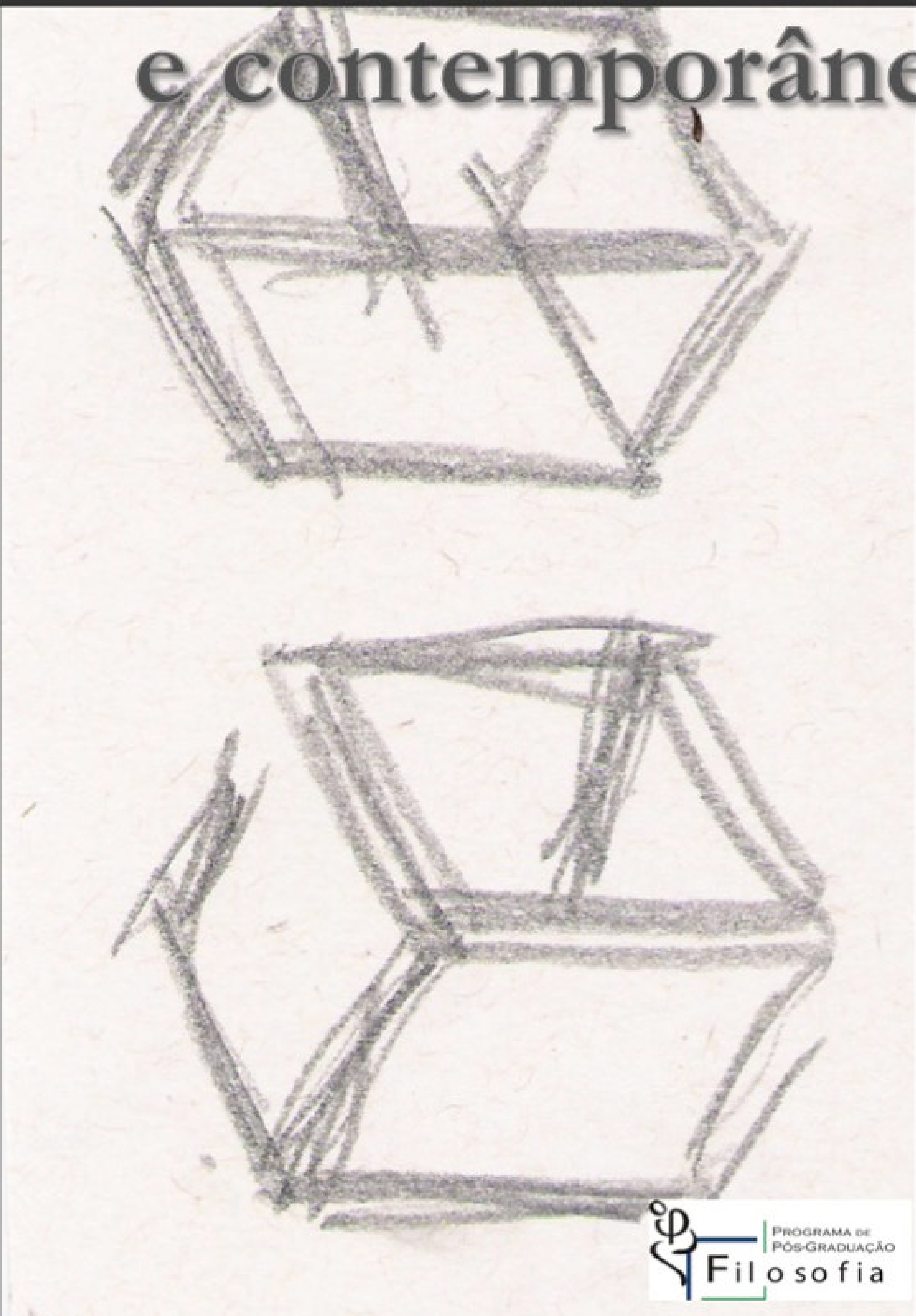
Revista de Filosofia moderna e contemporânea



Brasília, v. 6, n. 2, ago/dez 2018

Programa de Pós-Graduação Filosofia

UnB

revista de Filosofia moderna e contemporânea

A *Revista de Filosofia Moderna e Contemporânea* é uma publicação semestral do Departamento de Filosofia (FIL) e do Programa de Pós-graduação em Filosofia (PPGFIL) da Universidade de Brasília (UnB)

The *Journal of Modern and Contemporary Philosophy* is a semiannual publication of the Department of Philosophy and the Graduate Program in Philosophy of the University of Brasília (UnB)

Brasília, volume 6, número 2, 2º semestre de 2018

ISSN 2317-9570 (publicação eletrônica)

http://periodicos.unb.br/index.php/fmc/index

FICHA CATALOGRÁFICA

R454 Revista de Filosofia Moderna e Contemporânea / Universidade de Brasília, Instituto de Ciências Humanas, Departamento de Filosofia, Programa de Pós-graduação em Filosofia. / Alexandre Hahn, editor. – v. 5. n. 1 (jan./jun. 2017). – Brasília: Universidade de Brasília, 2013-.

Periodicidade semestral
Endereço eletrônico: http://periodicos.unb.br/index.php/fmc
Descrição baseada em: v. 5, n. 1 (jan./jun. 2017).
ISSN 2317-9570 (versão online)

1. Filosofia – Periódicos 2. Filosofia Moderna – História 3. Filosofia Contemporânea – História I. Departamento de Filosofia II. Programa de Pós-graduação em Filosofia

CDD - 109
CDU - 19

Contato
Revista de Filosofia Moderna e Contemporânea
Universidade de Brasília
Departamento de Filosofia
Programa de Pós-Graduação em Filosofia
ICC Ala Norte - Bloco B - 1º Andar (B1 624)
Telefone: (61) 3107-6677 / 6680 / 6679
Campus Universitário Darcy Ribeiro
70910-900 – Brasília – Distrito Federal – Brasil
E-mail: rfmc@unb.br

Indexação
CiteFactor, CLASE, Diadorim, DOAJ, DRJI, Latindex, Sumários, Periódicos, PhilBrasil, Phipapers, SHERPA/RoMEO

Qualis CAPES – B2

DOI: https://doi.org/10.26512/rfmc.v6i2

Sumário

revista de
Filosofia moderna
e contemporânea

Editorial

Apresentação do Dossiê

O *Colóquio UnB-USP de Lógica e Filosofia da Lógica* teve sua primeira edição em janeiro de 2016 como uma das atividades regulares do grupo de pesquisa *Lógica Filosófica e Filosofia da Lógica* (CNPq) cujos membros permanentes são professores de Lógica dos departamentos de Filosofia da *Universidade de Brasília* e *Universidade de São Paulo*.

A lógica combina métodos matemáticos e conceitos filosóficos, e a ideia do colóquio é refletir esse perfil. O evento tem como característica principal a realização de palestras longas que são em geral conduzidas por palestrantes que utilizam apenas giz e quadro. A organização do colóquio é realizada ora em Brasília, ora em São Paulo, com membros ligados aos dois centros. Sempre ocorre no mês de janeiro, em duas tardes consecutivas.

O presente número da *Revista de Filosofia Moderna e Contemporânea* (UnB) contém um dossiê dedicado à publicação de artigos escritos por participantes de edições prévias do colóquio, não retratando uma edição específica. Agradecemos ao professor Alexandre Hahn, editor deste periódico, por hospedar o dossiê e aos autores que contribuíram com seus artigos.

Alexandre Costa-Leite e Rodrigo Freire

(Organizadores do *Dossiê)*

* * *

Além dos trabalhos que compõem o *Dossiê*, o presente número também conta com outras contribuições recebidas em fluxo contínuo.

(1) Renato Kinouchi, professor adjunto da UFABC, discute, em seu artigo *Sobre as limitações do dilema do bonde para a avaliação dos riscos impostos por veículos autônomos*, a adequação da versão original desse dilema aos potenciais acidentes com veículos autônomos. Critica essa versão, por não incluir questões relativas aos riscos e incertezas envolvidos em tais acidentes, e propõe um modelo abstrato (probabilístico) que satisfaria os requisitos conceituais da noção padrão de risco.

(2) Em seu *Uma crítica popperiana ao raciocínio sociológico de Jean-Claude Passeron*, Sergio Tarbes, mestre em filosofia pela UnB, busca analisar os fundamentos empíricos do "raciocínio sociológico" baseado na utilização de linguagem natural e proposto por Passeron, empregando um método inspirado nas ideias de Popper.

(3) Ronaldo Manzi, psicanalista e doutor em filosofia pela USP, retoma, no artigo *Relendo* O Processo *de Kafka como a manifestação de uma patologia social*, a leitura de Agamben da referida obra, focada na suposta autocalúnia do personagem Joseph K. como redenção da culpa perante a lei. Propõe interpretação, em que busca mostrar que é o social que faz com que K. assuma a culpa, não havendo autocalúnia como entendia Agamben.

(4) *Schelling e a questão dos postulados práticos em "Cartas Filosóficas sobre o Dogmatismo e Criticismo"*, artigo de Marília Cota Pacheco, professora substituta do Departamento de Filosofia da UnB, apresenta a noção kantiana de incondicionado como uma das "raízes" da dialética da imaginação de Schelling. Sua intenção é mostrar que a resposta ao diagnóstico deste último, acerca da incapacidade do dogmatismo e do criticismo demonstrarem o processo pelo qual o múltiplo surge de uma unidade absoluta, é um postulado prático.

(5) Por fim, Marília Cota Pacheco nos apresenta a resenha da obra *Interpreting Schelling: Critical Essays*, organizada por Lara Osteric.

Gostaríamos de aproveitar o ensejo para agradecer a todos os autores, por terem honrado a nossa *Revista* com as suas produções, bem como aos membros do corpo editorial, avaliadores, editores e leitores de provas, pela fundamental colaboração na confecção da presente edição.

Os Editores

ISSN: 2317-9570

Mathematical Concepts as Natural Kinds

[Conceitos Matemáticos como Tipos Naturais]

Daniel Arvage Nagase*

Abstract: This paper presents an approach to mathematical terms similar to the approach developed by Kripke in order to deal with natural kind terms; in fact, I argue that mathematical terms *are* natural kind terms in the sense of Kripke. Thus, I suggest that, from a semantic perspective, such terms should be seen as primarily referring terms, and, from a metaphysical perspective, that good definitions of such terms should embody structural information about the exemplars from the kind in question.
Keywords: mathematical definitions, natural kinds, essence, structure

Resumo: Este artigo apresenta uma abordagem de termos matemáticos semelhante à abordagem desenvolvida por Kripke para lidar com termos de tipo natural; na verdade, eu defendo que termos matemáticos *são* termos de tipo natural no sentido de Kripke. Assim, sugiro que, a partir de uma perspectiva semântica, tais termos devem ser vistos principalmente como termos referentes, e, de uma perspectiva metafísica, que boas definições de tais termos devem incorporar informações estruturais sobre os exemplares do tipo em questão.
Palavras-chave: definições matemáticas, tipos naturais, essência, estrutura

There are at least two approaches to natural kinds: first, we can approach them from a *semantic* point of view, studying the semantics of terms which denote natural kinds. Second, we can approach them from a *metaphysical* point of view, investigating the *structure* of the exemplars of a given kind. In this paper, I will make some suggestions about how some of Kripke's ideas about natural kinds can be adapted to the mathematical case. I will start with the semantic approach and then use some of the issues arising from it to motivate the metaphysical approach as well. I qualify my remarks here as suggestions, instead of arguments, to indicate their provisional nature; much more needs to be done about this subject, and I will finish by pointing in what seems to me to be promising directions for further study.

*Doutorando em filosofia pela Universidade de São Paulo (USP). E-mail: dan.nagase@gmail.com.

DOI: https://doi.org/10.26512/rfmc.v6i2.19014

Semantics

Let us start then with a quick review of Kripke's ideas regarding the semantics of natural kind terms. Following traditional usage, I call the *semantic value* of a term the term's contribution to the determination of the truth value of the sentences in which it occurs. Before Kripke's work, it was common to identify the semantic value of a natural kind term with a description, or a cluster of descriptions, satisfied only by the extension of the term in question. In other words, it was common to identify the semantic value of a natural kind term with what could be called its *descriptive content*, how the term is used to describe a given feature of reality.

Kripke's main thesis, then, is that the semantic value of a natural kind term is *not* its descriptive content, but rather its reference—a natural kind term is used not to *describe* a given feature of reality, but to *name* it. That is, for Kripke, a natural kind term is not synonymous to a description, or cluster of descriptions, which is associated to the term by speakers of the language in which the term occurs. In particular, this means that its reference is not determined by whether some objects satisfy a given description (or cluster of descriptions), but rather by another mechanism. Since this mechanism is important for what follows, let us take a quick look at how Kripke motivates it.

Initially, according to Kripke, a term acquires reference by way of a "baptism ceremony", in which a person declares, for instance, "I will call this stuff by the term 'gold"' or "I will call by 'gold' this yellow, metallic stuff" (in the first case, the reference is fixed by a demonstrative indication, whereas in the second case it is fixed by description). After this "ceremony", the reference of a term is transmitted by a historical chain whose first link is the ceremony: as the speakers of a language intend to use the term *with the same reference as the baptizer*, this reference gets transmitted along the community, even if those employing the term are mistaken with respect to which description was used to fix this reference. That is why, to Kripke, reference takes precedence to description: we use a term first to refer to "this object here" or "to that object referred by (whomever)", and only later we associate a description to the term. Otherwise put, we use a term *because it refers to a given object* (because we want to speak about this object, and not another one), and not because we associate to it a certain description.

As a consequence, we obtain what is usually called the *stability of reference*: the reference of a natural kind term is *stable* across changes in the descriptions associated with the term. For example, perhaps in the past the description associated with the term "gold" was something like "yellow metal", whereas now it is "chemical element with atomic number 79". But this change in the associated description does not alter the semantic value of the term, as we still refer to the same substances as gold.

The above is a very brief sketch of Kripke's views (which I will not defend here), but it will suffice for my purposes. What I want to suggest is that this view can be applied to mathematical terms as well. In particular, I want to suggest that at least some interesting examples of mathematical terms—such as "continuous"—are examples of natural kind terms, in the sense sketched above. Of course, I do not want to suggest that *every* mathematical predicate is a natural kind term (which is absurd). But I do want to suggest that the widespread view according to which most, if not all, mathematical definitions are *stipulations* or *conventions* that determine the semantic value of the defined terms is mistaken. According to this view, the semantic value of a mathematical term would be the result of a *decision* on the part of the mathematical community, in such a way that there would be an arbitrary component in the corresponding concepts. Against this, I want to suggest that the degree of arbitrariness involved in the definition of a mathematical term is generally very low.

Let us consider a typical example, the notion of something being continuous. An idealized version of the history of this notion is as follows: initially, mathematicians isolated an interesting mathematical class, noticing that they all satisfied a property that they called "continuity" (very roughly, these objects were lines or curves that could be drawn without lifting the pencil from the paper). The first characterizations of this class were derived from what we could call its *phenomenal qualities*, such as "being generated by the uninterrupted movement of a point". However, especially in the 19th century, developments tied to the representation of infinite series of functions and the generalization of certain concepts from real to complex analysis led to a deeper investigation of this class of functions. There were at least two goals: first, to obtain a definition of continuity that only used intrinsically mathematical concepts and, second, to obtain a definition that could reveal a deep structural fact that explained why this class of functions is so fruitful. Both of these goals were achieved by the celebrated Weierstrass definition of continuity in terms of "espilons and deltas".

One lesson I want to draw from this narrative is the following. Suppose, for vividness, that one were to show Newton and Weierstrass paradigmatic examples of functions (say, the trigonometric functions, or a given polynomial). It is uncontroversial that they would agree in their classification of certain functions as continuous in the great majority of cases. That is, in the great majority of cases they would agree that the same functions were, or were not, continuous. But according to the descriptivist, who holds that the semantic value of a term is determined by the description associated to it, this agreement would be illusory, since Newton and Weierstrass associate different descriptions to the term "continuous", and hence actually mean different things by it. Indeed, according to this descriptivist, what we have above is *not* the history of the continuity notion, since there is no such a notion, but rather a succession of notions each attached to the term "continuous". That is, the term "continuous" would be *ambiguous*. On the other hand, if we adopt Kripke's views, there is no difficulty in saying that we do have the same notion across our narrative, a notion whose properties were progressively determined by mathematical investigation.

Another lesson to draw from the history of the mathematical notion of continuity is that there is little that is arbitrary in Weierstrass's definition. It is not that Weierstrass *stipulated* that, thereafter, a continuous function would be one that satisfied such and such a property. Indeed, that this was not the case is easily seen by the fact that practically every mathematician involved in the investigation of the foundations of mathematics in the 19th century agreed that a good definition of a continuous function would have to satisfy certain conditions (notably, it would have to imply the intermediate value theorem). Of course, such conditions were merely necessary, but not sufficient (not even jointly sufficient) conditions for a function to be continuous. But they did point to the demand that a good definition of continuity should *explain* why the continuous functions satisfied these conditions.

I cannot resist adding a further example of this idea, this time

ISSN: 2317-9570

a little too contentious. It seems to me that Turing's analysis of computability in terms of Turing machines fits exactly the pattern described above. Mathematicians had long been interested in algorithms, and had already isolated a large class of functions as being algorithmically computable. Some of these were paradigmatic, and Turing's definition captured exactly what made them paradigmatic, by describing the computation process behind such functions. This points to a *third* lesson to be drawn from the above narrative. A mathematical definition is successful *not* when it matches some kind of pretheoretic definition, or intuition, of a given concept. This would be to accept the descriptivist thesis, according to which the definition would merely spell out the descriptive content of the term being defined. Against the descriptivist, I hold that in most cases it may be that there was no "pretheoretic" or "intuitive" description corresponding to the term. So the widespread view, according to which Turing's thesis is "unprovable", because the "intuitive" notion of computability is not rigorous enough to allow one to prove the equivalence between it and Turing's rigorous definition, gets things exactly backwards. There was no intuitive notion of computability, but rather there were *paradigmatic instances*, and the rigorous definition is successful to the extent that it explains what makes such instances paradigmatic, what common characteristic holds together all those objects. Again, this is precisely what Turing did.

Metaphysics

The picture adumbrated above accords well with certain Aristotelian strands in Kripke's description of the scientific endeavor. According to this description, researchers start by grouping together certain objects that share given features (generally phenomenal qualities)—what the Aristotelians call a predicate profile. This predicate profile allows us to infer *that* such objects form a natural kind. In the second phase of the inquiry, one then looks for structural features—the *essence*, according to Kripke—of these objects, features that explain why the objects in question share this predicate profile. It is those structural features that ground the phenomenal qualities of the objects and explain why the initial sample constitutes a kind. Of course, in many cases the phenomenal qualities were misleading, and one is then led to rectify the initial characterization of the kind.

From the metaphysical point of view, it is common to level the charge against Kripke that the notion of essence he employs is somewhat obscure. This is a curious charge, since his notion of essence is perfectly in order—a property is essential to an object if it is necessary that the object has that property. The problem with Kripke's notion of essence is not that it is obscure, but rather that it is inadequate. In our particular mathematical context, this inadequacy is glaring, since every property of a mathematical object is (presumably) necessary, whence every property would be essential to the object. This points to a demand for a new notion of essence that could fulfill the explanatory role indicated above.

To develop such a notion of essence is an enormous task, one that I will excuse myself from discharging here. Instead, I would like to merely indicate some possible ideas for further development. First, it seems to me that one thing we can draw from the above narrative is that this notion of essence would have to be tied to the *structural characteristics* of the object. It is extremely difficult to make this more precise in a non-trivial manner, but some attempts in this direction have been carried out by those who, following Kit Fine's example, are articulating a kind of Aristotelian hylemorphism. The idea here is to develop a rich enough notion of *form* that could distinguish between a mere heap or aggregate from a full-blown integrated object, and use this notion of form to specify the essence of a kind as its formal features. In the mathematical case, these formal features would then explain why the kind in question is natural, by exhibiting a common property that explained why certain theorems hold of the given kind (e.g. Bolzano explicitly introduced his definition of continuity in order to explain the intermediate value theorem). This last aspect is related to some problems in the philosophy of mathematical explanation: an informative proof would be one that exploited an essential feature of a given kind in order to derive a further property of the kind.

Regardless of whether a viable notion of essence can be developed, it seems clear, however, that the search for structural characteristics of mathematical objects is a driving force behind much mathematical research. One area in which related ideas have been very fertile is model theory. Take, for example, the characterization of $\aleph_0$-categorical theories as those with finitely many n-types for each n. This characterization

ISSN: 2317-9570

points to an initial description of a class of theories (all their countable models are isomorphic) and proceeds to relate this to *structural* properties of the theories themselves (for each n, the space of n-types is finite), which explains why the initial description holds (this is very rough, but intuitively the space of n-types controls how much variation there can be between models of the same theory; so the smaller the space, the smaller the variation—the limit case here is when the space is compact and discrete, which implies that there can be *no* variation). The success of this program for the countable case suggests that perhaps we could attempt the general case; Shelah's classification program is precisely the extension of such ideas to the general case.

Such remarks are obviously very inchoate and were intended to be merely suggestive. But I do hope to have convinced the reader that there is at least *something* to this approach. If not, I must apologize for the time misspent!

References

FINE, Kit.“Essence and Modality”. *Philosophical Perspectives*, 8 (1994), pp. 1-16.

FINE, Kit. “Senses of Essence”. In: *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus*. Ed. by SINNOTT-ARMSTRONG, RAFFMAN, and ASHER. Cambridge: Cambridge University Press, 1994, pp. 53-73.

FINE, Kit.“The Logic of Essence”. *Journal of Philosophical Logic*, 24 (1995), pp. 241-273.

FINE, Kit.“Things and Their Parts”. *Midwest Studies in Philosophy*, 23 (1999), pp. 61-74.

FINE, Kit. “Semantics for the Logic of Essence”. *Journal of Philosophical Logic*, 29 (2000), pp. 543-584.

KRIPKE, Saul.*Naming and Necessity*. Cambridge, Massachusetts: Harvard University Press, 1980.

SHELAH, Saharon. *Classification Theory and the Number of Non-Isomorphic Models*. 2nd ed. Amsterdam: North-Holland, 1990.

Logicality as Necessary Instantiation

[Logicalidade como Instanciação Necessária]

Roderick Batchelor*

Abstract: We propose to define *logicality* of a simple notion as necessary instantiation. (And then logicality of a complex notion as logicality of all simple notions occurring in it.)
Keywords: logicality, logical constants, necessary existence, necessary instantiation.

Resumo: Propomos definir *logicalidade* de uma noção simples como instanciação necessária. (E então logicalidade de uma noção complexa como logicalidade de todas as noções simples nela ocorrendo.)
Palavras-chave: logicalidade, constantes lógicas, existência necessária, instanciação necessária.

1. Introduction

There seems to be a close link between *logicality* and *necessity*; and correlatively a close link between *empiricality* and *contingency*. Dramatically speaking, the logical realm is the realm of necessities, whereas the empirical realm is the realm of contingencies. (I mean logicality here in a broad sense; some might prefer to say 'logico-mathematicality'.)

