 1. Introdução

Na *Introdução* da obra **Idéia de uma história universal de um ponto de vista cosmopolita**, diz Kant que, de forma geral, os seres humanos não agem apenas de modo instintivo, como os animais, e nem também como “*razoáveis cidadãos do mundo (vernünftige Weltbürger)*”.<sup>1</sup> Daí se seguiria que um filósofo não pode pressupor nos seres humanos, tomados conjuntamente, nenhum “**propósito racional próprio**”.<sup>2</sup> Dessa forma, Kant quer descobrir “um **propósito da natureza**”<sup>3</sup> para a espécie humana, o qual “possibilite uma história segundo determinado plano da natureza para criaturas que procedem sem um plano próprio”.<sup>4</sup> Descobrir tal propósito é, então, um dos objetivos de *IaG*, a qual se desenvolve na forma de **proposições (Satz)**<sup>5</sup>, as *proposições sobre a história*.

## 2. O conceito de sociabilidade: a *insociável sociabilidade*

A **Primeira proposição** reza “todas as disposições naturais de uma criatura estão destinadas a um dia se desenvolver completamente e conforme a um fim”.<sup>6,7</sup> De acordo com isso, pergunta-se, *como* acontecerá esse desenvolvimento na espécie humana? A resposta está na **Quarta proposição**:

*o meio (Das Mittel)* de que a natureza se serve para realizar o desenvolvimento de todas as suas disposições é o **antagonismo** das mesmas na sociedade, na medida em que ele se torna ao fim a causa de uma ordem regulada por leis desta sociedade.<sup>8</sup>

Existe um antagonismo entre as disposições naturais da espécie humana, isto é, um conflito entre as diferentes tendências do ser humano e destas com as dos demais seres humanos. Mas é justamente ele quem promoverá o desenvolvimento completo das mesmas. Mas que antagonismo é esse propriamente? Responde Kant

eu entendo aqui por antagonismo (*Antagonism*) a *insociável sociabilidade (ungesellige Geselligkeit)* dos homens, ou seja, a tendência (*Hang*) dos

mesmos a entrar em sociedade que está ligada a uma oposição (*Widerstand*) geral que ameaça constantemente dissolver essa sociedade.<sup>9</sup>

A *sociabilidade* consiste numa *propensão* à sociedade.<sup>10</sup> É uma tendência a entrar em sociedade mas que, em *IaG* está ligada a uma oposição, ou seja, uma resistência a essa sociedade, que a coloca em risco de forma constante.

Há que se destacar que os seres humanos possuem a tendência a entrar em sociedade, ou seja, possuem *sociabilidade*. Kant não diz que a espécie humana é insociável, não, ela é sociável. Porém, essa tendência é acrescentada de uma *qualidade*, é uma sociabilidade *insociável*. Pois ao mesmo tempo que essa espécie possui a *tendência* a entrar em sociedade possui também uma oposição a ela, posição que, por ser geral, a coloca em risco continuamente.

Talvez se possa dizer, portanto, que a *insociável sociabilidade* consiste num conflito entre disposições contrárias de cada ser humano e também dessas entre os seres humanos. Esse é o *antagonismo* presente na espécie humana. E é ele quem promoverá o desenvolvimento completo das suas disposições naturais, conforme Kant.

Essa *disposição (Anlage)*, continua a **Quarta Proposição**, seria evidente na natureza humana<sup>11</sup> porque o ser humano tem “uma inclinação (*Neigung*) para *associar-se* porque se sente mais como homem [*Mensch*] num tal estado, pelo desenvolvimento de suas disposições naturais”.<sup>12</sup> Mas, por outro lado, “ele também tem uma forte tendência (*Hang*) a separar-se (isolar-se)”<sup>13</sup> por ter uma qualidade insociável que o leva querer fazer tudo em seu proveito, o que provocará “oposição de todos os lados”,<sup>14</sup> da mesma forma como fará oposição por supor que os demais também agirão *em causa própria*.

O ser humano tem uma inclinação a se *associar* porque se *sentiria melhor, mais* como ser humano em sociedade, em função de desenvolvimento de suas disposições naturais. Todavia, ele tem também uma tendência a se separar, porque essa sociabilidade é insociável, o que o leva a querer agir sempre de modo a obter *proveito próprio* e, assim, espera reação de todos os demais. Além disso, reagirá na mesma medida quando perceber um igual comportamento de qualquer outro, pois *julgando os outros como a si mesmo*, supõe que farão o mesmo.

No entanto, é a insociabilidade quem estimula a espécie humana a se desenvolver, pois é ela quem desperta as capacidades dos seres humanos, conforme *IaG*. E, assim, a *insociável sociabilidade*, é favorável na medida em que desperta “todas as forças do homem, [e] o leva a superar a sua tendência à preguiça”.<sup>15</sup> E, a partir daí, “dão-se então os primeiros verdadeiros passos que levarão da rudeza à cultura”.<sup>16</sup>

### 1.3 O conceito de *Direito*

A partir da **Quinta Proposição** aparece mais a ideia do *direito*. Examinar a abordagem do *direito* da **Quinta** à **Nona Proposição**, ou seja, até o final da obra, e em que sentido a sociabilidade nessa obra pode estar ligada ao direito são os objetivos a partir de agora.

Nesse sentido, iniciar-se-á por destacar uma extensa passagem da **Quinta Proposição**. Nela Kant assevera que

**o maior problema** (*Das größte Problem*) **para a espécie humana, a cuja solução a natureza a obriga** (*zwingt*), **é alcançar** uma sociedade civil **que administre universalmente o direito**. Como *somente em sociedade* e a rigor naquela que admite a máxima liberdade de seus membros e, conseqüentemente, um antagonismo geral de seus membros e, portanto, a mais precisa determinação e resguardo dos limites desta liberdade – de modo a poder coexistir com a liberdade dos outros; como somente nela *o mais alto propósito da natureza* (*die höchste Absicht der Natur*), ou seja, o desenvolvimento de todas as suas disposições, pode ser alcançado pela humanidade, a natureza quer que a humanidade proporcione a si mesma este propósito, como todos os outros fins de sua destinação: assim uma sociedade na qual a *liberdade sob leis exteriores* encontra-se ligada no mais alto grau a um poder irresistível, ou seja, uma **constituição civil** perfeitamente **justa**, deve ser a *mais elevada tarefa da natureza* (*die höchste Aufgabe der Natur*) para a espécie humana, porque a natureza somente pode alcançar seus outros propósitos relativamente à nossa espécie por meio da solução e cumprimento daquela *tarefa*.<sup>17, 18</sup>

O alcance de uma sociedade civil que administre o direito de forma universal é o maior *problema* para a humanidade resolver. E é a natureza quem impele a humanidade a chegar a tal solução. É que somente nessa sociedade, entende Kant, *o mais alto propósito da natureza*, que é o desenvolvimento de todas as suas disposições, pode ser alcançado pela humanidade.

Portanto, é somente *em sociedade* que o mais alto propósito da natureza para a humanidade poderá ser alcançado. Entretanto, não é em qualquer sociedade que ele poderá ser alcançado e sim apenas em uma sociedade que permitir, em primeiro lugar, a máxima liberdade e, como consequência, um antagonismo geral de seus membros e, então, possuir a mais precisa determinação e resguardo dos limites dessa liberdade; portanto, numa sociedade civil que administre o direito de forma universal.

Quer dizer, é somente naquela sociedade que, para proporcionar a máxima liberdade e, com isso, um antagonismo geral de seus membros, possui os limites dessa liberdade ao mesmo tempo determinados da forma mais precisa possível mas também garantidos, assegurados, de modo a liberdade de um possa coexistir com a liberdade dos outros; é nessa sociedade que o desenvolvimento de todas as disposições da humanidade pode ser alcançado. Propósito este que a natureza quer que a humanidade proporcione a si mesma.

Desse modo, entende Kant que, uma sociedade na qual a liberdade esteja sob leis exteriores e se encontra ligada a uma constituição civil perfeitamente justa tem de ser a mais elevada tarefa da natureza para a humanidade, porque a natureza somente pode alcançar os outros propósitos em relação à espécie humana através da solução e do cumprimento daquela tarefa.

Um aspecto que merece destaque é que uma coisa é o propósito, o mais alto por sinal, da natureza para a espécie humana, outra é a tarefa, a mais elevada

também, da natureza para essa espécie. O mais alto *propósito* é o desenvolvimento de todas as disposições humanas. Já a mais elevada *tarefa* é uma sociedade em que a liberdade esteja sob leis exteriores e ligada a uma constituição civil perfeitamente justa. E é somente com o cumprimento dessa tarefa que o propósito poderá ser alcançado.