The very *existence* of things or of instances of kinds of things is, typically, contingent in empirical cases and necessary in logical cases. Thus a paradigmatic empirical thing like the Eiffel Tower might of course not have existed; and a paradigmatic empirical concept like the concept of dog might not have had any instances. But a paradigmatic logical (or if you like logico-mathematical) thing like the empty set is no doubt a necessary existent (at least if it exists at all), and a paradigmatic logical concept like identity is necessarily instantiated.

One might then think of *defining* logicality as something like this – necessary existence or necessary instantiation. However, there are obvious counterexamples. The concept [dog or non-dog] is necessarily instantiated, but not logical; and the concept [identical and non-identical with] is logical, but not necessarily instantiated (indeed necessarily *un*instantiated).

But in such counterexamples we artfully construct certain *com-*

*Lecturer in Philosophy at the University of São Paulo. E-mail: roderick_batchelor@hotmail.com.

DOI: https://doi.org/10.26512/rfmc.v6i2.18891

plex concepts. Perhaps we can still use the idea above to define logicality for *simples*, and then somehow extend the definition to complexes? – I will argue in what follows that indeed we can.

2. A definition of logicality

By an *attributive* notion I mean a notion which may be 'attributed' or 'applied' to other notions so as to form a proposition. Possible examples of attributive notions are: the notions expressed by ordinary verbal expressions like 'smokes' or 'is parallel to'; propositional 'connectives', such as the notions expressed by 'and', 'Socrates believes that', etc.; quantifiers and other notions which produce propositions upon application to notions which are themselves attributive – thus [Something], [Every philosopher], etc.

By a *substance*-notion I mean a notion which is not attributive. Thus e.g. the notions expressed by 'Socrates' and by 'All men are mortal' are substance-notions.

I say that a notion is *complex* if at least one other notion enters into its constitution – thus e.g. [white horse] is of course a complex notion. And I say that a notion is *simple* if it is not complex.

By saying that an attributive notion is *instantiated* I mean that there is some *true* proposition consisting of the attributive notion in question applied to some notion or notions. Thus [white horse] is instantiated, but [flying horse] is not.

I will use also the concept of the *referent* of a simple substance-notion. (Different philosophers may wish to define this in different ways. I do not exclude that the referent be *identified* with the simple substance-notion itself.) By saying that a simple substance-notion is *instantiated*, then, I mean that its referent exists. (Again, different philosophers – e.g. 'actualists' on one side, 'possibilists' on the other – may wish to interpret this 'exists' in different ways.) (We will not need to define instantiation for *complex* substance-notions.)

We come finally to logicality. Let s be any *simple* notion (whether attributive notion or substance-notion); we then define:

s is logical $=_{df}$ Necessarily, s is instantiated.

Necessity here is 'metaphysical' necessity.

Now no doubt it should be the case that a *complex* notion is logical if and only if all the simple notions occurring in it are logical. (See BATCHELOR 2013, §§ 5–6.) So given any complex notion x, we

ISSN: 2317-9570

can define ‘x is logical’ as the conjunction of all propositions [s is logical] (in the sense already defined) for s a simple notion occurring in x. This completes the presently proposed definition of logicality.

It should be emphasized here that it would not be adequate to define logical notions in general (including *complex* ones) as necessarily instantiated notions. (Thus the title of the present note is elliptical (for the sake of lapidarity).) Indeed we have not even defined instantiation for complex substance-notions. But irrespectively of that, there are obvious counterexamples with complex attributive notions, as already indicated above in § 1.

3. Agreement with customary classification of specific notions

Now, how far does the presently proposed definition of logicality agree with the customary classification of specific notions or classes of notions as logical or non-logical? – *Very* far, it seems to me.

It will be enough to consider *simple* notions, or reasonable candidates for being simple notions, since the adequacy of our general definition relative to the adequacy of our definition of logicality for simple notions is straightforward.

Let us begin with what are usually reckoned as *being* (plausibly simple) logical notions.

The usual truth-functional connectives negation, conjunction, disjunction (or the Sheffer stroke etc.) are obviously necessarily instantiated. E.g. if all men are mortal then negation is instantiated by [Not all men are mortal]; and if not all men are mortal then negation is instantiated by [All men are mortal]. Only a ‘falsum’ truth-function, $\perp$, would be a counterexample; but surely it is implausible to think that there is a *simple* notion here; it seems much more plausible to think that $\perp(A)$ should be defined as (say) $A \wedge \neg A$ (and similarly for the binary, ternary etc. falsum connectives).

(Another way around an alleged simple falsum would be to re-define: s is logical $=_{df}$ Either it is necessary that s is instantiated or it is necessary that s is not instantiated – or in other words: Whether s is instantiated is not a contingent matter. It may also be argued on independent grounds that this is, conceptually, a more satisfactory definition: *empiricality* (of a simple) is here, naturally, equated to *contingency* of instantiation, and *logicality*, which is the contradictory of empiricality, is thus equated to *non-contingency* of instantiation. – The empirical realm is the realm of contingencies, and the logical realm the realm of non-contingencies. – However, it may also be argued on the other side that this (slightly more complex) reformulation is hardly worth the trouble, since surely there are no necessarily uninstantiated simple notions, so that the two formulations are in fact equivalent.)

Again, the universal and existential quantifiers, identity, metaphysical necessity and metaphysical possibility – these are all necessarily instantiated.

(For someone who thinks that there might be *nothing at all*, the above-suggested definition of logicality would inadequately count e.g. the existential quantifier as non-logical. But the following amended version would do, I suppose: s is logical $=_{df}$ Necessarily: if there is anything at all, then s is instantiated. – But now, if it is possible that there is exactly *one* thing, and say the quantifier [At least two things] is a simple notion, then this definition would still wrongly classify it as non-logical. But of course it is much more plausible to think that [At least two things] is a *complex* notion; and the same holds for the other obvious putative counterexamples which might be given in this connection.)

Also, *if* there are simple notions of set-membership, empty set, etc., or of natural numbers and usual numerical relations etc., *then* these will count as logical by our definition, as they should, given that as I said this is meant as definition of logicality in a broad sense (or 'logico-mathematicality'). (Assuming of course standard 'rigid' rather than eccentric 'contingent' Platonism.)

Coming now to simple notions usually reckoned as *not* being logical notions, it seems clear that they will indeed count as non-logical by our definition. No doubt simple empirical particulars do not necessarily exist; no doubt simple empirical properties and relations of particulars are not necessarily instantiated. And there is no good reason that I know of for thinking that there are any simple empirical propositional connectives at all, let alone necessarily instantiated ones; and similarly for other supposed simple higher-order empirical notions.

(In another paper (BATCHELOR 2011), I have suggested another, quite differently motivated, definition of logicality. *That* definition had some consequences (explicitly indicated in that other paper, § 2) which some philosophers have found too drastic. The *present* definition has no such consequences, and I hope that others may find it more acceptable. As for my giving two

ISSN: 2317-9570

different definitions of logicality, I do not see any problem about it: these are two different conditions both of which, I conjecture, pick out precisely what should count as logical notions – no claim being made in either case to an inevitable analysis leading to the unique 'essence' of the idea of logicality (which would be a hopeless goal in my opinion). So there is no problem if we have more than one such condition; it would be a problem if we had *less* than one.)

References

BATCHELOR, R. 'Topic-Neutrality', *Mind*, vol. 120, 2011, pp. 1–9.

BATCHELOR, R. 'Complexes and Their Constituents', *Theoria* (Lund), vol. 79, 2013, pp. 326–352.

Pluralismo, Monismo e Relativismo Lógico

[Logical Pluralism, Monism and Relativism]

Diogo Henrique Bispo Dias*

Resumo: Existe apenas uma lógica adequada? Ou há várias lógicas igualmente adequadas? O que significa, afinal, dizer que lógicas diferentes podem ser igualmente adequadas? E elas seriam adequadas com relação a quê? Este artigo pretende analisar as diferentes respostas a estas perguntas, ou seja, avaliaremos os argumentos centrais do debate entre pluralismo, relativismo e monismo lógico. Explicitaremos, por um lado, os principais pressupostos desta discussão e, por outro, suas ramificações filosóficas. Terminaremos indicando o desenvolvimento de um possível pluralismo lógico a partir da noção de paraconsistentização de lógicas, que será apresentada posteriormente. Não pretendemos refutar, de uma vez por todas, o monismo lógico. Mostraremos que os principais argumentos apresentados por um monista contra o pluralismo lógico são infundados e que, portanto, a existência de alguma forma de pluralismo lógico continua possível.
Palavras-chave: monismo lógico, pluralismo lógico, relativismo lógico, paraconsistentização

Abstract: Is there only one logic? Or are there several equally adequate logics? What does it mean, after all, that different logics can be equally adequate? And they would be adequate with respect to what? This article intends to analyze the different answers to these questions, that is, we will evaluate the central arguments of the debate between pluralism, relativism and logical monism. We will explain, on the one hand, the main assumptions of this discussion and, on the other hand, its philosophical ramifications. We will end by indicating the development of a possible logical pluralism using the notion of paraconsistentization of logics, which will be explained afterward. We do not intend to refute, once and for all, logical monism. We will show that the main arguments put forward by a monist against logical pluralism are unsound and therefore the existence of some form of logical pluralism remains possible.
Keywords: logical monism, logical pluralism, logical relativism, paraconsistentization

Introdução

O objetivo deste artigo é apresentar uma análise do debate entre pluralismo, relativismo e monismo lógico, evidenciando os principais pressupostos da discus-

*Doutorando no Programa de Pós-graduação em Filosofia da Universidade de São Paulo. E-mail: diogo.bispo.dias@gmail.com.

ISSN: 2317-9570

DOI: https://doi.org/10.26512/rfmc.v6i2.18883

são, bem como as ramificações filosóficas dessas diferentes posições. Além disso, indicaremos o desenvolvimento de um possível pluralismo lógico a partir da noção de paraconsistentização de lógicas, que será explicada adiante. A estratégia geral do texto é mostrar que os principais argumentos apresentados por um monista contra o pluralismo lógico são infundados e que, portanto, é possível a existência de alguma forma de pluralismo lógico.

Em termos gerais, pluralismo lógico é a tese de que há mais de uma lógica adequada, coerente ou, em algumas formulações, verdadeira. Dito de outra forma, isso significa que há vários modos de caracterizar adequadamente a noção de consequência lógica.

Relativismo, por sua vez,

> não é uma única doutrina, mas uma família de visões cujo tema comum é que alguns aspectos centrais da experiência, pensamento, avaliação ou, até mesmo, a realidade é, de alguma forma, relativo a outra coisa. Por exemplo, padrões de justificação, princípios morais, ou verdade, são ditos, às vezes, relativos à linguagem, cultura, ou constituição biológica. (SWOYER, 2003)[1]

Portanto, todo relativismo é uma instanciação do seguinte esquema geral: *Y* é relativo a *X*.

É importante notar que relativismo e pluralismo são distintos e independentes. A tendência de pensar que há uma relação necessária entre ambos surge porque, geralmente, os casos mais interessantes envolvem sobreposições das duas teses. Mas, a princípio, elas são independentes. Tomemos, por exemplo, o caso de regras de etiqueta. Se aceitarmos que etiqueta é relativa à cultura, mas que, no entanto, só há uma cultura, então somos relativistas, mas monistas com relação aos costumes. Portanto, relativismo só implica pluralismo se aceitarmos, também, que há mais de uma possibilidade de substituição da variável independente X, à qual Y é relativa. Por outro lado, é possível ser pluralista sem ser relativista. Basta defender que não há variável independente X à qual Y é relativa, mas que, não obstante, há diferentes tratamentos adequados de X.

No caso do relativismo lógico, a relação de consequência lógica ocupa a posição da variável de-

[1] Todas as traduções das citações são nossas.

ISSN: 2317-9570

pendente Y. A especificação da variável X, por sua vez, será discutida ao longo do artigo.

Do anterior, fica claro que relativismo lógico, assim conceituado, não significa que vale qualquer coisa em lógica, ou que todas as lógicas são igualmente adequadas. A tese relativista se aplica, por exemplo, à proposta de Tarski (1956) e Bolzano (1972) a respeito da possibilidade de diversas distinções não-equivalentes entre termos lógicos e não-lógicos. Aqui, a relação de consequência é relativa à divisão dos termos em lógicos e não-lógicos. E disso não se segue, como o próprio Tarski notou, que qualquer divisão é válida.

O pluralismo lógico defendido por J. C. Beall e Greg Restall (2006) também se caracteriza como uma forma de relativismo. Segundo estes autores, um argumento é válido se e somente se, em todo "caso" em que a premissa é verdadeira, a conclusão também é. Esta formulação é pluralista na medida em que se aceita, como faz os autores que há mais de um tratamento para esses "casos". Temos, também, um relativismo lógico ao reconhecer que a noção de consequência lógica é relativa aos possíveis "casos".

O monismo lógico, por outro lado, defende que há uma única lógica correta, adequada, ou verdadeira. Assim, há uma única e definitiva avaliação da validade de um dado argumento, ainda que, eventualmente, podemos desconhecer ou estarmos errados com respeito a esta avaliação.

O Debate

Há diversos problemas já na formulação destas distinções. E boa parte do debate se torna infrutífero caso tais problemas não sejam resolvidos ou, ao menos, explicitados. Em primeiro lugar, adequação, coerência e verdade são noções bem distintas. Sendo assim, é preciso separá-las e tratá-las isoladamente.

Usualmente, o conceito de verdade é tratado como um predicado de sentenças. Assim, parece não fazer sentido dizer que uma lógica é verdadeira. Quando alguns autores defendem esta tese[2], o que está em jogo, de fato, são pressupostos metafísicos. Em última instância, esses autores são realistas com respeito à lógica. Ou seja, quando afirmam que a lógica lida com preservação de verdade, eles acreditam que há uma noção absoluta de verdade, e que esta noção se funda em uma correspondência entre linguagem e

[2] Cf. READ (2006).

mundo. Neste caso, dizer que uma lógica é *a* lógica verdadeira é afirmar que existe apenas uma lógica que captura os aspectos formais desta correspondência.

As noções de adequação e correção, por outro lado, geralmente são invocadas quando não há o pressuposto de uma correspondência necessária entre linguagem e mundo, e se está meramente interessado na relação entre uma noção informal de consequência em determinado domínio e a noção formal de consequência em alguma lógica. Assim, uma lógica seria adequada ou correta se ela representa corretamente a relação informal de consequência em um dado domínio de objetos.

Nossa abordagem, como dito anteriormente, será considerar as críticas que um monista lógico propõe ao pluralismo e relativismo lógico, e mostrar que elas são, ou infundadas, ou insuficientes para rejeitar tais posições. Sendo assim, não nos preocuparemos em escolher uma noção particular de coerência, ou adequação. Utilizaremos as noções propostas pelos monistas, e mostraremos os problemas com as mesmas.

O próximo passo para avaliar o debate é distinguir níveis nos quais um pluralismo pode surgir. Em primeiro lugar, podemos ter um *pluralismo puro*, que consiste em afirmar a existência de distintas lógicas puras, isto é, formulações completamente abstratas da noção de consequência lógica, sem nenhuma preocupação com a eventual aplicação desta noção a algum domínio extralógico de objetos.

Podemos ter, também, um *pluralismo teórico*, que se preocupa com aplicações teóricas da lógica. Neste nível, ser um pluralista lógico consiste em defender que existem lógicas distintas igualmente adequadas para formalizar um mesmo domínio[3]. Por exemplo, em domínios inconsistentes, lógicas paraconsistentes são mais adequadas do que a lógica clássica, visto que, nesta, tudo se segue de contradições. Não obstante, muitas lógicas paraconsistentes são equivalentes à lógica clássica em domínios consistentes e, portanto, são igualmente adequadas nesta situação.

Por fim, há o chamado *pluralismo canônico*, que defende que mesmo quando o domínio a ser investigado é o da linguagem natural, isto é, investiga os cânones tradicionais de inferência, é possível ter mais de uma lógica adequada. Feitas as devidas distin-

[3]Isso não significa, obviamente, que, existe mais de uma lógica adequada para a formalização de todos os domínios.

Revista de Filosofia Moderna e Contemporânea, Brasília, v.6, n.2, dez. 2018, p. 21-36
ISSN: 2317-9570

ções, podemos levantar as seguintes perguntas: Alguma dessas formas de pluralismo está correta? Quais as consequências filosóficas destas posições? Analisemos caso a caso.

Alguns monistas lógicos defendem que um pluralismo puro, ainda que coerente, é simplesmente trivial[4]. Bueno (2002) afirma que, se um pluralismo puro é trivial, o é, no máximo, de um ponto de vista sociológico, ou seja, é trivial que hoje há várias lógicas puras. Mas, o fato de que isso é um fenômeno recente na lógica mostra que não se trata de algo filosoficamente trivial, e pode ser comparado com a revolução que as geometrias não-euclidianas causaram na matemática. Não obstante, o interessante não é meramente o fato de que há várias lógicas diferentes, mas como devemos interpretar essa pluralidade.

Do ponto de vista filosófico, esta pluralidade é extremamente importante. Por exemplo, Frege e Russell, dois dos maiores nomes da formulação da lógica clássica, defendiam uma visão universal da lógica, de tal modo que suas leis eram necessárias, *a priori*, e irrestritas com relação ao seu domínio de aplicação. Como consequência, nada poderia ser dito de fora da lógica. Assim, não só questões metalógicas estão ausentes nos escritos desses autores, mas a própria possibilidade de lógicas distintas e rivais é excluída de princípio[5]. Ainda que eles tenham efetuado mudanças profundas na noção de lógica, seu caráter universal permanece o mesmo desde, ao menos, Aristóteles. A aceitação de um pluralismo lógico, ainda que do ponto de vista puramente abstrato, é um marco na história da lógica, e muda significantemente o debate de questões centrais da lógica, tais como a determinação do significado dos conectivos lógicos, e a relação entre lógica e racionalidade. Outrossim, qualquer outro tipo de pluralismo lógico pressupõe a possibilidade de desenvolver múltiplas lógicas puras. Portanto, a existência de um pluralismo teórico não pode ser considerada meramente trivial.

O pluralismo teórico também é uma forma de relativismo, visto que uma lógica seria correta relativamente ao fenômeno que se pretende representar. Um monista lógico tende a argumentar que há critérios para escolher entre lógicas rivais na aplicação em um dado domínio[6]. Critérios

[4] Cf. PRIEST (2005).

[5] Isto está relacionado com a distinção entre lógica como cálculo e lógica como linguagem, proposta por van Heijenoort. Cf. van HEIJENOORT (1967).

[6] Cf/ PRIEST (2006), p. 195.

como simplicidade, falta de elementos *ad hoc*, adequação aos dados, etc. são geralmente usados. No entanto, não há garantia de que tais critérios estabeleçam a adequação de uma única lógica.

Tomemos, por exemplo, o conceito de simplicidade. Ao contrário do que parece, trata-se de um conceito muito complexo. O que significa ser simples? Ao avaliar uma lógica há, ao menos, três noções distintas de simplicidade[7], a saber: simplicidade com respeito à ontologia pressuposta, com respeito aos conceitos básicos, ou com respeito aos postulados lógicos. E, obviamente, esses três tipos de simplicidade podem entrar em conflito. Uma lógica pode ser mais simples que outra em termos de sua ontologia, mas conter conceitos básicos mais complexos; ou pressupor menos postulados lógicos, mas com uma ontologia mais forte que outra. Quando vários desses critérios são analisados em conjunto, a tarefa de determinar uma única lógica adequada parece menos factível ainda: uma lógica pode ser mais simples que outra, mas com menor adequação aos dados; pode ter menos elementos *ad hoc*, mas ser mais complexa etc. Assim, em princípio, não podemos descartar uma forma de pluralismo teórico.

Contra a existência do pluralismo canônico, Priest(2006) argumenta que: ou lógicas diferentes são rivais, ou não o são, isto é, ou elas discordam acerca da validade de certos argumentos, ou não discordam. Se não forem rivais, então não importa qual será a escolhida. Todas avaliam igualmente os argumentos e, portanto, para a tarefa de formalizar argumentos da linguagem natural, essas lógicas são equivalentes. Porém, se elas forem rivais, então apenas uma delas é correta. De qualquer forma, o resultado final é uma única lógica. Bueno (2002) defende que este é um falso dilema, pois postula um pressuposto que é rejeitado por um lógico pluralista, a saber: que quando duas lógicas rivais são diferentes, apenas uma delas está certa. A recusa desta tese consiste precisamente em um dos pontos essenciais do pluralismo lógico.

Um pluralista lógico defende que é possível que – ao menos – duas lógicas discordem e, não obstante, sejam igualmente adequadas para um determinado domínio. Isso pois, em primeiro lugar, essas lógicas podem concordar inteiramente neste domínio, mas discordar em domínios diferentes. Este é exatamente o caso da lógica paraconsistente e da ló-

[7] Cf. LEHRER (1990), capítulo 5.

ISSN: 2317-9570

gica clássica. Ambas concordam em domínios consistentes, mas diferem em domínios inconsistentes. Do mesmo modo, a lógica clássica e a intuicionista avaliam igualmente os argumentos em domínios finitos, mas discordam em domínios infinitos. Além disso, é possível que uma determinada lógica seja capaz apenas de formalizar partes de uma dada teoria, e outra lógica somente consiga formalizar outras partes desta teoria e, no entanto, podemos não conseguir determinar qual formalização é a mais adequada. Para ver como isso é possível, consideremos a seguinte objeção ao pluralismo:

> Suponha que alguém seja pluralista [com respeito à lógica]. Seja s alguma situação sobre a qual estamos raciocinando; suponha que s esteja em classes diferentes de situações, como K_1 e K_2. Devemos usar a noção de validade apropriada para K_1 ou K_2? Não podemos responder 'ambas' aqui. Tome alguma inferência que é válida em K_1, mas inválida em K_2, $\alpha \vdash \beta$, e suponha que saibamos (ou assumamos) que α vale em s; podemos, ou não, aceitar β? Ou sim, ou não: não pode haver pluralismo sobre isto. De fato, a resposta é que podemos, uma vez que s está em K_1, e a inferência preserva verdade em todas as situações em K_1. Em outras palavras, se sabemos que uma situação sobre a qual estamos raciocinando está na classe K, estamos justificados em raciocinar com a validade definida em uma classe restrita de situações K. (PRIEST, 2006, p. 203.)

O problema com esse raciocínio é que ele nos conta apenas parte da história. Ainda que na situação descrita consigamos saber qual lógica usar, há vários contextos em que tal decisão não é óbvia. Considere, por exemplo, $\alpha \vdash \neg\beta$ é válido em K_2. Assim, teríamos uma situação inconsistente[8], e não saberíamos, necessariamente, qual lógica seria mais adequada. Ademais, imaginem a seguinte situação análoga: temos esse cenário apresentado por Priest e, além disso, sabemos que $\alpha \vdash \delta$ é válida em K_2, mas inválida em K_1. Além disso, suponha que δ e β sejam igualmente importantes para a teoria

[8] O que é aceitável se estivermos no âmbito de uma lógica paraconsistente.

sendo formalizada. Assim, os tratamentos dados por K_1 e K_2 são igualmente adequados, e não temos, necessariamente, um critério para decidir qual deve ser usado.

Uma possível objeção a essa estratégia de discutir a adequação de lógicas relativa ao seu domínio de aplicação é a seguinte[9]: ainda que possamos falar que lógicas diferem em domínios diferentes, o que fazer quando precisamos raciocinar em múltiplos domínios? Ora, neste caso, temos um novo domínio e, portanto, podemos ter, novamente, mais de uma lógica adequada para formalizá-lo. Por exemplo, podemos raciocinar em um domínio construtivo e inconsistente. Assim, precisamos de alguma combinação entre uma lógica paraconsistente e uma intuicionista para estudá-lo. A objeção, então, afirmaria que a solução para a rivalidade entre lógicas seria combiná-las e produzir uma única lógica para formalizar um dado domínio. Assim, esta lógica resultante seria a única lógica adequada para tal tarefa.

A falha desta abordagem é que, a princípio, não há nenhum método geral para combinar lógicas. Isto significa que, dadas duas lógicas, pode não haver uma forma de combiná-las ou, pior ainda, essa combinação pode levar a resultados indesejados, como o *problema do colapso*[10]: a combinação entre duas lógicas pode ser equivalente a uma das lógicas iniciais, ou essa combinação pode gerar novas e indesejáveis interações entre os conectivos das lógicas inicias[11]. De um lado, isso pode fazer com que a lógica resultante seja considerada muito forte, no sentido de que ela tem um poder de inferência muito maior do que as lógicas iniciais. Por outro lado, esta combinação pode ser muito fraca, e falhar em inferir certas proposições importantes que eram captadas pelas lógicas iniciais. Portanto, a combinação de lógicas não exclui a possibilidade de algum pluralismo lógico. Aceitar ou rejeitar a ideia de que a lógica é relativa ao domínio de implicação tem sérias consequências para questões centrais da lógica. Vejamos isto em detalhes. Um dos precursores desta noção pluralista de variação do domínio é Newton da Costa, que afirma que:

> É claro que, para objetos comuns, como um livro ou uma pessoa, (...) [o

[9] Essa objeção foi levantada por Ingolf Max em conversa particular.

[10] Cf. CARNIELLI & CONIGLIO (2016), seção 5.

[11] Por exemplo, a combinação da disjunção de uma lógica com a conjunção de outra lógica pode gerar certas propriedades distributivas entre estes conectivos na lógica resultante. Essas propriedades podem ser – ou não – adequadas para o domínio sob investigação.

ISSN: 2317-9570

> princípio lógico de identidade] se aplica sem uma única dificuldade importante. Qualquer pessoa *A*, por mais que sofra múltiplas modificações ao longo de sua vida, se mantém, em um certo sentido, idêntica a si mesma: $A = A$. Isto parece ainda mais claro no que diz respeito a objetos abstratos: por exemplo, a igualdade $1 = 1$ parece evidente e incontestável (...). No entanto, as coisas não são tão simples quanto um realismo ingênuo nos leva a crer. Na física quântica, as partículas elementares, ao que tudo indica, infringem o princípio de identidade. Assim, Schrödinger afirma que a relação de identidade entre partículas não tem sentido (...). Pode ser que a posição de Schrödinger seja aceitável apenas temporariamente e que o futuro nos mostrará que ele está errado. No entanto, o fato é que a física quântica mostra a possibilidade de dialetizar a ideia de identidade e, por consequência, a própria lei que corresponde a ela. (da COSTA, 1997, pp. 120-1.)

Portanto, a lógica a ser usada varia de acordo com o domínio de fenômenos estudados. A resposta monista consiste em afirmar que

> como validade é preservação de verdade em todas as situações, se existem situações nas quais objetos podem ser não-idênticos, então o princípio de identidade não é, no final das contas, lógico. É apenas uma propriedade 'contingente' de alguns domínios, e pode ser invocada quando se está raciocinando sobre eles. (PRIEST, 2006, p. 198.)

É interessante notar que ambas as posições estão rejeitando uma característica considerada, por muitos, como fundamental da lógica, a saber: seu caráter *a priori*. Isso, pois, tanto defender que as leis lógicas são relativas a certos domínios, quanto afirmar que existem situações empíricas que podem nos levar a rejeitar uma lei lógica, implica em aceitar que estabelecer quais são as leis lógicas é uma tarefa *a posteriori*, visto que essas leis podem mudar de acordo com as propriedades físicas que encontramos nos objetos. Por outro lado, a primeira tese rejeita o pressuposto de que a lógica é neutra com relação a seu objeto, uma vez que as leis lógicas variam de acordo com os fenômenos estudados. Um monista, por sua vez, mantém o caráter puramente formal da lógica ao defender que se uma lei é relativa a certos domínios, então tal lei não pode ser lógica.

Não obstante, esta postura monista pode, em última instância, levar à conclusão de que não existe nenhuma verdade – ou lei – lógica. Não seria possível imaginar um conjunto de fenômenos que rejeite cada uma das consideradas leis lógicas[12]? Portanto, teríamos leis locais que ditam o comportamento dos fenômenos em determinados domínios, mas não haveria leis que valessem independentemente do fenômeno investigado.

Outrossim, sabemos que toda teoria tem alguma lógica subjacente. Se os resultados dessa teoria forem utilizados para reformular a lógica, será que não teríamos algum tipo de círculo vicioso? Além disso, enfrentaríamos o problema clássico de decidir, face a um experimento que contrarie nossa expectativa, qual parte da teoria devemos rejeitar (incluindo, agora, sua lógica subjacente).

Pluralismo e Significado

Passemos, agora, ao que consideramos a principal crítica ao pluralismo lógico, a saber: a tese de que, quando mudamos de lógica, mudamos o significado dos conectivos lógicos. A formulação clássica desta crítica está nas seguintes palavras de Quine:

> quem nega a lei do terceiro excluído está mudando o assunto. Isto não é o mesmo que dizer que ele está errado em fazê-lo. Ao rejeitar '$p \vee \neg p$' ele está, de fato, abandonando a negação clássica (...); e ele pode ter seus motivos. (QUINE, 1960, p. 100.)

O cerne desta posição está na ideia de que, quando mudamos a lógica, mudamos o significado dos conectivos lógicos. Obviamente, esta tese depende fundamentalmente de alguma teoria do significado estipulando quais os elementos que determinam o significado dos conectivos lógicos.

Lidaremos, neste texto, com duas formulações comuns na literatura acerca da especificação do significado dos operadores lógicos. Para os monistas lógicos que defendem a primazia do apa-

[12]Esta tese é conhecida como possibilismo ou não-necessitarianismo. Cf. ESTRADA-GONZÁLEZ (2011).

rato semântico da lógica, o significado dos conectivos é dado por suas condições de verdade. As primeiras formulações explícitas desta tese foram dadas por Frege (1964, I, §32) e Wittgenstein (1922, §4.024). Assim, condições de verdades diferentes implicam em significados diferentes e, portanto, em conectivos lógicos diferentes.

Há, também, lógicos que privilegiam o aparato sintático da lógica. Neste caso, o significado dos conectivos é determinado pelas suas regras de inferência. Esta ideia foi proposta, pela primeira vez, por Carnap (1937). Curiosamente, este livro também contém a primeira formulação do Princípio de Tolerância, que consiste, segundo alguns comentadores, na primeira defesa de um tipo de pluralismo lógico[13]. Independente do possível pluralismo presente nesta obra, Carnap aceitava que uma mudança nas regras de inferência de um conectivo implicava na mudança do seu significado.

Pretendemos mostrar que nenhuma dessas teses acerca do significado dos conectivos leva a um monismo lógico. Comecemos com a abordagem semântica. Tomemos, como exemplo, as lógicas LP[14] e K_3[15]. Ambas as lógicas contém as mesmas tabelas de verdade para seus conectivos:

α	β	$\alpha \to \beta$
1	1	1
1	i	i
1	0	0
i	1	1
i	i	i
i	0	i
0	1	1
0	i	1
0	0	1

Implicação

α	β	$\alpha \vee \beta$
1	1	1
1	i	1
1	0	1
i	1	1
i	i	i
i	0	i
0	1	1
0	i	i
0	0	0

Disjunção

α	β	$\alpha \wedge \beta$
1	1	1
1	i	i
1	0	0
i	1	i
i	i	i
i	0	0
0	1	0
0	i	0
0	0	0

Conjunção

[13]Para uma análise do pluralismo subjacente ao Princípio de Tolerância, bem como uma investigação do seu limite, cf. DIAS (2015).

[14]Cf. PRIEST (1979).

[15]Cf. PRIEST (2008), pp.122-4.

α	$\neg\alpha$
1	0
i	i
0	1

Negação

Formalmente, a diferença entre as duas lógicas está nos valores designados[16]. Em K_3, um argumento é válido quando é impossível que suas premissas tenham valor 1, e a conclusão tenha valor 0. Por outro lado, um argumento é válido em *LP* se é impossível que suas premissas tenham valor 1 ou i, mas sua conclusão seja 0. É simples verificar que a lei de não-contradição é válida em *LP*, mas inválida em K_3[17]. Mas, como as condições de verdade dos conectivos são as mesmas, segue-se que os conectivos das duas lógicas são os mesmos.

Também há abordagens sintáticas que permitem a formulação de diferentes lógicas, mas mantendo as mesmas regras de inferência para seus respectivos conectivos. Greg Restall (2014), por exemplo, desenvolve um aparato formal que permite manter as regras de inferência intactas, e variar algumas propriedades estruturais da lógica como, por exemplo, a possibilidade de um argumento ter múltiplas conclusões ou não. Quando permitimos múltiplas conclusões, obtemos a lógica intuicionista; caso contrário, temos a lógica clássica. Sendo assim, variando tais propriedades estruturais, obtemos lógicas distintas, mas com os mesmos conectivos.

Logo, nenhuma das formulações usuais do significado dos conectivos lógicos implica em um monismo lógico. Mas, o que dizer quando aceitamos que tal significado é dado pela conjunção das abordagens semânticas e sintáticas, isto é, quando o significado dos conectivos é determinado pelas suas condições de verdade *mais* suas regras de inferência. Neste caso, os contraexemplos propostos acima não se sustentam: os conectivos de K_3 e *LP* têm as mesmas tabelas de verdade, mas suas regras de inferência são diferentes; as lógicas clássica e intuicionista têm as mesmas regras de inferência – no aparato desenvolvido por Restall –, mas seus conectivos têm condições de verdade distintas.

Mesmo nesta situação, ainda existe a possibilidade de um pluralismo lógico, usando a noção de *paraconsistentização de lógicas*, cunhado por Alexandre Costa-Leite (2007)[18]. Paraconsistentizar uma

[16] A interpretação intuitiva do valor i também é diferente. Em *LP*, i significa 'verdadeiro e falso'; em K_3, i significa 'nem verdadeiro, nem falso'.

[17] De fato, K_3 não tem nenhuma fórmula válida.

[18] Ver, também, de SOUZA; COSTA-LEITE & DIAS (2016).

ISSN: 2317-9570

lógica significa transformar uma dada lógica inicial em uma lógica paraconsistente. É possível paraconsistentizar uma lógica usando tanto com uma abordagem semântica, quanto sintática[19]. Em linhas gerais, esta estratégia consiste em limitar a relação de consequência da lógica inicial, evitando a trivialização de conjuntos inconsistentes. Considere uma lógica explosiva L, isto é, uma lógica na qual, de premissas contraditórias, tudo se segue, e defina a sua contraparte paraconsistente P_L da seguinte forma:

Uma fórmula α é consequência lógica em P_L de um conjunto Γ de fórmulas se e somente se existe $\Gamma' \subseteq \Gamma$, consistente em L, tal que α é consequência lógica em L de Γ'.

Notem que a única alteração com respeito à lógica inicial se dá na definição de consequência lógica. Não há mudança nas regras de inferência da lógica inicial[20]. A paraconsistentização é obtida através de uma restrição na aplicação dos conectivos a subconjuntos consistentes de um certo conjunto inicial. Quando este conjunto inicial é consistente, as inferências permitidas em P_L são exatamente as mesmas que em L. Quando o conjunto é inconsistente, esta restrição evita a explosão da lógica inicial. De acordo com a tese de que o significado dos conectivos é dado pelas suas regras de inferência, essa restrição não contribui para a constituição deste significado. Portanto, os conectivos são os mesmos nas duas lógicas[21].

Encontramos uma situação análoga quando analisamos a paraconsistentização do ponto de vista semântico. Sem entrar em detalhes técnicos, o procedimento consiste em substituir a noção de *consistência* por *satisfabilidade*. Novamente, temos uma restrição na utilização dos conectivos, mas sem alterar suas condições de verdade. Logo, os conectivos são os mesmos da lógica inicial.

Diferentemente dos exemplos anteriores, em que a preservação dos conectivos era limitada ao aparato semântico ou sintático, a paraconsistentização apresentada acima é realizada de maneira uniforme pelas abordagens semântica e sintática. Dito de outra forma, esta paraconsistentização preserva correção e completude, isto é, se a lógica inicial for correta e completa, sua con-

[19]Para uma apresentação detalhada deste procedimento, cf. de SOUZA; COSTA-LEITE & DIAS (no prelo).

[20]Há, na verdade, uma mudança na definição de dedução.

[21]O fato de que as lógicas são equivalentes em contextos consistentes reforçam a ideia de que elas têm os mesmos conectivos.

[22]A prova deste resultado encontra-se em de SOUZA; COSTA-LEITE & DIAS (no prelo).

traparte paraconsistente também será[22]. Isto garante que os conectivos obtidos pela transformação semântica são os mesmos que aqueles obtidos sintaticamente. Portanto, se levarmos às últimas consequências a tese de que o significado dos conectivos lógicos é dado por suas regras de inferência e por suas condições de verdade, a paraconsistentização mostra que, ainda assim, é possível haver duas lógicas distintas, isto é, que discordem sobre a validade de um mesmo argumento e, não obstante, contém os mesmos conectivos lógicos.

Conclusão

Neste artigo, procuramos evidenciar os principais pressupostos da discussão entre pluralismo, monismo e relativismo lógico. A discussão envolve pressupostos metafísicos – realismo e antirrealismo lógico –, epistemológicos e pragmáticos – critérios epistêmicos e pragmáticos para escolher uma lógica, linguísticos – teoria do significado para as linguagens lógicas – e, obviamente, lógicos – *a prioricidade*, universalidade e, até mesmo, a própria função da lógica. Mostramos que os principais argumentos monistas falham em refutar a possibilidade de algum tipo de pluralismo lógico. Com respeito à tese do significado dos conectivos, exibimos como construir uma lógica paraconsistente que contém os mesmos conectivos da lógica inicial. Temos, assim, uma forma de pluralismo paraconsistente de lógicas. Este artigo não tem a pretensão de estabelecer, de uma vez por todas, que o monismo lógico está definitivamente superado, mas, sim, mostrar que o debate continua vivo. Encontramo-nos em uma situação análoga ao desenvolvimento das geometrias não-euclidianas no século XIX. Nas palavras de Coffa:

> Durante a segunda metade do século XIX, através de um processo que ainda espera explicação, a comunidade de geômetras chegou à conclusão de que todas as geometrias estavam aqui para ficar(...) Isto teve toda a aparência de ser a primeira vez que uma comunidade de cientistas concordou em aceitar de uma forma não meramente provisória todos os membros de um conjunto de teorias inconsistentes sobre um mesmo domínio(...) Cabe agora aos filósofos dar algum sentido epistemológico da atitude dos matemáticos com relação à geometria. O desafio foi um teste

> difícil para os filósofos, teste esse no qual (infelizmente) todos falharam. (COFFA, 1986, p. 17)

Estamos vivendo uma situação semelhante na lógica. Não apenas o surgimento de uma pluralidade de lógicas, mas, também, de uma pluralidade de abordagens à lógica levou ao questionamento e, potencialmente, a mudanças em questões centrais da lógica como, por exemplo, a possibilidade da existência de diversas lógicas igualmente adequadas. Esta é uma tese que não faria o menor sentido até pouco menos de cem anos atrás. Sendo assim, independente de um eventual vencedor desta querela, o debate nos permite aumentar nosso conhecimento sobre lógica o que, afinal, é o objetivo principal de todo lógico.

Referências

BEALL, J.C.; RESTALL, G. *Logical Pluralism*. Oxford: Oxford University Press, 2006.

BOLZANO, B. *Theory of Science*. Berkeley: University of California Press, 1972.

BUENO, O. "Can a Paraconsistent Theorist be a Logical Monist?" In: CARNIELLI, W.; CONIGLIO, M.E.; D'OTTAVIANO, I.M.L. (eds). *Paraconsistency: The logical way to the inconsistent*. New York: Marcel Dekker, 2002, pp. 535-552.

CARNAP, R. *The Logical Syntax of Language*. London: Kegan Paul, 1937.

CARNIELLI, W.; CONIGLIO, M.E. "Combining Logics". *The Stanford Encyclopedia of Philosophy*, 2016. Disponível em: http://plato.stanford.edu/archives/sum2016/entries/logic-combining. Acessado em: 25/09/2018.

COFFA, A.J. "From geometry to tolerance: sources of conventionalism in nineteenth-century geometry". In: COLODNY, R.G. (ed.). *From Quarks to Quasars: Philosophical problems of modern physics*. University of Pittsburgh Series, Volume 7. Pittsburgh: Pittsburgh University Press, 1986, pp. 3-70.

COSTA-LEITE, A. *Interactions of metaphysical and epistemic concepts*. Tese (Doutorado em Filosofia) - Université de Neuchâtel, Switzerland, 2007.

da COSTA, N.C.A. Logiques Classiques et Non Classiques: Essai Sur les Fondements de la Logique. Paris: Masson, 1997.

de SOUZA, E.G.; COSTA-LEITE, A.; DIAS, D.H.B. "On a paraconsistentization functor in the category of consequence structures". *Journal of Applied Non-Classical Logics*, v. 26, 2016, pp. 240-250.

de SOUZA, E.G.; COSTA-LEITE, A. DIAS, D.H.B. "Paradeduction in Axiomatic Formal Systems". *Logique & Analyse*. No prelo.

DIAS, D.H.B. "Carnap's Principle of Tolerance and logical pluralism". *Argumentos: Revista de Filosofia (Online)*, v. 13, 2015, pp. 225-236.

ESTRADA-GONZÁLEZ, L. "On the Meaning of Connectives (Apropos of a Non-Necessitarianist Challenge)". *Logica Universalis*, 5, 2011, pp. 115-126.

FREGE, G. *The Basic Laws of Arithmetic*. Berkeley: University of California Press, 1964.

LEHRER, K. *Theory of Knowledge*. Oxon: Routledge, 1990.

PRIEST, G. *Doubt Truth to be a Liar*. Oxford: Clarendon Press, 2006.

PRIEST, G. *An Introduction to Non-Classic Logic*. Cambridge: Cambridge University Press, 2008.

PRIEST, G. *Logic of Paradox*. Journal of Philosophical Logic 8, 1979, pp. 219-41.

QUINE, W. von O. *Word and Object*. Cambridge: Massachusetts: The MIT Press, 1960.

READ, S. "Monism: The One True Logic". In: DEVIDI, D.; KENYON, T (eds.). *A Logical Approach to Philosophy: Essays in Honour of Graham Solomon*. Dordrecht: Springer, 2006, pp. 193-209.

RESTALL, G. "Pluralism and proof". *Erkenntnis*, 79:2, 2014, pp. 279-291

SWOYER, C. "Relativism". *Stanford Internet Encyclopedia of Philosophy*, 2003. Disponível em: http://plato.stanford.edu/entries/relativism. Acessado em: 25/09/2018.

TARSKI, A. "On the concept of logical consequence". In: TARSKI, A. *Logic, Semantics and Metamathematics*. Oxford: Clarendon Press, 1956, pp. 409-420. van HEIJENOORT, J. "Logic as calculus and logic as language". *Synthese*, 17(1), 1967, pp. 324–330.

WITTGENSTEIN, L. *Tractatus Logico-Philosophicus*. London: Routledge, 1922.

ISSN: 2317-9570

Da Semântica para Demonstrações de Consistência e a Volta

[From Semantics to Consistency Proofs and Back]

Rodrigo A. Freire*; e Luiza S. P. Ramos**

Resumo: O presente artigo contém duas teses principais. Primeiro, que o ponto de partida em uma demonstração de consistência de um sistema formal é uma noção semântica. Essa tese é apresentada a partir de uma análise das etapas pelas quais uma demonstração de consistência passa, uma vez que um atributo de fórmulas com base em alguma interpretação deve ser estipulado já na etapa inicial. Para avançar cada uma das etapas, uma demonstração de consistência deve produzir um ganho de entendimento correspondente em relação ao sistema. Confirmamos a tese por um estudo de casos de três sistemas em que três demonstrações construtivas de consistência correspondentes são analisadas. O estudo é restrito a demonstrações construtivas, pois no caso modelo-teórico a tese é evidente. Estudamos a lógica de primeira ordem, a aritmética sem indução, também conhecida como aritmética de Robinson, e a aritmética com indução ou de Peano. Em seguida, consideramos a relação entre consistência e verdade na aritmética sob o prisma do estudo de casos. A partir disso, podemos formular a segunda tese: há uma concepção da verdade aritmética que não se compromete de partida com a consistência. Tal tese é motivada tendo em vista as limitações para demonstrações de consistência apresentadas.
Palavras-chave: demonstrações de consistência, aritmética, verdade

Abstract: The current article endorses two main thesis. First we claim that the starting point in a consistency proof of a formal system is a semantic notion. This thesis is supported by an analysis of the stages in which a consistency proof goes through once that already in the initial stage an attribute for formulas must be specified on a interpretative basis. In order to move one stage forward at a time, a consistency proof must give an additional information corresponding to the system. We corroborate the thesis with case studies of three systems by which their constructive consistency proofs are analysed. Each study is restricted to constructive proofs since in the model-theoretic case the thesis is trivial. We analyse first-order logic, arithmetic without induction, also known as Robinson's arithmetic, and arithmetic with induction or Peano's. Then, we consider the relation between consistency and truth in Peano arithmetic in light of the case study. In regard to this we can formulate our second thesis: there is a conception of arithmetic truth in which no commitment with consistency must be taken for granted. The motivation for this thesis is the limitation for achieving constructive consistency proofs presented in the article.
Keywords: consistency proofs, arithmetics, truth

*Professor do Departamento de Filosofia da Universidade de Brasília (UnB). E-mail: freirester@gmail.com.
**Mestranda no PPGM da UnB. E-mail: luizaspramos@gmail.com.