Só para referir, a **Sétima**, a **Oitava** e a **Nona Proposição** reiteram a **Quinta Proposição** quando dizem, a primeira, que é numa constituição civil que todos os germes (*Keime*) da humanidade podem ser desenvolvidos.<sup>19</sup> Já a **Oitava** considera “um *estado cosmopolita universal (allgemeiner weltbürgerlicher Zustand)*, como o *seio (Schoß)* no qual podem se desenvolver todas as disposições originais da espécie humana”.<sup>20</sup> E a última destaca que é numa constituição civil, o “estado no qual todos os germes (*Keime*) que a natureza nela colocou poderão desenvolver-se plenamente”.<sup>21, 22</sup> Esse é o estado no qual todas as disposições da espécie humana poderão se desenvolver plenamente.

Na passagem que encerra a **Quinta Proposição**, Kant retoma o tema da insociabilidade. Ele diz

toda cultura e toda arte que ornamentam a humanidade, a mais bela ordem social são frutos da *insociabilidade*, que por si mesma é obrigada a se disciplinar e, assim, por meio de um artifício imposto, a desenvolver completamente os *germes* da natureza.<sup>23</sup>

Passando-se para a **Sétima Proposição**, a respeito da *instauração* de uma *constituição civil perfeitamente justa*, Kant observa que “**o problema do estabelecimento de uma constituição civil perfeita depende do problema da relação externa legal entre Estados, e não pode ser resolvido sem que este último o seja**”.<sup>24</sup>

Como se vê, Kant não separa a questão do estabelecimento de uma constituição civil perfeita em um determinado Estado da questão da relação externa legal entre Estados. São questões interligadas. Um Estado não conseguirá estabelecer uma constituição civil perfeita sem resolver o problema da relação externa legal com os outros Estados.

Portanto, o cumprimento da mais elevada tarefa para que espécie humana, isto é, o estabelecimento de uma constituição civil perfeita, depende diretamente da solução do problema da *relação externa legal* entre os Estados<sup>25</sup>.

A conclusão da **Sétima Proposição** diz que

a mesma *insociabilidade* que obrigou os homens a esta tarefa é novamente a causa (*Ursache*) de que cada república (*Gemeinwesen*), em suas relações externas – ou seja, como um Estado em relação a outros Estados –, esteja numa liberdade irrestrita, e conseqüentemente deva esperar do outro os mesmos males que oprimiam os indivíduos e os obrigavam a entrar num *estado civil conforme leis*.<sup>26, 27</sup>

A natureza teria se servido outra vez da qualidade insociável presente nos seres humanos, “mesmo entre as grandes sociedades e corpos políticos desta

espécie de criatura, como um meio para encontrar, no seu inevitável *antagonismo*, um estado de tranquilidade e segurança".<sup>28</sup> Quer dizer, a mesma qualidade insociável existente nos indivíduos se verifica entre os Estados. Ela é novamente a causa de cada república em suas relações externas, ou de um Estado em relação a outros, estar numa liberdade irrestrita. E, sendo assim, um Estado espera do outro a mesma oposição e os mesmos problemas que obrigaram os indivíduos a entrar num estado civil conforme leis. Novamente é a insociável sociabilidade que conduzirá, neste caso os Estados, a um estado de tranquilidade e segurança.

Parece estar dito aí que o antagonismo presente na espécie humana, a insociável sociabilidade, se encontra assim tanto no indivíduo quanto nos Estados. Pois é ela que obriga tanto os indivíduos a entrar em um *estado civil conforme leis* e como os Estados a entrar numa *federação de nações*. Como é a insociável sociabilidade que obrigará os indivíduos a entrar num estado civil conforme leis e conduzirá os Estados a entrar numa federação de nações, há aí uma ligação entre sociabilidade e direito.

Já a **Oitava Proposição** observa que se pode

considerar a história da espécie humana, em seu conjunto, como a realização de um *plano oculto da natureza* (*verborgen Plans der Natur*) para estabelecer uma *constituição política* (*Staatsverfassung*) perfeita interiormente e, quanto a esse fim, também exteriormente perfeita, como o *único estado no qual a natureza pode desenvolver plenamente, na humanidade, todas as suas disposições*.<sup>29</sup>

Não se quer aqui discutir o que pode ser propriamente esse **plano secreto da natureza**, secreto num certo sentido porque os indivíduos o desconhecem,<sup>30</sup> entre outras questões, mas apenas destacar que é uma constituição política perfeita interior e exteriormente o único estado no qual a natureza pode plenamente desenvolver todas as disposições da humanidade.

#### 1.4 Insociável sociabilidade: uma sociabilidade legal?

Em **IaG** a sociabilidade ocorre sob a forma de uma *sociabilidade* que é *insociável*. Essa *sociabilidade*, como já visto, obrigará os indivíduos a entrar num *estado civil conforme leis* e conduzirá os Estados a entrar numa *federação de nações*,<sup>31</sup> numa palavra, ela conduzirá a uma constituição política perfeita interior e exteriormente. Tanto o primeiro estado quanto o segundo, e também o último, referem-se ao direito, e, como é a sociabilidade que conduzirá a eles, assim, depreende-se daí que há uma ligação entre sociabilidade e direito.

E, se ela conduz ao direito, se houver uma ligação entre direito e sociabilidade, talvez essa sociabilidade possa ser considerada como uma *sociabilidade legal*. A qual, segundo Valerio Rohden, pode ser entendida enquanto capacidade e disposição a se deixar guiar por princípios do Direito,<sup>32</sup> pois sem tal capacidade e a disposição, tanto um estado civil conforme leis quanto uma federação de nações nem poderiam ser instituídos.

Há que se destacar, em primeiro lugar, que esse conceito não ocorre na obra **IaG** e sim seis anos mais tarde (1790), a saber, na **Crítica do juízo**, a **KU**. A (única) passagem em que ocorre refere-se ao momento da constituição de uma sociedade, ou mais propriamente, de uma *coletividade duradoura*. Ela reza propriamente que

a época e os povos, nos quais o ativo impulso à *sociabilidade legal*, pela qual um povo constitui uma coletividade duradoura, lutou com as grandes dificuldades que envolvem a difícil tarefa de unir liberdade (e, portanto, também igualdade) à coerção (mais do respeito e da submissão por dever do que por medo).<sup>33</sup>

Sem intentar aqui desenvolver tal passagem, mas pode-se destacar, pelo menos, que é pela sociabilidade legal que um povo constitui uma coletividade duradoura e que ela, ou mais propriamente, o ativo impulso a ela, teve de lutar com grandes dificuldades para unir liberdade, e igualdade, a uma coerção mais no sentido de respeito e de submissão por dever do que por medo.

A esse respeito, Valerio Rohden, no artigo intitulado **Sociabilidade legal uma ligação entre direito e humanidade na 3ª Crítica de Kant**, afirma que

o homem não é originariamente social, mas sociável: é graças à sua *capacidade de impor-se uma obrigação racional que o homem é capaz de sociedade*. [...] Pela capacidade de obrigar-se, responsabilizar-se, ser cidadão, o homem tem condições de vida coletiva. [...] Sem essa capacidade e disposição de deixar-se guiar por eles, a sociedade é impossível. Com isso fica claro o sentido da *sociabilidade legal*: ela é a *capacidade e disposição a deixar-se guiar por princípios do Direito*.<sup>34</sup>

Segundo Rohden, o *sentido* da sociabilidade legal, que pode ser apreendido do texto de Kant, já que este propriamente não o desenvolve, é o da capacidade e disposição a deixar-se guiar por princípios do direito. Em outras palavras, ela é a capacidade de se impor uma obrigação racional, capacidade através da qual o ser humano é capaz de sociedade, de outra forma ainda, é a capacidade de se obrigar, de se responsabilizar, de ser cidadão, capacidade pela qual o ser humano tem condições de vida coletiva.

Como em **IaG** a sociabilidade parece conduzir ao direito, porque conduz a um *estado civil conforme leis*, a uma *federação de nações* e, mais propriamente a uma constituição política perfeita interior e exteriormente, parece possível daí depreender que essa sociabilidade é uma sociabilidade legal. Num certo sentido, ao menos, para entrar em um *estado civil conforme leis*, em uma *federação de nações* e para estar sob uma constituição política perfeita interior e exteriormente é necessária a capacidade de se guiar por princípios do direito e estar disposto a se deixar guiar por eles, quer dizer, por “princípios de convivência”,<sup>35</sup> sem essa capacidade a sociedade, e assim, o estado civil conforme leis, a federação de nações, a constituição política perfeita interior e exteriormente são impossíveis.