DOI: https://doi.org/10.26512/rfmc.v6i2.19201

1 - Considerações Gerais Sobre Demonstrações de Consistência

Partimos da noção de sistema formal, estabelecida a partir das investigações sobre os fundamentos da matemática que ocorreram no final do século XIX e no início do século XX. O problema da demonstração de consistência surge no contexto de tais investigações, principalmente a partir da desinterpretação dos sistemas axiomáticos. Para um estudo do desenvolvimento dos sistemas formais veja Arno e Raggio. Para mais detalhes técnicos e históricos veja Bourbaki, especialmente a *Note Historique*. A consistência de teorias matemáticas se torna objeto de um estudo matemático apenas a partir de tal desenvolvimento, como resume Kleene:

> É uma contribuição de Hilbert ter concebido uma nova abordagem direta, e ter reconhecido o que esta envolvia para a axiomatização. Este método direto está implícito no significado de consistência (pelo menos como agora o pensamos), nomeadamente que uma contradição lógica (uma proposição A e sua negação $\neg A$ serem ambas teoremas) não pode surgir da teoria deduzida a partir dos axiomas. Portanto, para demonstrar a consistência de uma teoria diretamente, deve-se demonstrar uma proposição sobre a teoria ela própria, isto é, especificamente sobre todas as possíveis demonstrações de teoremas na teoria. A teoria matemática cuja consistência queremos demonstrar torna-se assim ela própria objeto de um estudo matemático, o qual Hilbert denomina "metamatemática" ou "teoria da demonstração". [1]

Usamos como sinônimas as expressões 'sistema formal', 'sistema axiomático formalizado' e 'sistema dedutivo'. Demonstrar a consistência de um sistema formal para uma teoria clássica é demonstrar que nem toda fórmula do sistema é dedutível no sistema. Estamos interessados em elucidar a estrutura das demonstrações de

[1]Tradução nossa do original, *It is Hilbert's contribution now to have conceived a new direct approach, and to have recognized what it involves for the axiomatization. This direct method is implicit in the meaning of consistency (at least as we now think of it), namely that no logical contradiction (a proposition A and its negation notA both being theorems) can arise in the theory deduced from the axioms. Thus to prove consistency of a theory directly, one should prove a proposition about the theory itself, i.e. specifically about all possible proofs of theorems in the theory. The mathematical theory whose consistency it is hoped to be proved then becomes itself the object of a mathematical study, which Hilbert calls "metamathematics" or "proof theory".* (Kleene, pg.55.)

consistência, entender quais são os problemas que devemos resolver para alcançar tais demonstrações, avaliar as dificuldades impostas por esses problemas e seu impacto para a fundamentação da matemática.

Vamos começar pela apresentação de um esquema das etapas que compõem, em geral, tais demonstrações. Podemos destacar quatro etapas que constituem um esquema geral para entender demonstrações de consistência. Enumeramos abaixo essas quatro etapas, explicadas uma a uma a seguir.

1. Uma propriedade $\mathcal{P}$ de fórmulas do sistema é escolhida.
2. Uma classe Δ de fórmulas do sistema é determinada.
3. A correção com relação a $\mathcal{P}$ das fórmulas em Δ que são dedutíveis é estabelecida.
4. Uma fórmula em Δ que não satisfaz $\mathcal{P}$ é determinada.

A propriedade $\mathcal{P}$ de fórmulas do sistema deve ser escolhida criteriosamente, o sucesso das etapas seguintes depende de uma escolha apropriada no início. As escolhas apropriadas de $\mathcal{P}$ são, em geral, tais que se A é uma fórmula do sistema, então $\mathcal{P}(A)$ diz que A é o caso segundo uma interpretação definida. Nas demonstrações modelo-teóricas de consistência $\mathcal{P}(A)$ diz que A é válida em uma estrutura determinada. De forma similar, nas demonstrações construtivas partimos de um entendimento semântico, no entanto buscamos extrair uma propriedade construtiva, ou até mesmo finitária, desse entendimento. Reforçamos que o ponto de partida das demonstrações de consistência, o que guia a escolha da propriedade $\mathcal{P}$, mesmo no caso das demonstrações finitárias, é semântico, o que é explicado de modo claro por Kleene:

> Todas essas demonstrações de consistência dependem da disponibilidade de um modelo para os axiomas, assim como aquelas dadas antes do advento da teoria de Hilbert da demonstração. Mas dar um modelo para os axiomas em termos de uma aritmética intuitiva não estabelece, além de toda dúvida, que nenhuma contradição pode surgir na teoria que e deduzida a partir dos axiomas, a não ser que possa também ser demonstrado que os argumentos na teoria possam ser traduzidos em argumentos aritméticos intuitivos nos termos dos objetos usados no mo-

delo. [2]

Essa passagem também chama a atenção para a importância das etapas seguintes, especialmente do que estamos chamando de correção. Não basta escolher a propriedade $\mathcal{P}$. Há uma correspondência direta entre (i) dar um modelo em termos de uma aritmética intuitiva e demonstrar que os argumentos da teoria possam ser traduzidos em argumentos aritméticos intuitivos nos termos dos objetos usados no modelo, conforme a formulação de Kleene, e (ii) dar uma propriedade finitária $\mathcal{P}$ e demonstrar a correção com relação a essa propriedade das fórmulas dedutíveis, ou pelo menos parte delas, de acordo com nosso esquema.

Por isso, as segundas e terceiras etapas têm por finalidade isolar uma classe Δ de fórmulas suficientemente amplas e tal que toda fórmula em Δ que é dedutível no sistema cuja consistência está em estudo satisfaz $\mathcal{P}$. A classe Δ não precisa ser a classe de todas as fórmulas do sistema, mas precisa conter alguma fórmula para que a última etapa seja possível. Na etapa final determina-se uma fórmula em Δ que não satisfaz $\mathcal{P}$. Pela etapa anterior, tal fórmula também não é dedutível e, portanto, o sistema é consistente. Assim, o núcleo de uma demonstração construtiva de consistência é composto pela escolha de uma propriedade $\mathcal{P}$ que seja construtiva e pela demonstração construtiva da correção com relação a $\mathcal{P}$ de uma classe apropriada de fórmulas dedutíveis.

2 - Por Que Buscar Demonstrações Construtivas de Consistência

Como já ressaltamos, o problema da consistência está bem formulado, pelo menos a partir de Hilbert, com base na possibilidade de representar a matemática sem recurso ao infinito.[3] A tentativa de resolver esse problema preliminar é baseada na ideia que bastaria representar o uso dos símbolos, e isso pode ser feito por meio de regras de uso que são, também, de natureza simbólica. Qualquer representação de entendimento de outra natureza do simbolismo não seria requerida. Essa possibilidade pode ser vislumbrada obser-

[2]Tradução nossa do original, *These consistency proofs all depend on having a model for the axioms, as did those given before the advent of Hilbert's proof theory. But giving a model for the axioms in intuitive arithmetical terms does not establish beyond all doubt that no contradiction can arise in the theory deduced from the axioms, unless it can also be demonstrated that the reasonings in the theory can be translated into intuitive arithmetical reasonings in terms of the objects used in the model.* (Kleene, p. 475.)

[3]Veja Hilbert, especialmente as páginas 191 e 192.

ISSN: 2317-9570

vando que o pensamento matemático opera por meio de frases na sua atividade de definir e demonstrar. Definições são determinadas por frases descritivas e demonstrações por encadeamentos de frases. O caminho para representar de modo puramente simbólico a atividade matemática passaria, então, pela construção de um sistema de frases com a delimitação precisa das suas regras de uso.

O emprego das frases nas deduções do sistema é determinado pela estipulação dos axiomas, ou seja, de quais frases podem ser usadas como premissas, e das regras de inferência, que determinam quais passagens são válidas na atividade dedutiva. Também é necessário saber usar frases descritivas para introduzir novos símbolos a partir dos símbolos inicialmente estipulados. Um símbolo como 'π' seria introduzido no nosso sistema que representa a matemática com uma descrição que determina de modo completo, dentro do sistema, o uso do novo símbolo. Note que o significado do símbolo 'π' não é parte constituinte desse processo e o infinito não ocorre nessas estipulações.

A lógica de primeira ordem serve de base para uma representação livre de infinitude atual da atividade matemática a partir de seu sistema dedutivo de frases e regras simbólicas. Para que o sistema dedutivo da lógica de primeira ordem desempenhe o papel esperado não é aceitável usar o infinito em lugar algum no processo de constituição do mesmo e tampouco no correspondente processo de redução do raciocínio matemático normal aos procedimentos simbólicos desse veículo formal. Ou seja, constituir o sistema finitariamente é apenas o primeiro passo. Devemos também mostrar que o sistema funciona adotando a mesma restrição metodológica sobre o infinito. Para isso, há que se mostrar que as passagens do raciocínio matemático em geral estão representadas no sistema e ao mesmo tempo que o sistema não é capaz de deduzir frases além da conta. Essa última propriedade está fortemente ligada à possibilidade de demonstrar construtivamente a consistência do sistema.

O que seria dizer de um sistema de representação simbólica da matemática que ele não deduz frases além da conta? Uma primeira resposta seria que para a classe das frases do sistema que admitem de modo uniforme um significado finitário há uma demonstração construtiva que se uma frase dessa classe é deduzida no sistema, então seu significado finitário é verdadeiro, caso em que dizemos que o sistema é finitariamente correto com relação à classe. Uma análise nesse sentido é apresentada a seguir.

Vamos supor que M é uma imagem formal de parte da matemática, um sistema simbólico tal que parte das passagens do raciocínio matemático estão representadas em M. Sabemos que M é consistente se e somente se há uma frase do sistema que não é dedutível. Outra caracterização da consistência de M é que não há frase A do sistema tal que A é dedutível em M e $\neg A$ também é dedutível em M, pois qualquer frase é consequência tautológica de A e $\neg A$. Para cada frase A, não pode ser que ambas A e $\neg A$ estejam na conta do que é para ser deduzido. De todo modo, fica claro que se M é inconsistente, então M é capaz de deduzir frases além da conta. Isso é importante, mas seria a conversa válida? Será que bastaria que M fosse consistente para garantir que M não é capaz de deduzir o que não é para ser deduzido? A resposta aqui é não. O sistema pode ser consistente e, por exemplo, deduzir uma sentença A que é interpretada finitariamente como falsa. A mera consistência de M não basta.

Uma hipótese mais forte que a mera consistência de M é a hipótese que há uma demonstração construtiva da consistência de M. Seria isso ainda insuficiente? A resposta agora depende do que entendemos como aquilo que não é para ser deduzido. Entre as sentenças de M há a classe daquelas que admitem uniformemente um significado finitário. Agora, qual é essa classe? Vamos estipular, por enquanto, que é a classe das sentenças $\forall$-rudimentares (ver Boolos, p. 262, e Smorynski, p. 823). Tal estipulação implica que a hipótese acima é suficiente, conforme o teorema abaixo.[4]

Suponha que M representa parte suficiente da matemática e que há uma demonstração construtiva da consistência de M. Nesse caso, M é finitariamente correto com relação à classe das sentenças $\forall$-rudimentares.

A classe das sentenças $\forall$-rudimentares é tal que suas sentenças admitem uniformemente um significado finitário. Se A é uma sentença $\forall$-rudimentar, representamos por $\tilde{A}$ o enunciado finitário correspondente. Seja A uma sentença $\forall$-rudimentar dedutível em M. Então, há uma demonstração construtiva que A é dedutível em M, que consiste em exibir a dedução de A em M. Por outro lado, $\neg A$ é uma sentença $\exists$-rudimentar, e há uma demonstração construtiva que se $\neg A$ não é dedutível em M, então $\tilde{A}$ (ver Bo-

[4] O resultado encontra-se também em Smorynski, p. 824.

ISSN: 2317-9570

olos p.267). Pela hipótese do teorema, há uma demonstração construtiva que pelo menos uma entre A e $\neg A$ não é dedutível em M. Portanto, há uma demonstração construtiva de $\tilde{A}$.

O teorema acima não pode ser ampliado para abarcar a classe das sentenças $\exists$-rudimentares. De fato, considere o sistema Q da aritmética minimal definido em Boolos, p. 266. Há uma demonstração construtiva, e até mesmo finitária,[5] que se Q é consistente, então a sentença de Gödel para Q não é dedutível em Q. Isso é parte da demonstração do primeiro teorema de Gödel da incompletude (ver Boolos, p. 288). Há também uma demonstração finitária da consistência de Q, como veremos, logo da não-dedutibilidade de G em Q. Há ainda, apenas com métodos finitários, uma demonstração que se G não é dedutível em Q, então o sistema $Q[\neg G]$, obtido a partir de Q pela adição de $\neg G$ como axioma, é consistente. Portanto, há uma demonstração finitária da consistência de $Q[\neg G]$. Como o enunciado finitário $\tilde{G}$ correspondente a G diz que G não é dedutível em Q, o enunciado finitário correspondente a $\neg G$ diz que G é dedutível em Q. Concluímos que $\neg G$ não é finitariamente correta, apesar de haver uma demonstração finitária da consistência do sistema $Q[\neg G]$ onde $\neg G$ é dedutível. Como $\neg G$ neste contraexemplo é $\exists$-rudimentar, o teorema acima não se aplica a essa classe de sentenças.

Portanto, uma demonstração construtiva de consistência é suficiente para legitimar o uso de M para demonstrar sentenças $\forall$-rudimentares, mas insuficiente para o uso correspondente às $\exists$-rudimentares. Contudo, há outras consequências interessantes de uma demonstração construtiva de consistência. Uma dessas consequências é que, para demonstrar construtivamente a consistência de um sistema é preciso entender a estrutura fina do funcionamento do sistema. Vamos tentar explicar isso em termos do esquema de quatro etapas da seção anterior.

Uma demonstração construtiva de consistência de um sistema M que segue as quatro etapas listadas deve apresentar uma construção que mostra como a propriedade $\mathcal{P}$ se aplica às fórmulas dedutíveis em M que estão na classe Δ. Além disso, uma tal demonstração deve mostrar uma fórmula em Δ para a qual a propriedade construtiva $\mathcal{P}$ não se aplica. Como

[5] A distinção entre demonstração finitária e demonstração meramente construtiva que adotamos aqui é explicada em Shoenfield, p. 214.

já foi dito, a escolha de uma propriedade $\mathcal{P}$ adequada para essa construção demanda um entendimento construtivo da semântica de M com as fórmulas dedutíveis em M que estão na classe Δ, ou seja, um entendimento de que tipo de propriedade construtiva se aplica a elas. Vamos ilustrar o ponto com um exemplo simples na seção seguinte.

3 - A Demonstração da Consistência da Lógica de Primeira Ordem

Considere M um sistema dedutivo para a lógica clássica de primeira ordem sem igualdade. É possível demonstrar finitariamente que M é consistente, ou seja que alguma fórmula não é dedutível em M. A ideia da demonstração é fácil de explicar nesse caso. Basta observar que as fórmulas dedutíveis em M são válidas em todas as estruturas não-vazias. Em particular as fórmulas dedutíveis em M são válidas nas estruturas com um único indivíduo no domínio, e a validade em todas as estruturas com um único indivíduo no domínio pode ser caracterizada como uma propriedade finitária. Vejamos como essa ideia pode ser transformada em uma demonstração.

Seja A uma fórmula qualquer e b um nome próprio novo, e seja A^* a fórmula obtida a partir de A pela omissão de todos os quantificadores e substituição de todos os termos restantes por b. Se A é dedutível a partir de M, então A^* é uma tautologia.

Se A é dedutível em M, então A é um axioma do tipo $B_x[a] \to \exists xB$, ou é consequência tautológica de fórmulas dedutíveis, ou é do tipo $\exists xB \to C$, em que x não ocorre livre em C e $B \to C$ é dedutível. Podemos demonstrar por indução que A^* é uma tautologia em qualquer dos três casos acima.

Primeiro, o passo base. Se A é um axioma do tipo $B_x[a] \to \exists xB$, então A^* é $B^* \to B^*$, e isso é uma tautologia. Agora o passo indutivo. Se A é consequência tautológica de fórmulas dedutíveis B_1, ..., B_n, então A^* é consequência tautológica de B_1^*, ..., B_n^*. Por hipótese de indução, B_1^*, ..., B_n^* são tautologias, de onde segue que A^* é tautologia. Por outro lado, suponha que A é $\exists xB \to C$, em que x não ocorre livre em C e $B \to C$ é dedutível. Nesse caso, A^* é $B^* \to C^*$, e é tautologia por hipótese de indução.

A consistência do sistema M é consequência do teorema acima. De fato, a fórmula $B \wedge \neg B$ não é dedutível pois $(B \wedge \neg B)^*$ é $B^* \wedge \neg B^*$, que não é uma tautologia. Temos, portanto, uma demonstração finitária da consistência da lógica de primeira ordem. Vamos analisar

ISSN: 2317-9570

esta demonstração em detalhes.

Primeiro, identificamos a propriedade finitária $\mathcal{P}$. Afirmar que A satisfaz $\mathcal{P}$ é afirmar que sua transformada A^* é uma tautologia. Observamos que, assim definida, $\mathcal{P}$ é uma propriedade finitária equivalente à validade em todas as estruturas com um único indivíduo no domínio. Vamos demonstrar, por indução na complexidade de A, que A é satisfeita em qualquer estrutura com um único indivíduo se e somente se A^* é uma tautologia, ou seja, verdadeira segundo o método usual das tabelas de verdade para qualquer atribuição de valor de verdade para suas subfórmulas atômicas.

Sejam p_1, ..., p_m símbolos de predicado e $\mathcal{N}$ uma estrutura com um único indivíduo α e que interpreta esses símbolos. Há exatamente duas opções para a interpretação de cada símbolo de predicado p_i, com $1 \leq i \leq m$: Ou a interpretação de p_i é a extensão vazia, ou é a extensão total $\{\langle\alpha, ..., \alpha\rangle\}$, determinada pela aridade de p_i, correspondendo naturalmente é atribuição de falso ou de verdadeiro. Desse modo, a estrutura $\mathcal{N}$ corresponde a uma linha da tabela de verdade para m símbolos proposicionais.

Com isso, para qualquer fórmula A cujos símbolos de predicado estão entre p_1, ..., p_m, temos que A é satisfeita em $\mathcal{N}$ se e somente se A^* é verdadeira na linha da tabela de verdade correspondente a $\mathcal{N}$. Se A é atômica, da forma $p_i(t, u, v, ...)$, então A^* é $p_i(b, b, b, ...)$. Se a interpretação de p_i é a extensão total, então, como b e t, u, v, ... denotam α, tanto A quanto A^* são satisfeitas em $\mathcal{N}$ e A^* é verdadeira na linha correspondente. Se a interpretação de p_i é a extensão vazia, então A e A^* não são satisfeitas e A^* é falsa na linha correspondente. Se A é $\neg B$ ou é $B \vee C$, então A^* é $\neg B^*$ ou é $B^* \vee C^*$ e o resultado segue da hipótese de indução nesses casos. Finalmente, se A é $\exists x B$, então A^* é B^*. Agora, A é satisfeita em $\mathcal{N}$ com seus termos denotando α se e somente se B é satisfeita em $\mathcal{N}$ quando seus termos denotam α. Por hipótese de indução, B é satisfeita em $\mathcal{N}$ quando seus termos denotam α se e somente se B^* é verdadeira na linha correspondente a $\mathcal{N}$, e o resultado segue do fato que A^* é B^*.

Concluímos, a partir do parágrafo acima, que A é válida em todas as estruturas apropriadas com um único indivíduo se e somente se A^* é uma tautologia. Tal fato não é usado na demonstração de consistência, mas explica de onde surge a ideia subjacente. É importante se perguntar por que usamos a propriedade $\mathcal{P}$, "a transformada é uma tautologia", no lugar da propriedade "ser válida nas estruturas com um único indivíduo". Acontece que a noção geral de validade em estruturas não é finitária, por exemplo, a cláusula sobre a interpretação da quantificação existencial em uma estrutura não é finitária. Por isso, precisamos primeiro eliminar o elemento infinitário, a interpretação da quantificação, para apresentar a propriedade $\mathcal{P}$ no âmbito finitário. Sem a eliminação dos quantificadores isso não seria possível. Argumentamos que trata-se de um padrão: é necessário eliminar a quantificação para realizar as etapas de uma demonstração construtiva de consistência.

4 - A Consistência da Aritmética de Primeira Ordem

As etapas das demonstrações de consistência que analisaremos aqui também são articuladas a partir de uma noção de validade. Tal noção deve por um lado ser adequadamente restrita para admitir uma caracterização finitária ou pelo menos construtiva, e por outro lado ampla o suficiente para acomodar a correção de uma classe apropriada de fórmulas.

Vamos começar analisando a demonstração da consistência de um sistema para a aritmética sem o axioma esquema da indução, o sistema N apresentado em Shoenfield, página 22. Trata-se de uma teoria aberta, ou seja, tal que os axiomas não-lógicos são livres de quantificadores. Nesse caso, como consequência do primeiro teorema epsilon[6], se uma fórmula A livre de quantificadores é dedutível em N então A é consequência tautológica de instâncias de axiomas também livres de quantificadores de N. A partir desse resultado, uma demonstração finitária da consistência de N pode ser obtida e inserida nas quatro etapas do nosso esquema.

Primeiro, considere $\mathcal{P}$ a propriedade de validade finitária definida como: Uma fórmula A é finitariamente válida se e somente se o valor de verdade de qualquer instância fechada de A é computável e verdadeiro. Dizemos que o valor de verdade de qualquer instância fechada de A é computável e verdadeiro quando o procedimento uniforme que calcula o valor de verdade de uma dada instância de A segundo as tabelas de verdade para os conectivos lógicos e os algoritmos para sucessor, soma, produto, relação de igualdade e relação de ordem dá como resultado o valor verdadeiro. Observamos que os axiomas abertos de N, o que inclui os axiomas não-lógicos e os axiomas da igualdade,

[6]Veja a terceira seção do verbete https://plato.stanford.edu/entries/epsilon-calculus/

são finitariamente válidos.

Segundo, seja Δ a classe das fórmulas abertas, ou seja, livres de quantificadores. Como vimos acima, se uma fórmula A em Δ é dedutível em N, então A é consequência tautológica de instâncias de axiomas abertos de N. Mas uma consequência tautológica de fórmulas finitariamente válidas é também finitariamente válida. As instâncias dos axiomas abertos de N são finitariamente válidas, e disso segue que as fórmulas abertas dedutíveis em N são todas finitariamente válidas. Ou seja, que vale a correção com relação a $\mathcal{P}$ das fórmulas em Δ que são dedutíveis em N. Finalmente, a fórmula aberta $\neg x = x$ não é finitariamente válida, portanto não é dedutível em N, o que mostra sua consistência.

A demonstração acima tem dois pontos cruciais. O primeiro é que há um procedimento uniforme, com base nas tabelas de verdade e nos algoritmos para as operações e relações aritméticas básicas, para calcular o valor de verdade de qualquer sentença livre de quantificadores escrita na linguagem de N. O segundo, é o uso do primeiro teorema epsilon, que estabelece a correção com relação a $\mathcal{P}$ das fórmulas em Δ que são dedutíveis em N. A demonstração deste teorema é a parte mais difícil, que consiste em eliminar quantificadores. Notamos que a possibilidade de eliminar a quantificação é o elemento comum de ambos os pontos, sem o qual não alcançaríamos a propriedade de validade finitária ou a correção.