Apesar de o conceito de sociabilidade legal não ocorrer propriamente em **IaG**, ele parece encontrar-se aí, já que somente sob as condições de uma

*constituição política* é que as todas disposições humanas se desenvolverão e, se se concordar com Rohden que a sociabilidade legal é um pré-requisito para uma *constituição política*, logo, poderia-se depreender que tal conceito está aí implicado.

Foi o que se procurou mostrar.

## Referências

KANT, Immanuel. **Idéia de uma história universal de um ponto de vista cosmopolita**. (tradução de Rodrigo Naves e Ricardo Terra). São Paulo: Brasiliense, 1986, (Coleção Elogio da Filosofia). **IaG**, AA 08.

KANT, Immanuel. **Crítica da faculdade do juízo**. Rio de Janeiro: Forense Universitária, 2002. **KU**, AA 05.

CHIODI, P. *La filosofia kantiana della storia*. **Rivista di Filosofia**, (58), 1967.

GIANOTTI, José Arthur. Kant e o espaço da história universal. In: **Idéia de uma história universal de um ponto de vista cosmopolita**. (tradução de Rodrigo Naves e Ricardo Terra). São Paulo: Brasiliense, 1986, (Coleção Elogio da Filosofia).

ROHDEN, Valerio. Sociabilidade legal uma ligação entre direito e humanidade na 3ª Crítica de Kant. **Analytica**, Rio de Janeiro, v. 1, n. 2, 1994.

## Notas

<sup>1</sup> **IaG**, AA 08: 17. KANT, 1986, p. 10. *Itálicos acrescentados*.

<sup>2</sup> **IaG**, AA 08: 18. KANT, 1986, p. 10. **Negritos de Kant**.

<sup>3</sup> **IaG**, AA 08: 18. KANT, 1986, p. 10. **Negritos de Kant**.

<sup>4</sup> **IaG**, AA 08: 18: 06-12. KANT, 1986, p. 10.

<sup>5</sup> Conforme **IaG**, AA 08: 18. KANT, 1986, p. 11.

<sup>6</sup> **IaG**, AA 08: 18. KANT, 1986, p. 11. **Negritos de Kant**. Isso está de acordo com a *doutrina teleológica da natureza*, pois, do contrário não se terá uma natureza regida por leis e sim “um jogo sem finalidade” (**IaG**, AA 08: 18. KANT, 1986, p. 11) da mesma (*eine zwecklos spielende Natur*) (Conforme **IaG**, AA 08: 18: 25-26. Conforme KANT, 1986, p. 11.).

<sup>7</sup> No ser humano, porém, conforme a **Segunda Proposição**, “aquelas disposições naturais que estão voltadas para o uso de sua razão devem desenvolver-se completamente apenas na espécie e não no indivíduo” (KANT, 1986, p. 11). É, que, segundo Kant, cada ser humano precisaria de uma vida muito longa para aprender a usar plenamente todas as suas disposições naturais e, como a natureza só lhe concedeu uma vida curta, “ela necessita de uma série talvez indefinida de gerações” (**IaG**, AA 08: 19.07. KANT, 1986, p. 11) [que possam transmitir umas às outras os seus avanços para assim conduzir, na espécie, “o germe da natureza àquele grau de desenvolvimento que é completamente adequado ao seu propósito” (**IaG**, AA 08: 19. 09-10. KANT, 1986, p. 11)]. Portanto, o desenvolvimento completo das disposições naturais do ser humano só se dará na espécie.]

<sup>8</sup> *IaG*, AA 08: 020. 26-28. KANT, 1986, p. 13. **Negritos de Kant. Itálicos acrescentados.**

<sup>9</sup> *IaG*, AA 08: 020. 30-33. KANT, 1986, p. 13.

<sup>10</sup> Conforme *KU*, AA 05:296. *KU*, B 162.

<sup>11</sup> Conforme *IaG*, AA 08: 020. Conforme KANT, 1986, p. 13.

<sup>12</sup> *IaG*, AA 08: 020-021. KANT, 1986, p. 13.

<sup>13</sup> *IaG*, AA 08: 021. KANT, 1986, p. 13.

<sup>14</sup> *IaG*, AA 08: 021. KANT, 1986, p. 13.

<sup>15</sup> *IaG*, AA 08: 021. KANT, 1986, p. 13.

<sup>16</sup> *IaG*, AA 08: 021. KANT, 1986, p. 13.

<sup>17</sup> *IaG*, AA 08: 022.06-21. KANT, 1986, p. 14-15. **Negritos de Kant. Itálicos acrescentados.**

<sup>18</sup> Tanto a **Sétima** quanto a **Nona Proposição** reiteram tal coisa quando dizem que é numa constituição civil que todos os germes (*Keime*) da humanidade podem ser desenvolvidos (Conforme *IaG*, AA 08: 025. 36-37. KANT, 1986, p. 18.) e que é numa constituição civil, o “estado no qual todos os germes (*Keime*) que a natureza nela colocou poderão desenvolver-se plenamente” (*IaG*, AA 08: 030.16-18. KANT, 1986, p. 23.).

<sup>19</sup> Conforme *IaG*, AA 08: 025. 36-37. KANT, 1986, p. 18.

<sup>20</sup> *IaG*, AA 08: 028.34-37. KANT, 1986, p. 22.

<sup>21</sup> *IaG*, AA 08: 030.16-18. KANT, 1986, p. 23.

<sup>22</sup> A título de comentário, a obra **Sobre o dito comum: isto pode ser correto na teoria mas não serve para a prática** (*Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis*) (*TP*) ratifica isso. Em seu **Capítulo Terceiro**, intitulado “*Da relação da teoria à prática no direito dos povos, considerada do ponto de vista filantrópico-universal, isto é, cosmopolita*”, em uma nota e que explica justamente esse título, Kant afirma que “não se vê de imediato como é que uma pressuposição **filantrópica**-universal aponta para *uma constituição cosmopolita* e como esta, por sua vez, *aponta para uma fundação de um direito dos povos, enquanto o único estado em que as disposições da humanidade que tornam a nossa espécie digna de amor podem ser desenvolvidas de um modo conveniente*. – A conclusão desta terceira secção evidenciará isto”. *TP*, AA 08: 307 (tradução de Christian Hamm, não publicada). **Negritos de Kant. Itálicos acrescentados.**

<sup>23</sup> *IaG*, AA 08: 022. 32-35. KANT, 1986, p. 15. *Itálicos acrescentados.*

<sup>24</sup> *IaG*, AA 08: 024.02-05. KANT, 1986, p. 16. **Negritos de Kant.**

<sup>25</sup> Não se dissertará aqui a respeito *da relação externa legal entre Estados*, entre outras questões que poderiam daí decorrer, pois tal exame fugiria do objetivo do presente texto ou, ao menos, desviaria o foco deste em direção ao problema do *Estado* em Kant propriamente, e outras questões a este pertinentes, questões importantes em Kant, mas que não fazem parte deste.

<sup>26</sup> *IaG*, AA 08: 024.07-12. KANT, 1986, p. 16-17. *Itálicos acrescentados.*

<sup>27</sup> Tal passagem é ratificada na obra *TP*. Nesta Kant afirma que “do mesmo modo que, enfim, a violência universal e a miséria dela decorrente levaram necessariamente um povo à resolução de se submeter à coação que a própria *razão* lhe prescreve como meio, a saber, à lei pública, e a entrar numa constituição **civil**, assim também a miséria resultante das guerras permanentes, em que os Estados, por sua vez, procuram enfraquecer ou subjugar uns aos outros, deve finalmente levá-los, mesmo contra vontade, a entrar numa constituição **cosmopolita**”. *Gemeinspruch*, AA 08: 310. **Negritos de Kant. Itálicos acrescentados.**

<sup>28</sup> *IaG*, AA 08: 024.13-16. KANT, 1986, p. 17. *Itálicos acrescentados.*

<sup>29</sup> *IaG*, AA 08: 027.02-07. KANT, 1986, p. 20. *Itálicos acrescentados.*

<sup>30</sup> Conforme *IaG*, AA 08: 17. KANT, 1986, p. 10.

<sup>31</sup> Conforme *IaG*, AA 08: 024. KANT, 1986, p. 17. *Itálicos acrescentados.*

<sup>32</sup> Conforme ROHDEN, Valerio. Sociabilidade legal uma ligação entre direito e humanidade na 3ª Crítica de Kant. *Analytica*, Rio de Janeiro, v. 1, n. 2, 1994, p. 104.

<sup>33</sup> *KU*, AA 05: 355. 32-36. KANT, Immanuel. **Crítica da faculdade do juízo**. Rio de Janeiro: Forense Universitária, 2002, *KU*, B 262-263. **Negritos de Kant. Itálicos acrescentados.**

<sup>34</sup> ROHDEN, 1994, p. 104.

<sup>35</sup> ROHDEN, 1994, p. 104.