Passemos ao caso da aritmética de primeira ordem com indução. Novamente seguimos a referência Shoenfield, em que a aritmética com indução é apresentada na página 204. Vamos denotar esse sistema por PA. Como o axioma esquema da indução não é livre de quantificadores, os axiomas de PA não são finitariamente válidos no sentido acima. Para demonstrar a consistência de PA é preciso apresentar uma propriedade $\mathcal{P}$, demonstrar a correção com relação a $\mathcal{P}$ das fórmulas em uma classe Δ dedutíveis em PA, e encontrar uma fórmula em Δ que não satisfaz $\mathcal{P}$. Para uma demonstração finitária, a propriedade $\mathcal{P}$ deveria ser finitária, assim como a demonstração da correção das fórmulas em Δ dedutíveis e da incorreção de alguma fórmula em Δ. Infelizmente, o segundo teorema de Gödel da incompletude faz parecer duvidoso que isso seja possível. Contudo, há demonstrações de consistência da aritmética de primeira ordem com indução que vão pouco além do finitário. A mais conhecida é devida a Gentzen, mas o próprio Gödel produziu uma demonstração assim, que foi apresentada como uma reformulação da demonstração de Gentzen.

Vamos analisar uma variante da demonstração de Gödel da consistência de PA devida a Shoenfield, que apresenta uma interpretação do sistema PA para a aritmética com indução diretamente em uma teoria de funcionais recursivos de tipo finito.[7] A linguagem da teoria não tem quantificadores, mas tem uma quantidade enumerável de tipos, e estoques de pronomes para cada um desses tipos. A definição dos termos e constantes (símbolos para funcionais) é indutiva, de forma a dar esquemas de introdução de constantes a partir das constantes básicas 0 e S (intencionalmente "zero" e "sucessor") e da descrição dos tipos. Além de esquemas puramente combinatórios, há um esquema de introdução de símbolos para funcionais por recursão. As fórmulas atômicas são equações apenas entre termos de tipo o, que intencionalmente se referem a números. As demais fórmulas são obtidas por negações e disjunções apenas.

A interpretação de PA é então definida associando para cada fórmula A do sistema uma fórmula transformada A^* do tipo $\forall\overline{x}\exists\overline{y}\phi$, em que ϕ é uma fórmula da teoria descrita acima e $\overline{x}$ e $\overline{y}$ são sequências de pronomes. Exatamente como no caso da demonstração da consistência da lógica de primeira ordem, uma fórmula A e sua transformada A^* possuem o mesmo significado pretendido, o que não é usado na demonstração. No próximo passo, um teorema de correção é demonstrado relacionando a demonstrabilidade de A com a validade da respectiva fórmula transformada A^* na teoria dos funcionais. A demonstração do teorema de correção é construtiva, e dele segue o resultado de consistência.

O único elemento não-finitário da demonstração de Gödel da consistência de PA é a noção de validade da teoria dos funcionais. Mas essa noção não está tão distante do finitário quanto, por exemplo, a noção de validade padrão da aritmética de primeira ordem, esta sim fortemente não-finitária. O motivo para isso é que não há fórmulas quantificadas na teoria dos funcionais, portanto a noção de validade associada não inclui cláusulas para quantificadores. Na verdade, a única cláusula relevante é aquela que corresponde às fórmulas atômicas, equações entre termos de tipo o. A semântica dessas equações não

[7] Vamos apresentar apenas um breve esboço da teoria de funcionais e da demonstração de consistência suficiente para nossos propósitos. Para mais detalhes recomendamos Shoenfield, pp. 214 - 222.

é finitária porque símbolos para funcionais de tipo superior podem ocorrer nos termos de tipo o, o que torna a interpretação do termo geral de tipo o não-finitária. Tal elemento semântico está presente na demonstração de consistência e parece que não pode ser removido. Por outro lado, as relações entre funcionais recursivos de tipo finito parecem suficientemente determinadas e, como a semântica do termo geral de tipo o não vai muito além do âmbito finitário, a consistência de PA baseada na demonstração com teoria dos funcionais é plausível.

A demonstração que acabamos de descrever é facilmente colocada no esquema geral de demonstrações de consistência. Primeiro, dizemos que uma fórmula A de PA tem a propriedade $\mathcal{P}$ se a fórmula A^* correspondente, que é do tipo $\forall\overline{x}\exists\overline{y}\phi$, é válida no sentido que há um procedimento finitário e uniforme que produz termos $\overline{y}$ a partir de termos $\overline{x}$ de modo que a instância associada de ϕ seja válida no sentido da teoria dos funcionais. Depois, tomando como Δ o conjunto de todas as fórmula, demonstramos finitariamente a correção com relação é $\mathcal{P}$ de todas as fórmulas dedutíveis. Por fim, demonstramos que a fórmula $\neg 0 = 0$ não possui a propriedade $\mathcal{P}$. De todas essas etapas, apenas a definição da propriedade $\mathcal{P}$ apresenta algum elemento não-finitário.

Podemos dizer que essa demonstração é construtiva apesar de não ser estritamente finitária. Novamente, temos aqui dois pontos cruciais. O primeiro é o estabelecimento da teoria dos funcionais e da noção de validade nessa teoria, a partir da qual a propriedade $\mathcal{P}$ é definida. O segundo é demonstrar a correção com relação a $\mathcal{P}$ das fórmulas são dedutíveis em PA. Para isto é preciso mostrar, com um procedimento finitário e uniforme, como produzir para toda fórmula dedutível A termos apropriados a partir de termos dados, de modo que a instância associada da matriz da fórmula transformada A^* seja válida no sentido da teoria dos funcionais. Notamos também que a eliminação da quantificação está presente em cada parte da demonstração pois sem isso não alcançaríamos a propriedade "quase-finitária" $\mathcal{P}$ ou a correção.

A partir de nossa análise podemos destacar como elemento comum nas demonstrações de consistência tanto a eliminação de quantificadores como a observação de que a propriedade $\mathcal{P}$ é derivada de uma semântica padrão. Os dois pontos estão relacionados e não sem razão. Uma vez que uma demonstração de consistência é obtida a partir da correção de algumas fórmulas dedutíveis com relação a $\mathcal{P}$, é natural derivar a propriedade $\mathcal{P}$ de uma semântica padrão para a qual já temos correção. Contudo a principal componente não-finitária de uma semântica padrão é a cláusula correspondente à quantificação. Portanto, para derivar de uma semântica padrão uma propriedade $\mathcal{P}$ que seja pelo menos próxima do âmbito finitário, é preciso eliminar quantificadores. No caso de PA isso aparentemente não pode ser feito sem custo e a quantificação é substituída por algo ainda não-finitário, como os funcionais de tipo superior na demonstração analisada.

É razoável que o interesse na busca por demonstrações de consistência construtivas para sistemas como PA persista apesar do segundo teorema da incompletude de Gödel, que impede demonstrações de consistência que sejam formalizáveis em tais sistemas. A análise dos casos acima mostra que eliminar quantificadores substituindo-os por algo cuja interpretação seja mais próxima do finitário é parte importante da estratégia de demonstração de consistência dentro do quadro delimitado por Gödel. As etapas do esquema geral proposto são cumpridas construtivamente apenas após alguma eliminação de quantificadores. Isso ajuda a explicar a dificuldade de demonstrar construtivamente a consistência de um sistema como ZFC, para o qual parece muito difícil substituir a quantificação por algo mais próximo do finitário e mais facilmente interpretado. No entanto, o esforço para avançar nas demonstrações construtivas de consistência frequentemente leva a um melhor entendimento sobre a estrutura quantificacional do sistema e sua interpretação.

Ainda, observamos que tais demonstrações, além do seu interesse fundacional, enriquecem a disciplina da lógica matemática enquanto método para atacar problemas matemáticos, tendo como fruto a área conhecida como teoria da demonstração. Como corolário da demonstração de consistência de PA e da demonstração do Teorema de Herbrand temos, por exemplo, um resultado devido a Kreisel que estabelece que para bloquear potenciais contraexemplos para uma fórmula fechada em forma normal prenexa que é dedutível em PA basta considerar os funcionais recursivos. Esse resultado já foi usado para extrair informação construtiva de demonstrações matemáticas não-construtivas. Os recentes avanços desta área podem ser encontrados em Kohlenbach.

ISSN: 2317-9570

5 - Considerações Finais Sobre Consistência e Verdade Aritmética

Podemos considerar a relação entre consistência e verdade na aritmética sob o prisma do estudo de casos acima. Em linhas gerais, dizemos de uma sentença declarativa que ela é verdadeira com relação a um padrão de correção se uma concordância entre a sentença e o padrão é obtida. Caso contrário, dizemos que a sentença é falsa com relação ao padrão de correção instituído. Nesse sentido, um discurso sobre verdade de sentenças da aritmética deve vir acompanhado de uma elucidação dos elementos centrais em torno dos quais esta concepção geral se articula. Logo, antes de falar em verdades aritméticas precisamos explicar o que é uma sentença (declarativa), que tipo de coisa é um padrão de correção, como um padrão é instituído e em que consiste a relação de concordância entre sentença e padrão.

Outros tipos de sentença podem ser considerados, mas vamos nos ocupar aqui apenas com o discurso sobre verdade de sentenças em uma linguagem de primeira ordem, como a linguagem de PA. De um modo geral, o problema de definir precisamente as sentenças de uma teoria matemática não oferece grande dificuldade. É muito mais difícil explicar satisfatoriamente que tipo de coisa é um padrão de correção em uma teoria matemática, como ele é instituído e em que consiste a relação de concordância entre sentença e padrão. Uma primeira tentativa seria a concepção de verdade como validade em uma classe de estruturas: Um padrão de correção em uma teoria matemática é uma classe de estruturas formada pelos modelos padrão da teoria; esse padrão é instituído pela intenção do matemático de falar dessas estruturas ao formular sua teoria, e uma sentença concorda com esse padrão se é satisfeita em todas as estruturas da classe.

Um esclarecimento deve ser feito sobre o que entendemos por estrutura, no nosso caso, estruturas da aritmética de primeira ordem. Uma tal estrutura é constituída por um domínio qualquer de indivíduos, uma relação binária (nesse domínio) que interpreta o símbolo de igualdade =, um indivíduo que interpreta o símbolo 0 e duas operações binárias que interpretam os símbolos +, ., para adição e multiplicação. Reforçamos que não há restrição adicional para os dados de uma estrutura, ou seja, o domínio, a relação binária e as operações. É usual pedir que o número de indivíduos no domínio seja maior que zero e que a relação que interpreta igualdade seja a relação de identidade no domínio, mas mesmo esses requerimentos podem ser dispensados. Claro que nem todas as estruturas são modelos padrão intencionados.

Dado o esclarecimento, ao retornar à tentativa, observamos os seguintes problemas. Não é claro que ao formular uma teoria a intenção do matemático de se dirigir aos modelos padrão é suficiente, ainda que esses modelos sejam entendidos por meio de uma metateoria. Não basta ter uma intenção para garantir que as sentenças da teoria adquiram um sentido específico. Estruturas que servem de base para a interpretação da linguagem são usualmente entendidas através de uma teoria matemática. Nesse caso, o entendimento proposto sobre verdade de sentenças em uma teoria matemática só poderia ser dado em outra teoria matemática, uma metateoria, capaz de definir tanto as fórmulas da teoria objeto quanto as estruturas apropriadas. Um discurso sobre verdade de sentenças aritméticas acaba por ser realizado em uma metateoria, portanto mediado por outra teoria matemática. Tal mediação é indesejada pois a aritmética é uma teoria matemática básica e parece natural que possamos entender sua verdade sem ter que recorrer à outra teoria.

Se tentamos resolver esse problema supondo que o entendimento de modelos padrão intencionados prescinde de uma metateoria, então a situação se agrava por outro lado. Pois teríamos agora dois tipos de entendimento dos modelos padrão, um mediado pela metateoria e outro direto. E assim, ou esse entendimento direto dos modelos padrão concorda com as representações teóricas destes, ou o conhecimento matemático sobre os modelos padrão estaria errado. Como a segunda alternativa não é admissível se temos como hipótese que o conhecimento matemático sobre os modelos não pode estar errado, temos que nos comprometer com a concordância entre o entendimento direto de modelos e o entendimento por meio de uma metateoria. Mas tal comprometimento parece inadequado e poderia trazer consequências posteriores, portanto preferimos buscar outra alternativa.

Portanto, vamos trabalhar em detalhes uma proposta alternativa para a relação de verdade na aritmética. A finalidade dessa nova proposta é eliminar a referência à classe de modelos padrão preservando o ideal almejado. Uma teoria matemática, segundo a proposta acima, *dirige-se*, através da intenção do matemático, à classe de modelos padrão, almejando,

ISSN: 2317-9570

com isso, alcançar exatamente as sentenças satisfeitas em todos os modelos padrão. A chave para a formulação da nova proposta é o entendimento de que essa *direção* que a teoria possui, ou seja, a noção que a teoria almeja alcançar as sentenças satisfeitas em todos os modelos padrão, pode (i) ser caracterizada sem referência à classe de modelos padrão e (ii) desempenhar o papel de padrão de correção. Se isso for o caso, então os problemas mencionados acima não atingem a nova proposta, enquanto os méritos da concepção de verdade como validade em uma classe de estruturas são mantidos.

Há um problema sério com a tese segundo a qual uma classe de modelos padrão tem prioridade sobre o critério objetivo de seleção de estruturas correspondentes. Essa tese toma por base que pelo menos uma classe de modelos padrão de uma teoria matemática está determinada em primeiro lugar, e, que a axiomatização da teoria é uma articulação subordinada aos modelos padrão que é obtida na tentativa de descrevê-los. Isso pode ser plausível para algumas teorias como, por exemplo, a aritmética, em que há uma representação simples e precisa dos indivíduos, predicados e operações de um modelo padrão. Contudo, o mesmo não ocorre para outras teoria, como a teoria de conjuntos. Não há representação simples, que não assuma de antemão a própria teoria de conjuntos e que determina exaustivamente um modelo dessa teoria, nem é o caso que uma classe de modelos padrão da teoria de conjuntos está determinada (em uma metateoria) e antecede sua axiomatização. Como acreditamos ser importante alcançar um entendimento da noção de verdade que seja uniforme para as teorias matemáticas, esse problema deve ser levado em consideração aqui.

Partimos, então, da ideia que se uma teoria matemática dirige-se a uma classe de modelos padrão, então essa direção que a teoria possui pode (i) ser caracterizada sem referência à classe de modelos padrão e (ii) desempenhar o papel de padrão de correção. A propriedade (i) pode ser entendida do seguinte modo: Cada classe de modelos padrão é caracterizada por um critério objetivo de seleção de estruturas que pode ser formulado como uma lista de princípios, os *princípios diretivos* correspondentes à classe dada. Ou seja, cada classe de modelos padrão pode ser vista como a classe das estruturas que estão em conformidade com um critério correspondente. Contudo, essa formulação sugere que a classe de modelos padrão continua desempenhando o papel principal, enquanto os princípios diretivos são secundários. Nós vamos inverter a ordem de prioridade, dando aos princípios diretivos o papel principal. Assim não é preciso se comprometer ontologicamente com a classe de modelos padrão para apresentar o critério que deveria ser satisfeito por cada um deles, pois não precisamos de modelos padrão para prescrever o que seria um.

Consideramos, portanto, que uma lista de princípios diretivos tem prioridade sobre a classe de modelos padrão correspondente, e a axiomatização de uma teoria matemática deve ser guiada por tal lista. Um padrão de correção primário para uma teoria matemática é a lista de princípios diretivos correspondente, a classe de modelos padrão correspondente deve ser entendida como um padrão de correção secundário e apenas na medida em que representa os princípios. Esse padrão é instituído na prática histórica da matemática, na medida em que os princípios são adotados, implícita ou explicitamente, e regem essa prática.

Vamos apresentar agora uma lista de princípios diretivos para a aritmética e, em seguida, uma explicação do papel da mesma na aritmética. Essa lista de princípios constitui uma direção a ser seguida pela teoria. Podemos dizer que a prática da aritmética está sujeita à adoção de princípios diretivos instituídos historicamente, como aqueles apresentados aqui, e que essa prática regida por princípios antecede a axiomatização.

1. Cada número é denotado por um único numeral, sendo este um objeto sintético obtido pela repetição, possivelmente nula, de um símbolo primitivo. Cada numeral denota um único número e o número zero é denotado pelo numeral nulo.

2. Dados dois numerais s e t, a adição dos números denotados por s e t é denotada pelo numeral obtido pela repetição do símbolo primitivo determinada pela concatenação de t e s.

3. Dados dois numerais s e t, a multiplicação dos números denotados por s e t é denotado pelo numeral obtido pela repetição de s determinada por t, ou seja, por uma repetição de s para cada ocorrência do símbolo primitivo em t.

A lista de princípios acima definitivamente não é uma descrição de um modelo padrão dos "números verdadeiros"; ela apenas prescreve a direção seguida pela aritmética, ou seja, o que essa teorização almeja descrever. Esta lista prescreve o que um domínio qualquer de objetos com operações deve obrigatoriamente satisfazer para poder ser considerado um modelo padrão da aritmética, uma estrutura que a aritmética almeja descrever. É importante ressaltar que não faz sentido falar em valor de verdade do critério prescrito por essa lista de princípios uma vez que esta lista não desempenha função descritiva para ser verdadeira ou falsa. Podemos apenas dizer que os princípios diretivos foram instituídos historicamente pela relevância que a investigação na direção apontada apresenta, e que, uma vez instituídos, prescrevem objetivamente a direção da aritmética.

Os axiomas da aritmética devem ser escolhidos de forma que o sistema dedutivo resultante seja uma aproximação, tão boa quanto possível, das sentenças verdadeiras de acordo com o padrão de correção dado pelos princípios diretivos. A intenção de satisfazer tal aproximação resulta em uma escolha de um sistema dedutivo que seja um correlato formal da lista de princípios. Acreditamos que este é o modo pelo qual chegamos aos axiomas do sistema dedutivo PA, mesmo que implicitamente. Para o caso do axioma esquema da indução é válida uma explicação adicional: Esse axioma é obtido a partir do primeiro princípio, na tentativa de descrever uma situação em que os números são todos denotados por numerais.

A lista de princípios não constitui um âmbito externo à aritmética, e sim faz parte da aritmética enquanto sistema teórico interpretado. A aritmética, assim compreendida, é constituída por duas camadas: A camada do critério dado pela lista de princípios, em que é dada a direção da teoria, e a camada dos sistemas formais para a aritmética, que deve estar em concordância com os princípios e que constitui o âmbito dedutivo associado. Uma sentença formal da aritmética se encontra na segunda camada definida acima, e dizemos que ela é verdadeira se descreve corretamente a direção prescrita pelos princípios dados na primeira camada. Essa noção de verdade aritmética pode ser analisada matematicamente, e o resultado da análise matemática corresponde com o esperado. Para um tratamento mais técnico e abrangente desta proposta veja Freire.

Um fato relevante sobre a noção de verdade aritmética que acabamos de estabelecer é que ela não implica a consistência dos sistemas formais correspondentes. Mesmo que para alguns sistemas a demonstração de consistência seja plausível, uma concepção de verdade de teorias matemáticas deve ser uniforme, como dissemos anteriormente, e se aplicar igualmente às teorias para as quais não temos tais demonstrações. Se o sistema formal *PA* é inconsistente, então, em virtude do mesmo fato, o critério estabelecido pela lista de princípios é necessariamente vazio, e não aponta para direção alguma. Isso não ocorre com a concepção de verdade como satisfação em uma estrutura padrão que prescinde de uma metateoria, porque se essa estrutura satisfaz *PA*, então *PA* é consistente. Julgamos que essa característica da concepção de verdade é desejável, pois, tendo em vista as limitações para demonstrações de consistência expostas na seção anterior, faz sentido deixar em aberto a possibilidade da inconsistência.

Referências

BOOLOS, George, Burgess, John, Jeffrey, Richard, *Computabilidade e Lógica*, Tradução de Cezar Mortari, São Paulo: Editora Unesp, 2012.

BOURBAKI, Nicolas, *Théorie des Ensembles*, Paris: Hermann, 1970.

FREIRE, Rodrigo, Interpretation and Truth in Set Theory, em *Trends in Logic: Vol. 47, Contradictions, from Consistency to Inconsistency*, Berlim: Springer, 2018.

HILBERT, David, On the Infinite, em *Philosophy of Mathematics: Selected Readings*, editado por Paul Benacerraf e Hilary Putnam, Segunda Edição, Cambridge University Press, 1983, pp. 183 - 201.

KLEENE, Stephen, *Introduction to Metamathematics*, Amsterdam: North-Holland Publishing Company, 1952.

KOHLENBACH, Ulrich, *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*, Berlim: Springer-Verlag, 2008.

RAGGIO, Andres, A Evolução da Noção de Sistema Axiomático, *Philosophos - Revista de Filosofia*, vol. 8, no. 1, 2003, pp. 95 - 119.

SHOENFIELD, Joseph, *Mathematical Logic*, Reading, Massachusetts: Addison-Wesley Publishing Company, 1967.

SMORYNSKI, Craig, The Incompleteness Theorems, em *Handbook of Mathmatical Logic*, editado por Jon Barwise, Amsterdam: North-Holland, 1977, pp. 821 - 865.

TARSKI, Alfred, Mostowski, Andrzej, Robinson, Raphael, *Undecidable Theories*, Amsterdam: North-Holland Publishing Company, 1953.

VIERO, Arno, *Sistemas Axiomáticos Formalizados: A Questão da Desinterpretação e da Formalização da Axiomática*, Campinas: Coleção CLE, 2011.

ISSN: 2317-9570

Análise de uma Fundamentação da Verdade de Sentenças Aritméticas

[Analysis of Grounds for the Truth of Arithmetical Sentences]

Edgar L. B. de Almeida*

Resumo: O tema deste trabalho é a verdade de proposições matemáticas e seu objetivo é avaliar, no contexto aritmético, um dos elementos presentes em Freire (2018) e também considerado por Ramos e Freire nesta edição da Revista de Filosofia Moderna e Contemporânea: a fixação do valor de verdade de asserções aritméticas a partir de princípios diretivos que regem a prática da disciplina. O método de análise visa à elucidação da contribuição dos princípios diretivos para a fixação do modelo padrão da aritmética e considera três métricas distintas. A partir desses resultados a proposta fundada nos princípios diretivos é comparada com três abordagens alternativas, presentes na literatura especializada. O resultado dessa comparação é bastante favorável à abordagem formulada por Freire e Ramos nesta Revista.
Palavras-chave: aritmética, modelo padrão, condições de verdade.

Abstract: The main subject of this work is the truth of mathematical assertions and its aim is to evaluate, in the arithmetical context, one of the elements featured by Freire in (2018) and in this issue of the Revista de Filosofia Moderna e Contemporânea (in joint work with Ramos): a strategie to fix the truth-condition of arithmetical propositions based on the directive principles that govern the practice of this matter. The method of analysis aims to elucidate the contribution of the directive principles to fix the standard model of arithmetics and takes in consideration three differents metrics. From this results the approach based in principles is compared with other three proposals well known in the literature. The result of this comparison is far favorable to Freire and Ramos' approach in this issue.
Keywords: arithmetic, standard model, truth-condictions.

1 - Introdução

O escopo deste trabalho - análise de uma proposta de fixação do valor de verdade das sentenças aritméticas - é caro à filosofia da matemática e será investigado com o suporte de sistemas axiomáticos. Contemporaneamente esses sistemas desempenham papel de relevo em discussões de cunho filosófico-matemático pois atuam como ferramentas de des-

*Professor do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo e Pesquisador Colaborador do Departamento de Filosofia da Universidade de Brasília. E-mail: edgar.almeida@ifsp.edu.br

DOI: https://doi.org/10.26512/rfmc.v6i2.22096

crição e interpretação da matemática.

Sistemas axiomáticos possuem dois componentes: sistemas formais e modelos. Modelos são de natureza semântica-veritativa, enquanto sistemas formais são de natureza sintático-dedutiva. Grosso modo, os modelos codificam as condições de verdade das proposições do ramo da matemática descrito pelo sistema axiomático e os sistemas formais codificam a linguagem, os métodos de dedução e os termos primitivos da área sob investigação axiomática.

Neste contexto, fixado um sistema formal, há duas possibilidades para os modelos: ou há um único modelo - a menos de isomorfismos - ou há uma infinidade de modelos não isomorfos. Quando ocorre o primeiro caso, diz-se que o sistema formal é *interpretado* e, no outro caso, que trata-se de um sistema formal *não interpretado*. Essa nomenclatura é estendida, de modo natural, também aos sistemas axiomáticos: um sistema axiomático é interpretado quando seu sistema formal é interpretado. Grupos são um exemplo canônico de sistema axiomático não interpretado; afinal, toda estrutura que atesta a veracidade dos axiomas de grupo é um modelo genuíno do sistema formal de grupos. O mesmo não ocorre com a aritmética; a posição compartilhada por muitos matemáticos e filósofos é que os sistemas axiomáticos da aritmética são interpretados: há, a menos de isomorfismos, um *único* modelo para a aritmética. Este modelo é o *modelo padrão* da aritmética.

O entendimento de que há um modelo padrão para a aritmética acarreta consequências imediatas para a análise da verdade nessa disciplina: se há um modelo privilegiado, então as condições de verdade das proposições aritméticas estão codificadas nesse modelo. E, uma vez que as condições de verdade das proposições são preservadas nas sentenças que expressam a proposição, o valor de verdade da proposição pode ser aferido pela inspeção recursiva da sentença. Deste modo, se a estipulação do modelo padrão da aritmética estiver alicerçada em bases robustas, então o mesmo poderá ser dito sobre a verdade das proposições aritméticas. Entretanto, é reconhecidamente problemático fornecer bases segundo as quais um modelo da aritmética possa, justificadamente, ser considerado *o* modelo padrão.

Ramos e Freire apresentam neste volume da Revista de Filosofia Moderna e Contemporânea uma posição que, dentre outras consequências, funda a verdade de disciplinas matemáticas interpretadas, isto é, disciplinas codifi-

ISSN: 2317-9570

cadas por sistemas axiomáticos interpretados, em um esquema em "dupla camada". Essa posição também pode ser vista, com maiores detalhes e em outro contexto, em Freire (2018).

Uma das camadas é composta pelo sistema formal e, a outra, pelos princípios diretivos que regem a disciplina. Nesse esquema os modelos não estão alienados das camadas, mas também não são determinados por uma única camada particular. Este trabalho analisa em qual medida os princípios diretivos propostos para a aritmética viabilizam a fixação do modelo padrão da aritmética e, consequentemente, o valor de verdade das sentenças aritméticas. Para isso é conduzida uma análise rigorosa tanto dos princípios diretivos quanto dos métodos de análise dos princípios. É importante destacar que, embora os princípios diretivos não constituam um sistema formal e não sejam substitutos de sistemas formais, são passíveis de investigações lógicas e formais. De fato, é desenvolvida uma análise lógica da contribuição dos princípios diretivos para a fixação do valor de verdade das sentenças aritméticas. A noção de lógica aqui empregada é aquela introduzida por Lindström (1969) e aqui denominada *sistema semântico*.

Uma vez avaliados os princípios diretivos em um sistema semântico, procede-se a avaliação dos sistemas semânticos segundo três métricas, as quais visam quantificar e qualificar aspectos matemáticos, dedutivos e modelo-teóricos envolvidos na análise. O resultado dessas investigações estabelece, de modo claro e preciso, em qual medida a adoção dos princípios diretivos contribui para a fundamentação do modelo pretendido da aritmética. Adicionalmente, a análise fornece um subproduto bastante interessante: uma fundamentação original para verdade dos axiomas da aritmética de Peano. Outra consequência da análise é a sugestão que linguagens infinitárias podem desempenhar papel relevante no estudo de questões fundacionais. Por fim, os resultados aqui obtidos acerca da fixação do modelo padrão da aritmética são comparados com três propostas presentes na literatura especializada - o resultado da comparação é bastante favorável às propostas tecidas nesta Revista.

2 - Análise dos princípios diretivos da aritmética

Um sistema semântico $\mathbb{L}$ é um par $(L^{\Sigma}, \models)$ em que a primeira componente é um conjunto de sentenças em uma assinatura Σ e a segunda componente é a relação de satisfatibilidade entre a classe das estruturas adaptadas para a assinatura e as sentenças dessa assinatura. Sistemas semânticos não necessariamente são munidos de uma contraparte formal, isto é, não necessariamente há um sistema formal correlato para a relação de satisfatibilidade. Entretanto, este não é o caso para os sistemas semânticos aqui empregados. O sistema formal correlato a um sistema semântico é composto por linguagem formal, conjunto recursivo de axiomas e coleção finita de regras de inferência. Exemplos de tais sistemas semânticos são as formulações canônicas da lógica de primeira ordem, da lógica de segunda ordem e da lógica infinitária que admite sentenças formadas por disjunções enumeráveis e ocorrência de quantificadores apenas em número finito. Tais sistemas serão denotados, respectivamente, $\mathbb{L}_{\omega\omega}$, $\mathbb{L}_2$ e $\mathbb{L}_{\omega_1\omega}$. Apresentações sistemáticas desses sistemas podem ser vistas em Ebbinghaus (1985) e Ebbinghaus, Thomas, Flum (1996).

Formalização é o processo de reescrita de sentenças da linguagem natural em uma linguagem formal. Uma das principais características da formalização é a eliminação de ambiguidades e vaguidades que, eventualmente, estejam presentes em sentenças da linguagem natural. Outros aspectos são elencados por Kreisel (1967), para quem a formalização desempenha papel destacado em processos cognitivos, filosóficos e relativos à prática matemática - isto porque a formalização de um conceito: (i) é parte importante do processo psicológico de entendimento; (ii) é componente essencial de programas fundacionais, tais como o de Hilbert e (iii) desempenha papel de relevo no trabalho diário do lógico-matemático. Exemplos dessa última afirmação são a análise da estrutura cumulativa de conjuntos desenvolvida por Zermelo, a explicação do porquê um problema matemático é um problema aberto com relação a uma teoria e o entendimento do processo de memorização de provas matemáticas. Estas são algumas das razões pelas quais a legitimidade e correção da formalização da linguagem natural não é posta sob suspeita neste trabalho.

Um primeiro passo em direção à análise rigorosa dos princípios diretivos é a formalização dos mesmos. E, ainda que a formalização seja considerada apenas da perspectiva da eliminação de ambiguidades e vaguidades, isso não acarretaria que os princípios diretivos devam ser considerados vagos ou ambíguos. Assumimos, como Freire (2018), que os princípios diretivos da aritmética governam a prática aritmética - não

ISSN: 2317-9570

há vaguidade ou imprecisão nesta afirmação. A formalização não tem por foco a eliminação de ambiguidades; é um primeiro movimento em direção à avaliação rigorosa do papel dos princípios na fixação do modelo padrão da aritmética.

Uma objeção que pode ser levantada quanto à formalização dos princípios diz respeito a uma pretensa circularidade: a aritmética é governada por princípios diretivos que são formalizados em sistemas semânticos forjados em metateorias impregnadas de noções aritméticas, e tais sistemas são empregados na análise dos princípios diretivos. Mas não há circularidade alguma. Isto porque os princípios diretivos da aritmética precedem conceitualmente os sistemas formais empregados na análise. Esta observação continua válida mesmo quando os sistemas nos quais os princípios diretivos são analisados internalizam completamente a aritmética ou quando a metateoria desses sistemas é a própria aritmética.

Na formalização dos princípios será empregada a assinatura $\Sigma = \{+, \cdot, s, 0\}$, em que 0 é símbolo de constante, s é símbolo de função unária e tanto $+$ quanto $\cdot$ são símbolos de operações binárias. A interpretação pretendida dos símbolos é a usual: 0 é o numeral que denota o número zero, s é o símbolo de função tal que, para cada numeral t, temos que st é o numeral que denota o número que é o sucessor do número denotado por t. Por sua vez $+$ é o símbolo de operação binária tal que, dados os numerais t e v, temos que $t + v$ é o numeral que denota a soma dos números denotados por t e v. Analogamente, $\cdot$ é o símbolo de operação binária tal que, dados os numerais t e v, temos que $t \cdot v$ é o numeral que denota o produto dos números denotados por t e v. A não estipulação de uma interpretação pretendida aos símbolos da assinatura inviabiliza a formalização dos princípios. Isso porque a atribuição de símbolos da linguagem formal a termos, relações e expressões da linguagem natural é componente importante da formalização. Não é possível formalizar o discurso normativo aritmético, ou qualquer discurso, se não é atribuída uma interpretação privilegiada aos membros da assinatura da linguagem empregada na formalização da linguagem natural.

A concatenação de sequências de símbolos será representada por $\frown$ e, com isso, a concatenação das sequências u e v será denotada $u^\frown v$. Os numerais são os termos $s^m(t)$, definidos por recursão a partir dos Σ-termos t pelas seguintes cláusulas:

$$s^0(t) = t \quad \text{e} \quad s^{m+1}(t) = s^\frown s^m(t).$$

Uma vez que o único termo da assinatura Σ em consideração é o símbolo de constante 0, adotamos a definição recursiva dos numerais em que $s^0(0) = 0$ e $s^m(t)$, para $m > 0$, é a expressão formada pela repetição de m símbolos s à esquerda do símbolo 0. Como usual, os numerais 0, $s0$, $ss0$, etc, serão representados por $\mathbf{0}$, $\mathbf{1}$, $\mathbf{2}$, etc.

O principal objetivo dessa seção é analisar em qual medida os princípios diretivos da aritmética contribuem para a fixação do modelo padrão da aritmética. Visando este fim, diremos que uma coleção de sentenças Γ com assinatura Σ em um sistema semântico $\mathbb{L}$ *fixa a estrutura* $\mathfrak{A}$ quando $\mathfrak{A}$ é modelo de Γ e qualquer outro modelo de Γ é isomorfo a $\mathfrak{A}$. Uma vez que, dada qualquer estrutura $\mathfrak{A}$ e qualquer conjunto X equipotente ao domínio de $\mathfrak{A}$, há um modo canônico de obter uma estrutura $\mathfrak{B}$, isomorfa a $\mathfrak{A}$ e com domínio X, não é possível fixar uma estrutura a menos de sua classe de isomorfismos. Posto de outro modo, uma coleção de sentenças fixa uma estrutura quando a coleção de sentenças é, no sentido da teoria de modelos, categórica.

A contribuição dos princípios diretivos para a fixação da estrutura padrão será avaliada do seguinte modo. Primeiro, escolhemos um sistema semântico $\mathbb{L}$ cuja linguagem seja suficientemente expressiva para formalizar, ao menos parcialmente, os princípios diretivos a partir da assinatura Σ. Uma Σ-estrutura $\mathfrak{A}$ é *correta segundo os princípios diretivos da aritmética* quando $\mathfrak{A}$ é um modelo do conjunto de sentenças que corresponde a formalização dos princípios diretivos. Neste sentido, os princípios constituem um critério de seleção, dentre todas as Σ-estruturas, daquelas que são corretas segundo os princípios diretivos. Se quaisquer dois membros na classe das estruturas corretas segundo os princípios diretivos são isomorfos, então os princípios diretivos da aritmética fixam as estruturas aritméticas corretas. Se os princípios diretivos da aritmética fixam as estruturas aritméticas corretas e $\langle \omega, +, \times, suc, 0\rangle$ é uma estrutura correta, então os princípios diretivos da aritmética *fixam o modelo padrão da aritmética*.

Fixado o modelo pretendido, é natural que se proceda a uma meta-análise dos princípios diretivos, isto é, que os sistemas semânticos que analisam os princípios diretivos da aritmética sejam submetidos, eles próprios, à análise. Tais sistemas serão avaliados com base em aspectos dedutivos, semânticos e matemáticos. Com relação ao primeiro desses aspectos, uma lógica é avaliada enquanto ciência da dedução

ISSN: 2317-9570

e o caráter dedutivo encontra-se intrinsecamente associado ao aparato formal. Desta perspectiva, o critério de avaliação dos sistemas semânticos é a completude. Sistemas semânticos nos quais as sentenças válidas sejam demonstráveis no sistema formal correlato são mais bem avaliados do que os sistemas semânticos que não apresentam esta característica.

O caráter semântico está relacionado ao papel de instrumento de identificação de estruturas matemáticas e noções correlatas que os sistemas semânticos podem assumir. Esse aspecto será avaliado, prioritariamente, segundo as propriedades semânticas de Löwenheim-Skolem e compacidade. Secundariamente, serão consideradas a propriedade de interpolação de Craig e definibilidade de Beth. Uma característica de sistemas semânticos munidos da propriedade de Löwenheim-Skolem é a impossibilidade de fixar estruturas que possuam domínio infinito. Já os sistemas munidos da propriedade de compacidade não distinguem estruturas com domínio finito, mas arbitrariamente grande, de estruturas com domínio infinito. Sistemas semânticos munidos dessas propriedades serão mais bem avaliados do que aqueles que não as possuem.[1]

Sistemas semânticos são impregnados de noções matemáticas e o ambiente natural de avaliação dessas noções é a teoria de conjuntos ZFC. Uma forma usual de avaliar quão impregnado de noções matemáticas encontra-se um determinado conceito, expressável em ZFC, é pela aferição da classe de modelos de ZFC para a qual o conceito sob avaliação é absoluto - quanto maior a classe de modelos, menos impregnada de noções matemáticas encontra-se o conceito em análise. Nos casos considerados neste trabalho, a coleção de sentenças é definida por recursão a partir de uma relação bem fundada e absoluta para modelos transitivos de ZFC. Nestes casos, a coleção de sentenças é absoluta para modelos transitivos de ZFC, como assevera o teorema IV.5.6 em Kunen (1992). Resta, portanto, avaliar para quais classes de modelos a relação de satisfatibilidade é absoluta.

Para comodidade do leitor, uniformidade de notação e sistematização dos resultados, os princípios diretivos para a aritmética presentes no trabalho de Ramos e Freire nesta mesma edição da Re-

[1] A razão dessa postura é o endosso ao ponto de vista que distinções entre cardinalidades finitas e infinitas, bem como distinções entre duas cardinalidades infinitas exigem que sejam adotados compromissos ontológicos fortes, o que afasta os sistemas capazes de tais distinções do escopo da lógica em sentido estrito.

vista de Filosofia Moderna e Contemporânea são listados a seguir.

1º Princípio Diretivo (PD_1): *cada número é denotado por um único numeral, que é um objeto sintático obtido pela repetição, possivelmente vazia, de um símbolo primitivo.*

2º Princípio Diretivo (PD_2): *cada numeral denota um único número.*

3º Princípio Diretivo (PD_3): *dados dois numerais s e t, a soma dos números denotados por s e t é denotada pelo numeral obtido pela repetição do símbolo primitivo determinado por s sobre t.*

4º Princípio Diretivo (PD_4): *dados dois numerais s e t, o produto dos números denotados por s e t é denotado pelo numeral obtido pela repetição da repetição s determinada por t.*

2.1 - Análise dos princípios diretivos em primeira ordem

Linguagens de primeira ordem são bastante expressivas, o que é atestado pelo fato de parte significativa da matemática contemporânea ser descrita em uma dessas linguagens - a linguagem da teoria de conjuntos ZFC. Mas não é possível formalizar, em primeira ordem e com uma única sentença, a expressão "a todo número corresponde um único numeral".[2] Uma vez que o princípio diretivo PD_1, apresentado através de uma expressão em linguagem natural, não pode ser formalizado por uma única sentença em linguagem de primeira ordem, dizemos que este princípio não admite uma *formalização plena* em primeira ordem. Entretanto, pode-se apelar ao esquema de indução presente nas axiomatizações em primeira ordem da aritmética de Peano e estabelecer uma correlação entre o princípio diretivo PD_1 e as infinitas instâncias do esquema de indução. Essa correlação pode ser entendida como uma *formalização parcial* do princípio diretivo PD_1.

Uma maneira de enunciar o princípio da indução é: considere uma propriedade que se aplica aos números; se o número zero goza

[2]Uma dificuldade em formalizar essa expressão deve-se ao fato que, na abordagem sob análise, números são vistos como os elementos do domínio de estruturas corretas segundo os princípios diretivos. Uma formalização da expressão em discussão deve condensar a informação de que todos os elementos no domínio de uma estrutura correta segundo os princípios desempenham papel de número. Com isso, dada uma estrutura correta $\mathfrak{A}$, para todo x, se x é elemento do domínio de $\mathfrak{A}$, então $x = 0^{\mathfrak{A}}$ ou $x = (s0)^{\mathfrak{A}}$ ou $x = (ss0)^{\mathfrak{A}}$ ou $\cdots$ Uma tal disjunção infinta não pode ser formalizada em primeira ordem, embora seja facilmente formalizada na linguagem do sistema semântico infinitário que consideramos no corpo deste trabalho. Ademais, se houvesse uma sentença de primeira ordem φ que formalizasse a expressão em análise, então uma estrutura $\mathfrak{A}$ seria modelo de φ se, e somente se, o domínio de $\mathfrak{A}$ fosse infinito e enumerável, o que claramente não pode ser o caso.

ISSN: 2317-9570

desta propriedade e, se sempre que um número possui tal propriedade seu sucessor também a possui, então todos os números naturais possuem a propriedade em consideração.

Por um lado, o princípio diretivo PD_1 pode ser empregado para justificar o princípio da indução. Suponha PD_1 e que há uma propriedade acerca dos números naturais tal que (i) o número zero goza dessa propriedade, (ii) se um número possui tal propriedade, então seu sucessor também a possui e (iii) há um número que não possui a propriedade. Pela condição (i) o número denotado por **0** goza dessa propriedade e, pela condição (ii), se o número denotado por **k** possui a propriedade, então o número denotado por s**k** também possui tal propriedade. Decorre da definição recursiva que os numerais são esgotados pelas condições (i) e (ii) e, consequentemente, o número em (iii) que não goza da propriedade não pode ser denotado por um numeral, o que contraria PD_1.

Por outro lado, o princípio da indução pode ser usado para explicar o princípio diretivo. Admita o princípio da indução e que a propriedade considerada é “ser denotado por um numeral". A propriedade é desfrutada pelo número zero, pois este pode ser denotado pelo numeral **0**. O passo indutivo também é o caso, pois se um número é denotado pelo numeral **k**, seu sucessor pode ser denotado por s**k**. Logo, é o caso que todo número é denotado por um numeral.

A partir da interpretação fornecida aos termos primitivos da linguagem aritmética os demais princípios diretivos podem ser formalizados, plena e facilmente, em primeira ordem. O princípio diretivo PD_2 estabelece que numerais distintos denotam números distintos. Os números são elementos do domínio das estruturas corretas segundo os princípios e os numerais são termos gerados recursivamente. Disto, uma formalização do segundo princípio diretivo é a expressão

$$\forall m \forall n \left(m \neq n \rightarrow s^m(0) \neq s^n(0)\right).$$

O princípio PD_3 estabelece que a soma de dois números é denotada pela soma dos numerais que denotam cada um dos números. Os numerais são definidos pela concatenação do símbolo s e, desse modo, é natural estabelecer que a soma de numerais é dada pela justaposição, à esquerda, da sequência de símbolos s ao outro numeral. Por exemplo, a soma do numeral $s0$ com o numeral $ss0$ é a concatenação $s^\frown ss(0)$, isto é, $sss0$. A partir dessas considerações é razoável formalizar PD_3 por

$$\forall m \forall n \left(s^{m+n}(0) = s^m(0) + s^n(0)\right).$$

A formalização de PD_4 é análoga ao do PD_3 e escrita como

$$\forall m \forall n \left(s^{m \cdot n}(0) = s^m(0) \cdot s^n(0)\right).$$

Admitindo-se que os princípios diretivos regem a prática aritmética, a formalização aponta para uma justificativa plausível e fundamentada dos axiomas de Peano. De fato, se a formalização dos princípios diretivos da aritmética em primeira ordem é uma axiomatização da arimética de Peano, os axiomas são asserções aritméticas verdadeiras não porque eles são verdades evidentes acerca dos números; afinal, não parece não problemático afirmar que todas as infinitas instâncias do princípio da indução são verdades evidentes. Tampouco parece razoável que os axiomas sejam vistos como verdadeiros pelo fato de constituírem uma adaptação, no que tange a expressividade, da veracidade dos axiomas da aritmética estipulados por Peano na linguagem de segunda ordem. Uma vez que se aceite a tese de que há princípios que regulamentam a aritmética e estes são explicitados por PD_1-PD_4, os axiomas da aritmética de Peano em primeira ordem não devem ser postos sob suspeição porque são formalizações, não plenas, dos princípios diretivos que regem a aritmética. E uma vez que seja aceita a proposta de que os princípios diretivos regem a prática aritmética, conclui-se que a adequação dos axiomas de Peano também é fundada na prática.

Conversamente, se os axiomas da aritmética de Peano são considerados não problemáticos, então não se pode advogar contra a adequação dos princípios PD_1-PD_4 pois estes são traduções, para a linguagem natural, daqueles.

Fixada uma linguagem e formalizados os princípios diretivos, procede-se a investigação do papel destes com relação a fixação da estrutura padrão. A classe de estruturas corretas segundo os princípios diretivos pode, certamente, ser considerada uma noção matemática. E o ambiente para a análise de noções matemáticas é, por excelência, a lógica de primeira ordem. As razões para esse reconhecimento são consonantes com as métricas que propomos para a avaliação dos sistemas semânticos. Do ponto de vista dedutivo, sistemas semânticos de primeira ordem são completos; quanto ao critério matemático, assim como o conjuntos das sentenças, a relação de satisfatibilidade é absoluta para todos os modelos transitivos de ZFC, resultado folclórico presente, por exemplo, na seção 5 do capítulo 3 de Drake (1974). Do ponto de vista semântico, valem as propriedades de Löwenheim-Skolem, compacidade, definibilidade de Beth e interpolação de

Craig.

Quando as teses que os princípios diretivos regem a prática aritmética e que a prática deve fundamentar a fixação do modelo pretendido são endossadas, o fato da análise dos princípios em sistemas semânticos de primeira ordem não fixar o modelo pretendido pode ser creditado à limitação expressiva da linguagem de primeira ordem - e, se este for o caso, linguagens mais expressivas devem contribuir para a fixação do modelo pretendido. Mas, embora a formalização dos princípios diretivos em primeira ordem não fixe o modelo pretendido da aritmética, a abordagem normativa motiva uma diferenciação importante na classe de modelos dos sistemas formais aritméticos. Essa diferenciação se faz necessária pois nenhum sistema formal, em primeira ordem, é capaz de diferenciar o modelo pretendido dos demais modelos do sistema formal. Naturalmente, o elemento fundante da identificação do modelo padrão de um sistema formal não podem ser as pretensões de quem formata o sistema. Uma saída é oferecida pelos princípios diretivos, segundo o entendimento de que um modelo não padrão é um modelo do sistema formal que não se conforma a algum dos princípios diretivos.

Por exemplo, consideremos a teoria cuja linguagem é a linguagem da aritmética de Peano acrescida de um símbolo de constante c e cujos axiomas são os da aritmética de Peano, estendidos com uma coleção recursiva de axiomas da forma $c \neq \mathbf{0}$, $c \neq \mathbf{1}$, $\mathbf{c} \neq \mathbf{2}, \ldots$ Esta teoria é consistente e tem modelo $\mathfrak{C}$, cujo reduto apropriado é modelo da aritmética de Peano. Mas a estrutura $\mathfrak{C}$ não é um modelo pretendido da aritmética, pois não é correta com relação ao princípio diretivo PD_1.

Sabidamente, elemento importante para a não fixação da estrutura padrão pela formalização dos princípios diretivos, em primeira ordem, é o teorema ascendente de Löwenheim-Skolem. Na sequência, os princípios diretivos serão formalizados em linguagens mais expressivas do que a de primeira ordem e o resultado será avaliado em sistemas semânticos nos quais não vale a versão ascendente de Löwenheim-Skolem.

2.2 Análise dos princípios diretivos em segunda ordem.

Posto que a linguagem de primeira ordem é um fragmento da linguagem de segunda ordem, toda sentença da primeira é uma sentença da segunda e, consequentemente, os princípios diretivos $PD_2 - PD_4$ podem ser formalizados pelas mesmas expressões empregadas nas considerações acima, relativas à primeira ordem. Quanto ao princípio diretivo PD_1, em segunda ordem é possível a formalização plena, isto é, expressar PD_1 em uma única sentença da linguagem formal. Uma apresentação dos princípios diretivos são as sentenças a seguir, em que P é uma variável de predicado que, possivelmente, emprega parâmetros.

$$\text{Princípio PD}_1:\ \forall P\left[P(0)\wedge\forall x\left(P(x)\rightarrow P(sx)\right)\rightarrow\forall x\, P(x)\right].$$

$$\text{Princípio PD}_2:\ \forall m\forall n\left(m\neq n\right)\rightarrow s^m(0)\neq s^n(0)\Big).$$

$$\text{Princípio PD}_3:\ \forall m\forall n\left(s^{m+n}(0)=s^m(0)+s^n(0)\right).$$

$$\text{Princípio PD}_4:\ \forall m\forall n\left(s^{m\cdot n}(0)=s^m(0)\cdot s^n(0)\right).$$

Claramente, a formalização dos princípios diretivos corresponde aos axiomas da aritmética de Peano em segunda ordem. Portanto, de modo análogo ao discutido anteriormente, a adequação dos axiomas da aritmética de Peano em segunda ordem pode ser asseverada pois corresponde à formalização dos princípios diretivos. Segundo essa perspectiva, não são os axiomas de Peano que motivam ou estabelecem os princípios diretivos da aritmética; os princípios legislam sobre a aritmética e devem ser vistos como anteriores ao sistema formal. Disto, o fato dos axiomas corresponderem à formalização dos princípios deve ser considerada como uma evidência favorável à correção dos axiomas, não o inverso.

Uma resposta afirmativa para a questão da fixação da estrutura padrão da aritmética a partir dos princípios diretivos é fornecida pelo teorema de categoricidade de Dedekind: toda Σ-estrutura que é modelo da formalização dos princípios diretivos da aritmética em segunda ordem é isomorfa à estrutura padrão $\langle\omega,+,\times,suc,0\rangle$. A demonstração desse resultado pode ser vista em Almeida (2017; teorema 1.2.1).

Uma vez que a linguagem de segunda ordem é bastante corriqueira na prática matemática e a formalização dos princípios nessa linguagem fixa a estrutura padrão da aritmética, poder-se-ia argumentar que a apresentação de bases para a fixação do modelo canônico da aritmética está concluída. Mas há uma crítica forte e imediata a este encaminhamento: linguagens de segunda ordem estão sujeitas a diversas críticas.

A situação é particularmente problemática pois há duas interpretações para o escopo das variáveis de segunda ordem, a interpretação ampla e a interpreta-

ISSN: 2317-9570

ção de Henkin - conforme Shapiro (1991) - das quais redundam sistemas semânticos de segunda ordem distintos. A possibilidade de mais uma semântica para a linguagem formal não é estritamente problemática. De fato, há duas interpretações possíveis para as variáveis livres de uma fórmula da linguagem de primeira ordem, a interpretação condicional e a interpretação de generalidade, que podem ser vistas, por exemplo, em Freire (2009) e Kleene (1967). Mas, em primeira ordem, a escolha da semântica não é particularmente problemática, pois ambas são inter-interpretáveis. O mesmo não ocorre com as interpretações de segunda ordem ampla e de Henkin.

Se a semântica de Henkin é adotada, a interpretação dos quantificadores é sujeita a críticas quanto à arbitrariedade da interpretação, isto porque o escopo dos quantificadores sofrem restrições arbitrárias. Por exemplo, se $\Sigma = \{P\}$, em que P é uma variável de predicado unário e $\mathfrak{A}$ é uma Σ-estrutura de domínio A, então a relação

$$\mathfrak{A} \models \forall P \ \varphi(P)$$

não deve ser entendida como "para todo subconjunto $C \in \mathcal{P}(A)$, é o caso que $\mathfrak{A} \models \varphi(P)[C]$"; deve, sim, ser entendida como "para todo subconjunto $C \in d \subseteq \mathcal{P}(A)$, é o caso que $\mathfrak{A} \models \varphi(P)[C]$". Isto constituí uma restrição completamente arbitrária no escopo do quantificador, pois a restrição d na coleção de subconjuntos de A é arbitrária. Além deste, há outro problema com a semântica de Henkin: o sistema semântico resultante pode ser interpretado no sistema de primeira ordem, algo que influencia e contamina a análise em segunda ordem.

A semântica ampla suporta a categoricidade da aritmética de Peano e possui uma alegada naturalidade semântica, pois não restringe de modo arbitrário o escopo de quantificação. Entretanto, essa interpretação não é livre de críticas. Uma delas é apontada por Putnam (1980) no contexto de estabelecer os referentes dos elementos do vocabulário da linguagem de segunda ordem:

> [A] interpretação 'pretendida' do formalismo de segunda ordem não é fixada pelo uso do formalismo (o próprio formalismo admite os chamados modelos de Henkin [$\cdots$]), e [assim] torna-se necessário atribuir a mente poderes especiais para 'agar-

[3]Tradução nossa do original: *($\cdots$) the 'intended' interpretation of the second-order formalism is not fixed by the use*

> rar' noções de segunda ordem.[3]

Uma vez que a lógica de segunda ordem é passível de discussão quanto à interpretação dos quantificadores e, como aponta Putnam, devem ser feitos apelos a propriedades subjetivas para a adoção da semântica e fixação do referente, é razoável considerar que os sistemas semânticos de segunda ordem não são a ferramenta ideal para a análise dos princípios diretivos da aritmética.

Relacionada e adicionalmente a essas críticas, os sistemas semânticos de segunda ordem são muito mal avaliados pelas métricas fixadas. Do ponto de vista dedutivo, não há aparato sintático que seja completo para a semântica ampla. Da perspectiva semântica, o sistema não é dotado das propriedades de Löwenheim-Skolem, compacidade, definibilidade de Beth ou interpolação de Craig.

Por fim, uma fonte recorrente de críticas às linguagens de segunda ordem é a expressividade dessas linguagens. Afinal, segundo Tharp (1975), o elevado poder expressivo permite que questões matemáticas sofisticadas sejam reduzidas à pura semântica.

> O poder expressivo desta lógica [de segunda ordem], que é demasiado grande para admitir um procedimento de prova, é adequado para expressar declarações conjuntistas. Questões abertas típicas, como a hipótese do contínuo ou a existência de grandes cardinais, são facilmente declaradas como questões da validade de fórmulas de segunda ordem. Assim, os princípios dessa lógica fazem parte de uma área ativa e um pouco esotérica da matemática. Parece haver um sentimento justificável de que esta teoria deve ser considerada matemática, e que lógica deve ser mais auto-evidente e menos aberta.[4]

Se os princípios diretivos fixam o modelo padrão e se tais princípios são considerados anteriores a algumas noções matemá-

of the formalism (the formalism itself admits the so-called 'Henkin models' $(\cdots)$, and it becomes necessary to attribute to the mind special powers of 'grasping second-order notions'. (1980; p. 481).

[4]Tradução nossa do original: *"The expressive power of this logic [second-order], which is too great to admit a proof procedure, is adequate to express set-theoretical statements. Typical open questions, such as the continuum hypothesis or the existence of big cardinals, are easily stated as questions of the validity of second order formulas. Thus the principles of this logic are part of an active and somewhat esoteric area of mathematics. There seems to be a justifiable feeling that this theory should be considered mathematics, and that logic - one's theory of inference - is supposed to be more self-evident and less open."* (1975; p. 38).

ISSN: 2317-9570

ticas e aos sistemas formais, então é certamente desejável que a ferramenta de análise dos princípios faça o mínimo de referência a tais noções. E, embora seja reconhecidamente problemático caracterizar precisamente a extensão de “mínimo de referência a noções matemáticas", a citação acima revela a percepção que, quando se trata da lógica de segunda ordem, o mínimo é extrapolado. Essa crítica à permeabilidade do sistema semântico por noções matemáticas é consonante com a análise da relação de satisfatibilidade, pois esta é absoluta apenas para modelos supertransitivos de ZFC.[5]

Quando formalizados em segunda ordem, os princípios diretivos fixam a estrutura padrão da aritmética. Mas a análise do sistema semântico que viabiliza a fixação da estrutura padrão fomenta objeções sérias e que podem obscurecer a real contribuição dos princípios diretivos. O emprego de uma linguagem menos expressiva certamente é valoroso.

2.3 - Análise dos princípios diretivos em linguagem infinitária

Na linguagem formal ora em consideração as variáveis são, exclusivamente, variáveis individuais. E são admitidas fórmulas de comprimento infinito enumerável, mas nas quais o número de ocorrências de variáveis ligadas é finito. Tal linguagem, diferentemente da linguagem de segunda ordem, não está muito presente no discurso matemático, embora ocorram exceções marcantes como, por exemplo, apresentações canônicas dos grupos de torção. A baixa expressividade, quando comparada a segunda ordem, associada a propriedades metateóricas que a aproximam bastante da primeira ordem, podem ser consideradas explicações da pouca aplicação. Uma apresentação dessa linguagem e do correspondente sistema semântico $L_{\omega_1\omega}$ pode ser vista em Keisler (1971).

Assim como na linguagem de segunda ordem, a formalização dos princípios diretivos na linguagem infinitária é plena. O princípio PD_1 pode ser formalizado pelo emprego de um quantificador universal e uma disjunção enumerável. E, por se tratar de uma extensão da linguagem de primeira ordem, os demais princípios podem ser formalizados pelas mesmas expressões formais empregadas em primeira ordem. Entretanto, será apresentada outra formalização, com o emprego de sentenças com comprimento enume-

[5]Consequência da operação “tomar as partes” na semântica de segunda ordem e de II.4.24 em Kunen (2011).

rável e sem apelo ao emprego de quantificadores.

Princípio PD_1 :
$$\forall x \bigvee_{m \in \omega} \big(x = s^m(0)\big).$$

Princípio PD_2 :
$$\bigwedge_{m \in \omega} \bigwedge_{n \in \omega \setminus \{m\}} \big(s^m(0) \neq s^n(0)\big).$$

Princípio PD_3 :
$$\bigwedge_{m \in \omega} \bigwedge_{n \in \omega} \big(s^m(0) + s^n(0) = s^{m+n}(0)\big).$$

Princípio PD_4 :
$$\bigwedge_{m \in \omega} \bigwedge_{n \in \omega} \big(s^m(0) \cdot s^n(0) = s^{m \cdot n}(0)\big).$$

A correção da formalização dos princípios nas expressões acima é imediata. O princípio PD_1, por exemplo, afirma que todo número é denotado por um numeral. Assim, um número qualquer é denotado por 0 ou $s0$ ou $ss0$ ou ..., e essa disjunção enumerável pode ser formalizada pela primeira expressão acima. A formalização dos demais princípios é também é imediata. O teorema de categoricidade para $L_{\omega_1 \omega}$, a seguir, responde afirmativamente ao problema da fixação da estrutura padrão com base nos princípios.

Teorema 1 *Toda Σ-estrutura que satisfaz a formalização dos princípios diretivos da aritmética em $\mathbb{L}_{\omega_1 \omega}$ é isomorfa à estrutura padrão.*

Prova. Sejam $\langle \omega, +, \times, suc, 0 \rangle$ o modelo padrão e $\langle A, \oplus, \otimes, mu, c \rangle$ uma Σ-estrutura $\mathfrak{A}$. Definimos a função h de domínio ω e codomínio A pelas cláusulas recursivas (i) $h(0) = c$ e (ii) $h(suc(a)) = mu(h(a))$. Grosseiramente, o 'zero' da estrutura canônica é levado, por h, no zero da estrutura $\mathfrak{A}$ e h mapeia a sucessão de modo homogêneo. Por indução em ω mostramos facilmente que h é injetiva. A sobrejetividade de h é consequência direta de $\mathfrak{A}$ ser modelo da formalização de PD_1 - de fato, claro está que a interpretação do Σ-símbolo de função s é mu e, consequentemente, pela formalização de PD_1, se $a \in A$, então $a = c$ ou $a = mu(b)$, para algum $b \in A$. A verificação que h preserva as propriedades estruturais da soma e produto, isto é, que (i) $h(a + b) = h(a) \oplus h(b)$ e (ii) $h(a \times b) = h(a) \otimes h(b)$, é feita facilmente a partir das formalizações dos princípios diretivos PD_3 e PD_4. Consequentemente, h é um isomorfismo de $\langle \omega, +, \times, suc, 0 \rangle$ em $\langle A, \oplus, \otimes, mu, c \rangle$. ■

Formalizados na linguagem infinitária, os princípios diretivos fixam o modelo pretendido da aritmética. E o sistema semântico empregado é bem avaliado pelas métricas adotadas. De fato, pelo critério semântico a lógica infinitária $L_{\omega_1 \omega}$ não padece dos mesmos problemas da lógica de segunda ordem. Não há propostas consolidadas de interpretações concorrentes para os quantificadores. Além disso, esses sistemas são dotados da propriedade de Löwenheim-Skolem, bem como das propriedades de interpolação de Craig, definibilidade de Beth e um resultado que mimetiza, em muitas aplicações, a compacidade.[6] Se são admitidas apenas fórmulas de comprimento enumerável e deduções a partir de um conjunto enumerável de sentenças, há uma contraparte sintática para $\mathbb{L}_{\omega_1\omega}$ que é adequada do ponto de vista dedutivo, isto é, há um aparato de prova que é correto e completo com relação a semântica, conforme Keisler (1971; cap. 4). Do ponto de vista matemático, $\mathbb{L}_{\omega_1\omega}$ se revela muito menos impregnado de noções matemáticas do que os sistemas de segunda ordem. De fato, a relação de satisfatibilidade é absoluta para qualquer modelo transitivo de ZFC que contenha a coleção dos conjuntos hereditariamente contáveis.[7]

O emprego de linguagens infinitárias em análises da aritmética não está isento de críticas. Por exemplo Button e Walsh (2015) destacam que

> Nós podemos obter a categoricidade adicionando-se à aritmética de Peano em primeira ordem uma disjunção contável dizendo: "tudo é zero, ou o sucessor de zero, ou o sucessor do sucessor de zero, etc.". Mas para capturar essa proposta, nós precisamos capturar o 'etc'.; e isso parece exatamente o desafio original de capturar a (ou uma) sequência de números naturais.[8]

No entanto, essa crítica não se aplica a análise desenvolvida aqui, pois não se trata do problema da 'captura' da sequência numérica por um sistema formal, nem de extensões da aritmética de Peano com sentenças de comprimento enumerável. A posição aqui analisada é que a prática aritmética é regida por princípios que são fundantes da verdade aritmética. O papel desempenhado pelos sistema semântico $\mathbb{L}_{\omega_1\omega}$, no que diz respeito à fixação da estrutura padrão, limita-se à formalização e análise dos princípios di-

[6] Embora a compacidade não valha em $L_{\omega_1\omega}$, no capitulo 3 de Keisler (1971) há um resultado que desempenha papel análogo ao da compacidade quando se trata de estender resultados modelo-teóricos de $L_{\omega\omega}$ para $L_{\omega_1\omega}$.

[7] Ver Ebbinghaus (1986; 3.2.2). Mais precisamente, esse resultado diz respeito a satisfatibilidade com relação a conjuntos admissíveis - neste trabalho consideramos o caso particular dos hereditariamente contáveis.

[8] Nossa tradução do original: *"(We can) obtain categoricity by adding to first-order Peano arithmetic a countable disjunction saying: "everything is either zero, or the successor of zero, or the successor of that, etc.". But to grasp this proposal, we need to grasp that 'etc.'; and that looks exactly like the original challenge of grasping the (or a) natural number sequence."* (2015; p.12).

ISSN: 2317-9570

retivos; não é analisado um conjunto com infinitos números ou a sequência infinita dos números naturais. São, sim, analisados os princípios diretivos que, por hipótese, regem a prática aritmética. E a análise desses princípios *no* sistema semântico $\mathbb{L}_{\omega_1\omega}$ explicita em qual sentido eles fixam a estrutura canônica. A análise *do* sistema semântico $\mathbb{L}_{\omega_1\omega}$ nos mostra em qual sentido esse sistema é adequado para a análise dos princípios diretivos: este sistema fixa o modelo padrão e é bem comportado segundo métricas semânticas, dedutivas e matemáticas.

Isso nos permite afirmar que, sob essas condições, a estrutura canônica da aritmética é fixada pelos princípios diretivos. E, dentre todas as estruturas adequadas à linguagem da aritmética e que satisfazem os axiomas de um sistema formal aritmético, os princípios diretivos permitem selecionar quais são as estruturas corretas com relação à prática matemática.

Claro que se a semântica da lógica infinitária for formalizada na teoria de conjuntos em primeira ordem, o teorema de categoricidade deixa de valer, mas essa situação é incontornável por se tratar de uma característica do ambiente de análise da semântica infinitária e, não necessariamente, da própria semântica.

Um comentário em favor da adequação do sistema semântico $\mathbb{L}_{\omega_1\omega}$ como ambiente de análise de noções matemáticas pode ser tecido a partir das considerações de Bays (2001), relativos a primeira ordem:

> [As] noções de finitude e recursão são necessárias para descrever a teoria de modelos de primeira ordem, uma vez que as fórmulas da lógica de primeira ordem podem ser de comprimento finito e arbitrário, mas não infinito, e a relação de satisfação é definida recursivamente.[9]

Para uma teoria de modelo infinitária exige-se que fórmulas de comprimento infinito enumerável sejam aceitas como não problemáticas e que a recursão transfinita enumerável seja aceita como não problemática. Essas exigências são completamente razoáveis, uma vez que em primeira ordem são aceitas fórmulas com comprimento arbitrariamente grande (e vale a compacidade!) e o discurso

[9] Tradução nossa do original: *"[T]he notions of finitude and recursion are needed to describe first-order model theory, since first-order formulas can be of arbitrary finite length, but they cannot be infinite, and first-order satisfaction is defined recursively."* (2001; p. 345).

ISSN: 2317-9570

sobre coleções enumeravelmente infinitas permeia parte significativa da matemática e filosofia da matemática contemporânea.

As considerações tecidas nesta seção acerca do papel dos sistemas semânticos enquanto ferramentas de análise dos princípios diretivos oferecem uma perspectiva de investigação que pode ser de interesse para a filosofia da lógica. Isso porque, pelo teorema ascendente de Lowenheim-Skolem, teoremas de categoricidade para teorias munidas de modelos infinitos devem ser desenvolvidos em extensões da lógica de primeira ordem. Entretanto, o apelo a estas extensões é, em geral, visto como uma petição de princípio, pois estas extensões envolvem o apelo à noções matemáticas que o resultado de categoricidade procurou garantir em primeiro lugar.

Na análise conduzida neste trabalho não há petição de princípio, pois o que se investiga são as consequências do paradigma segundo o qual os princípios diretivos legislam sobre a aritmética. A adoção deste paradigma fomenta questões acerca de qual linguagem deve ser considerada adequada para a formalização e análise dos princípios. A resposta para essas questões indica que um ambiente bastante adequado é fornecido pelos sistemas semânticos $\mathbb{L}_{\omega_1\omega}$. Afinal, estes são os mais bem comportados segundo as métricas adotadas e tanto fixam o modelo pretendido da aritmética quanto permitem a formalização plena desses princípios.

Na sequência são apresentadas, e comparadas à abordagem via princípios diretivos, estratégias alternativas de fixação do modelo padrão da aritmética.

3 - Análise comparativa

Nesta seção são esboçadas três propostas de fixação do valor de verdade de sentenças aritméticas. A primeira, McGee (1997), aborda o processo de apreensão da linguagem matemática e consequente fixação do valor de verdade de sentenças. Já a abordagem de Gaifman (2003) fixa o modelo padrão a partir de um critério de minimalidade na classe de modelos do sistema formal. Por fim, o estruturalismo modal de Hellman (1993) fixa a estrutura padrão com apelo a uma linguagem modal bastante expressiva.

A apresentação dessas posições visa a dois objetivos: (i) identificar convergências e divergências entre algumas propostas presentes na literatura e a abordagem normativa ao mesmo problema e (ii) delinear algumas características da interpretação do autor deste trabalho com relação a proposta de Freire (2018) para a fixação da verdade aritmética.

3.1 - Vann McGee, em *How We Learn Mathematical Language*

McGee analisa as condições de verdade de sentenças matemáticas dentro do paradigma realista com o objetivo de elucidar algumas das dificuldades enfrentadas por essa tradição na determinação do valor de verdade das sentenças. Uma dessas dificuldades pode ser posta nos seguintes termos:

> A concepção realista supõe que o sentido dos termos matemáticos é fixado com suficiente precisão para garantir que uma sentença tenha valor de verdade determinado. Agora, seja qual for o significado de uma expressão linguística, ela o possui em virtude dos pensamentos e práticas dos seres humanos. Nem todo significado é dependente de pensamentos e práticas de seres humanos - o fato de que um céu vermelho pela manhã significa tempestade não é uma questão de convenção - mas o fato de que o numeral '7' se refere ao quarto número primo é uma questão de como escolhemos usar o símbolo. Então, deve haver algo que pensamos, fazemos ou dizemos que fixa o significado pretendido de termos matemáticos. Como somos capazes de fazer isso?[10]

Embora McGee se dedique às sentenças da aritmética e da teoria de conjuntos, serão aqui consideradas apenas as primeiras. Com isso, uma resposta ao problema acima estará bem encaminhada quando houver uma explicação de como nossos pensamentos e práticas fixam o significado dos termos aritméticos de modo preciso e quando for estabelecido para cada sentença aritmética seu valor de verdade.

O problema atacado por McGee não diz respeito à atribuição de referente aos termos aritméticos. Para ele, é ponto pacífico: nossas práticas e pensamentos ao empregar a linguagem aritmética são incapazes de nos permitir discernir entre cópias isomorfas de

[10]Tradução nossa do original: *"[T]he realistic conception supposes that the meaning of mathematical terms is fixed with sufficient precision to ensure that a sentence has a determinate truth value. Now whatever meaning a linguistic expression has it possesses in virtue of the thoughts and practices of human beings. Not all meaning is thus dependent on human thought and action - the fact that a red sky in the morning means stormy weather isn't a matter of convenction - but the fact that the numeral '7' refers to the fourth prime is a matter of how we have chosen to use a symbol. So there must be something we think, do, or say that fixes the intended meaning of mathematical terms. How are we able to do this?"* (1997; p. 35-36).

estruturas aritméticas, o que torna impossível fixar, a menos de isomorfismos, o referente dos termos da linguagem aritmética. Em particular, não há referentes privilegiados para numerais.

Entretanto, a impossibilidade de fixar *o* referente de um termo aritmético não acarreta a impossibilidade de fixar as condições de verdade de sentenças aritméticas. Mas, então, como sentenças da linguagem aritmética adquirem determinado valor de verdade? Para responder a isso, McGee observa que aprendemos a linguagem da aritmética quando aprendemos a contar objetos concretos tais quais, por exemplo, palitos de sorvete. Mas, uma vez que a prática de contar palitos permite a acomodação de modelos diferentes do pretendido, "algo mais" é necessário; este algo mais, dirá McGeee, é uma teoria matemática. Assim, aprender aritmética é aprender uma linguagem aritmética *e* aprender uma teoria matemática.

O papel e características dessa teoria matemática são claras:

> [O] que precisamos para explicar o fato de que as sentenças aritméticas têm um valor de verdade determinado é uma teoria com as seguintes características: (i) a teoria, juntamente com os fatos matemáticos, deve determinar um valor de verdade exclusivo para cada sentença aritmética, de tal maneira que os axiomas aritméticos aceitos sejam classificados como verdadeiros, e (ii) a teoria deve ser aprendida por seres humanos.[11]

Mas qual seria esta teoria? A resposta óbvia - a coleção de todas as sentenças aritméticas válidas na estrutura padrão da aritmética - é pouco promissora. Isto porque é difícil sustentar que seres humanos aprendam uma teoria que não seja, ao menos, recursiva. Por este motivo, não é razoável supor que a teoria aprendida quando se aprende aritmética é uma teoria de primeira ordem T que inclui as sentenças aritméticas verdadeiras e segundo a qual uma sentença aritmética φ é verdadeira se, e somente, φ é consequência lógica de T. Uma saída promissora, aponta McGee, poderia ser a lógica de segunda ordem

[11]Tradução nossa do original: *"[W]hat we require to account for the fact that purely arithmetical sentences have a determinate truth value is a theory with the following two characteristics: (i) the theory, together with the mathematical facts, must determine a unique truth value for each arithmetical sentence, in such a way that the accepted arithmetical axioms are classified as true, and (ii) the theory must be learnable by human beings."* (1997; p. 43).

e, se este fosse o caso, aprender aritmética seria aprender uma teoria: a aritmética de Peano em segunda ordem.

Admitida essa solução, a determinação do valor de verdade das sentenças aritméticas está esclarecido: uma sentença aritmética é verdadeira apenas no caso de ser consequência lógica da teoria e, falsa, apenas no caso de sua negação ser consequência lógica da teoria. Essa solução é aceitável? McGee dirá que, uma vez que a aritmética de Peano em segunda ordem é categórica e formada apenas por axiomas simples, acrescidos do axioma esquema de indução em segunda ordem, se não houver objeções à lógica de segunda ordem, a resposta é sim.

Mas, reconhece McGee, pairam muitas suspeitas sobre a lógica de segunda ordem, o que nos obriga a não nos afastarmos dos recursos da lógica de primeira ordem. Submetendo-se a tais restrições, McGee propõe que aprendemos aritmética através de esquemas e com o auxílio de regras de substituição.

> O que aprendemos quando aprendemos um vocabulário matemático não é um conjunto fixo de axiomas de primeira ordem, mas sim um conjunto fixo de axiomas de primeira ordem e *esquemas de axiomas*. A aritmética de Peano em primeira ordem, como geralmente é apresentada, consiste no *axioma esquema de indução*, que é a expressão aberta obtida do axioma de indução de segunda ordem, excluindo o quantificador universal inicial, juntamente com uma pequena lista de fatos aritméticos rudimentares. Qualquer expressão (fechada) que pode ser obtida a partir do axioma esquema de indução, substituindo uma expressão aberta pela variável de segunda ordem livre (que, nesse contexto, é chamada de *variável esquemática*), evitando colisões de variáveis ligadas e, então, prefixando quantificadores universais de primeira ordem, é um *axioma de in-*

[12]Tradução nossa do original: *"[W]hat we learn when we learn a mathematical vocabulary is not a fixed set of first-order axioms, but rather a fixed set of first-order axioms and* axiom schemata. *First-order Peano Arithmetic, as usually presented, consists of the* Induction Axiom Schema, *which is the open sentence gotten from the Second-order Induction Axiom by deleting the initial universal quantifier, together with a short list of rudimentary arithmetical facts. Any (closed) sentence that can be gotten from the Induction Axiom Schema by substituting an open sentence for the free second-order variable (which, in this context, is called a* schematic letter*), avoiding collisions of bound variables, then prefixing universal first-order quantifiers, is an* Induction Axiom." (1997; p. 57).

ISSN: 2317-9570

dução.[12]

Quanto ao aprendizado da aritmética,

> [O] que nos é ensinado quando aprendemos a linguagem da aritmética são os axiomas de Peano, entendidos de modo que podemos substituir no axioma esquema de indução qualquer sentença aberta que nos agrade. Podemos substituir qualquer expressão aberta do inglês. Além disso, se estendemos o inglês dos dias de hoje com vocabulário adicional, esperamos poder ser capazes de substituir em qualquer expressão aberta do idioma estendido [$\cdots$]. Nossa compreensão da linguagem da aritmética é tal que antecipamos que o axioma esquema de indução, como as leis da lógica, irão persistir em todas essas mudanças. Não há um único conjunto de axiomas de primeira ordem que expresse plenamente o que aprendemos sobre o significado da notação aritmética quando aprendemos o axioma esquema de indução, já que sempre somos capazes de gerar novos axiomas de indução, pela expansão do idioma.[13]

Com as passagens acima, parece claro que McGee advoga que ao aprendermos aritmética, estamos aprendendo axiomas de primeira ordem e uma "regra" de livre instanciação de um esquema de indução. Esta regra seria poderosa o suficiente para a fixação, a menos de isomorfismos, do referente dos termos aritméticos. Além disso, o aprendizado dessa regra não estaria sujeito às principais objeções da lógica de segunda ordem.

Concordamos integralmente com a posição de McGee acerca da impossibilidade de fixação, a menos de isomorfismo, do referente dos termos aritméticos. Adicionalmente, endossamos sua visão de que a impossibilidade de fixação do referente não impede que sentenças aritméticas tenham valor de verdade determinado e, principalmente, que o valor de verdade das sentenças aritméticas deve ser fixado pela prática,

[13]Tradução nossa do original: *"Rationally reconstructed, what we are taught when we learn the language of arithmetic is the Peano axioms, so understood that we can substitute into the Induction Axiom Schema any open sentence we like. We can substitute any open sentence of English. Moreover, if we extend present-day English by adjoining additional vocabulary, we expect to be able to substitute in any open sentence of the extended language (...)* (1997; p. 58).

em harmonia com o modo pelo qual aprendemos aritmética. Entretanto, vemos com ressalvas a solução por ele indicada e acreditamos que proposta de Freire (2018), fundada nos princípios diretivos, fornece uma resposta muito mais natural ao problema da fixação do valor de verdade.

De fato, não nos parece fundado na prática que, ao aprendermos aritmética, estamos aprendendo *os axiomas de Peano* e que os *entendemos* de tal maneira que um esquema de indução pode ser instanciado por qualquer sentença que nos agrade, quer seja do léxico atual, quer seja de qualquer extensão do léxico atual. O próprio McGee reconhece a falta de conexão entre sua proposta de aprendizado da linguagem aritmética e os processos educacionais segundo os quais se aprende, desde a tenra idade, aritmética.

> Enquanto psicologia do desenvolvimento, a história [que foi contada] tem pouca plausibilidade, uma vez que, de fato, ser treinado para afirmar e concordar com frases matemáticas complexas desempenha pouco ou nenhum papel na aquisição, por crianças, do vocabulário matemático. [O que] descrevemos [foi] um processo através do qual o vocabulário matemático é adquirido por uma idealização realista de um aprendiz da linguagem, alguém que executa o mesmo tipo de atividade mental que executamos, mas as realiza sem falhas e incansavelmente. Para que se obtenha sucesso na resolução de nosso problema filosófico, *não* é requerido que o processo que descrevemos espelhe o processo pelo qual uma criança real aprende o vocabulário; é apenas requerido que o processo não requeira habilidades mentais que sejam de tipos diferentes daquelas realmente observadas em salas de aula.[14]

Em sua defesa, McGee argumenta que suas preocupações não são de natureza pedagógica mas, sim, filosóficas; em especial, ele

[14]Tradução nossa do original: *"As developmental psychology, the story we are about to tell has little plausibility, since, as a matter of fact, being trained to assert and assent to complex mathematical sentences plays little if any role in children's acquisition of mathematical vocabulary. [...] We describe a process by which mathematical vocabulary is acquired by a realist's idealization of a language learner, one who carries out the same sorts of mental activities we carry out but performs them tirelessly and flawlessly. For the account to succeed in solving our philosophical problem, it is not required that the process we describe resemble the process by which real-life children learn the vocabulary; it is only required that the process not require mental abilities that are different in kind from those observed in actual classrooms."* (1997; p. 41-42).

tem interesse em entender como seres com nossas capacidades cognitivas são capazes de entrar em contato com o universo realista. Ademais, se sua proposta resolve um problema filosófico acerca da aquisição da linguagem, não há razões para supor que problemas pedagógicos, relativos ao modo segundo o qual crianças aprendem a linguagem aritmética, serão contemplados pela elucidação filosófica. Por fim, ele argumentará que, se há um descompasso entre a proposta e o processo de aprendizagem, este descompasso é esperado e ocorre sempre que se faz uma reconstrução racional de processos complexos.

Concordamos com McGee que soluções a problemas filosóficos não necessariamente apresentam consequências, ou podem ser estendidos, para a pedagogia ou outras áreas do conhecimento humano. Também endossamos a tese de que reconstruções racionais de processos complexos e dinâmicos apresentam distorções em relação a criação e desenvolvimento dos processos. Mas consideramos que sua hipótese do aprendizado da aritmética envolver o aprendizado de axiomas e um esquema que pode ser livremente instanciado, no léxico presente e futuro, é deveras artificial. Adicionalmente, acreditamos que a adoção de um quadro normativo para a aritmética, governado pelos princípios diretivos elencados, fornece uma solução mais crível para a reconstrução racional da aquisição da linguagem aritmética. Realmente, nos parece muito mais razoável sustentar que crianças, quando expostas a aritmética, entram em contato com um sistema notacional regido pelos princípios em vez do quadro proposto por McGee.

Isto porque parece-nos razoável considerar que, quando aprendermos aritmética, o que quer que aprendamos sobre os números é aprendido através de um sistema notacional. Não somos apresentados a números destituídos de notação; números são representados por numerais e, quanto à manipulação numérica, aprendemos a manipular números ao aprendermos a manipulação da linguagem notacional. Afinal, quando a soma de números é aprendida? Quando é aprendido o algoritmo de transformação de notações correspondente a soma. O mesmo ocorre para a multiplicação: aprender a multiplicar números é aprender o algoritmo da multiplicação dos numerais. Uma vez aprendido o sistema notacional, a comutatividade da soma e da multiplicação são consequências naturais dessas operações, não consequências de instanciações do esquema de indução em alguma extensão da linguagem aritmética.

Julgamos que, ao supormos que o aprendizado de um sistema notacional regido pelos princípios diretivos da aritmética é o vetor do aprendizado aritmético, nos aproximamos dos processos cognitivos e pedagógicos reais de aprendizado da aritmética. Além disso, pensamos que esta mesma suposição fornece uma solução ao problema de fixação do referente dos termos aritméticos de modo mais plausível que aquela oferecida por McGee. De fato, para fixação do valor de verdade das sentenças aritméticas McGee fixa o referente de termos aritméticos e, para realizar isto a menos de isomorfismos e sem recorrer a lógica de segunda ordem, considera um esquema segundo o qual livres instanciações em expansões ilimitadas da linguagem são permitidas.

A análise desenvolvida na seção anterior revela que a proposta aqui analisada demanda por muito menos recursos. Em vez de aprendermos axiomas e um esquema de indução que admite instanciações misteriosas, aprendemos um sistema notacional muito simples e lógica de primeira ordem, acrescida de disjunção enumerável - não parece crível que disjunções enumeráveis são mais problemáticas do que quaisquer instanciações que nos agrade em quaisquer extensões do léxico atual ou futuro.

3.2 - Haim Gaifman, em *Non-Standard Models in a Broader Perspective*

Gaifman defende que os modelos não padrão da aritmética e da teoria de conjuntos são fontes genuínas de interesse filosófico e matemático. Do ponto de vista matemático, sustenta - e concordamos - que os modelos não padrão são estruturas matemáticas interessantes em si mesmas e objetos legítimos de estudo, além de possuírem diversas aplicações. Estas aplicações ocorrem, por exemplo, quando tais modelos são empregados no estudo de sistemas dedutivos, quando motivam a análise não padrão de Robinson e quando atuam como guias heurísticos para investigações acerca do infinito.

Do prisma filosófico, Gaifman reconhece que Skolem - quem introduziu modelos não padrão nas disciplinas matemáticas - obteve diferentes conclusões conceituais a partir desses modelos: enquanto os modelos não padrão da aritmética são uma evidência da limitação dos sistemas formais de primeira ordem em capturar o modelo pretendido, os modelos não padrão da teoria de conjuntos fornecem suporte a uma posição cética acerca de conjuntos não enu-

ISSN: 2317-9570

meráveis. A despeito do reconhecimento da posição de Skolem, Gaifman defende que modelos não padrão não podem ser empregados para fomentar ceticismos quanto ao realismo. Esta postura é adotada em oposição ao argumento de Putnam (1980), segundo o qual o realismo é posto em dificuldades quando confrontado com modelos não padrão da teoria de conjuntos. O foco da presente análise não é a crítica de Gaifman ao Putnam mas, sim, algo que a precede: a caracterização de modelo padrão oferecida pelo primeiro.

Após uma breve digressão histórica, em que analisa desenvolvimentos em aritmética, no conceito de função e em geometria, Gaifman identifica três noções para os quais haveria uma interpretação pretendida: a sucessão numérica, a noção de boa ordenação e o conceito de conjunto construtível. Ele, então, procura responder à seguinte questão: *Qual, se algum dentre alguns modelos dados, é o modelo padrão?*[15]

O modelo pretendido da sucessão numérica é o modelo padrão da aritmética e Gaifman defende, com inspiração no trabalho de McGee acima discutido e numa condição de minimalidade, a seguinte caracterização:

[O modelo padrão] é o menor modelo, incluído como um segmento inicial em qualquer outro modelo. Se um determinado modelo é não padrão, isso será revelado por um segmento inicial apropriado que é fechado sob a função sucessor. Formalmente, a caracterização é expressa pelo esquema indutivo:

$$\text{(I)}\ \ P(0) \wedge \forall x \Big[N(x) \rightarrow \Big(P(x) \rightarrow P(x+1)\Big)\Big] \rightarrow \forall x\Big(N(x) \rightarrow P(x)\Big),$$

sendo que '$N(x)$' representa 'x é um número natural' e '$P(\cdot)$' representa qualquer predicado. Qualquer fórmula bem formada da linguagem em uso pode ser um substituto para '$P(\cdot)$'. O conceito de sequência de números naturais, contudo, *não* é dependente da linguagem. ... [para eliminar a dependência] o esquema indutivo deve ser entendido como o metacomprometimento: (II) Qualquer predicado não vago, em qualquer idioma, pode

[15] Tradução nossa do original: *"Which, if any, of some given models, is the standard one?"* (2003; p. 15).

ser substituído por ‘$P(\cdot)$’ em (I).[16]

Para Gaifman, o modo como este metacomprometimento deve ser assumido é sintetizada pelo *slogan: (Como Vann McGee diz,) se o próprio Deus criar um novo predicado, então esse predicado pode ser substituído por P.*”[17] Não endossamos a proposta de Gaifman e isso não se deve apenas ao seu endosso às ideias de McGee.

A proposta de adoção de um critério de minimalidade para a fixação do modelo padrão não nos parece razoável pois, como o próprio Gaifman aponta, a existência de modelos não padrão é um fenômeno recente na história da matemática - apresentados por Skolem em 1934 para a aritmética e em 1922 para a teoria de conjuntos. Concordamos com Skolem que modelos não padrão da aritmética evidenciam uma limitação dos sistemas formais de primeira ordem e, consequentemente, não nos parece crível sedimentar o modelo padrão enquanto *o menor modelo*. Isto porque, se os modelos não padrão são manifestações de sistemas formais de primeira ordem, nos parece que carece de sentido considerar todos os modelos não padrão desses sistemas formais para, então, fixar aquele que é segmento inicial de todos esses modelos e identificá-lo como a interpretação pretendida do discurso aritmético padrão praticado por Diofanto, Gauss, Fermat, etc.

Outra crítica à abordagem de Gaifman é que o tratamento por ele oferecido não é uniforme. Por exemplo, ao analisar o conceito de boa ordenação, a interpretação pretendida é dada pela classe de todos os conjuntos ‘verdadeiramente’ bem ordenados. Gaifman entende que boa ordenação é uma generalização natural do esquema (I) citado acima, com a condição do predicado $N(x)$ ser substituído por $\mathrm{Ord}(x)$, em que $\mathrm{Ord}(x)$ expressa que “x é ordinal”. Disto, e do fato que todo conjunto bem ordenado é isomorfo a um único segmento inicial da classe dos ordinais, ele propõe que essa classe seja a interpretação pretendida da

[16]Tradução nossa do original: *“[The stantard model] is the smallest model, included as an initial segment in any other model. If a given model is non-standard, then this will be revealed by a proper initial segment that is closed under the successor function. Formally, the characterization is expressed by the inductive scheme:* (I) $P(0) \wedge \forall x(N(x) \rightarrow (P(x) \rightarrow P(x+1))) \rightarrow \forall x\ (N(x) \rightarrow P(x))$ *where ‘$N(x)$’ stands for ‘x is a natural number’, and where ‘$P(\cdot)$’ stands for any predicate. Any wff of the language we are using can be substituted for $P(\cdot)$ The concept of the sequence of natural numbers is, however,* not *language dependent. The absoluteness of the concept can be secured, if we help ourselves to the full (standard) power set of some given infinite set; for then we can treat ‘P’ as a variable ranging over that power set. [...] The inductive scheme should be therefore interpreted as an open ended meta-commitment: (II) Any non-vague predicate, in whatever language, can be substituted for ‘$P(\cdot)$’ in (I).”* (2003; p. 15-16)

[17]Tradução nossa do original: *“As Vann McGee expresses it, if God himself creates a new predicate, then this predicate can be substituted for ‘P’.”* (2003; p. 15).

ISSN: 2317-9570

noção de boa ordem. Essa solução aponta para a adoção de um critério de maximalidade para a fixação do modelo padrão da noção de boa ordem.

Sem prejuízo a outras críticas que possam ser feitas, propor a fixação do modelo padrão de noções matemáticas a partir de critérios que oscilam entre maximalidade e minimalidade não parece uma boa prática, e pode ser confundido com a não adoção de critérios objetivos.

Quanto à teoria de conjuntos, Gaifman sustenta que se a interpretação padrão da boa ordem por ele formulada é aceita, então qualquer construção indutiva baseada na boa ordem padrão deve ser vista como bem delimitada e não ambígua. Desse modo a classe **L** dos construtíveis, que é obtida a partir dessas premissas, é a interpretação pretendida do conceito de conjunto construtível. Gaifman reconhece a debilidade dessa posição e, em sua defesa, apela aos números naturais:

> O argumento baseia-se apenas na plausibilidade de uma construção transfinita, percorrendo todos os ordinais, onde cada estágio não é problemático. No entanto, a sugestão é atraente e é reforçada pela observação de que não existem resultados de independência conhecidos para **L**, do tipo que proliferou na teoria dos conjuntos nos últimos quarenta anos. **L** parece a este respeito mais como os números naturais.[18]

Não concordamos com a premissa de Gaifman, pois julgamos que a interpretação padrão da noção de boa ordem é passível de críticas. Adicionalmente, ao comparar os conjuntos construtíveis com os números naturais, Gaifman volta a adotar um critério de minimalidade. Por fim, a suposição de um modelo pretendido para os construtíveis não parece encontrar respaldo na história da disciplina.

Se aceitamos que, além da aritmética, outras disciplinas matemáticas possuem uma interpretação privilegiada, então um método que seja aplicado uniformemente às diversas disciplinas matemáticas é preferível às estratégias que tenham escopo de aplicação particular. Se esta pos-

[18]Tradução nossa do original: *"The argument rests solely on the plausibility of a transfinite construction, running through all ordinals, where each stage is non-problematic. Yet the suggestion is appealing and it is reinforced by the observation that there are no known independence results for* **L**, *of the kind that have proliferated in set theory in the last forty years.* **L** *seems in this respect more like the natural numbers."* (2003; p. 17).

tura é endossada, então defendemos que, se a fundamentação normativa para a verdade adotada no caso aritmético for estendida para outras disciplinas matemáticas, então o método de fixação do modelo padrão empregado no caso aritmético pode ser estendido a outras disciplinas que tenham modelos pretendidos. Este, de fato, é o caso como a teoria de conjuntos, cujos resultados obtidos em Freire (2018) são harmônicos com a análise aqui desenvolvida para a aritmética.

3.3 - Geoffrey Hellman, em *Mathematics Without Numbers*

Hellman (1993) vê a matemática como *a livre exploração de possibilidades estruturais, perseguidas por meios dedutivos (mais ou menos) rigorosos*.[19] Essa posição é alinhada ao estruturalismo, escola em filosofia da matemática para a qual o elemento central das teorias matemáticas são as relações entre os objetos matemáticos, e não a natureza interna desses objetos. O endosso ao estruturalismo conduz ao entendimento de que, em aritmética, o que importa são as relações estruturais entre os itens de uma progressão arbitrária, e não a identidade individual.

Para Hellman, o estudo dos números naturais é um ponto de partida legítimo para análises estruturais. Uma razão para isso é o fato dos números estarem bastante enraigados na prática e pensamento matemáticos. Outra razão é o entendimento de que o estudo dos números é genuinamente estrutural, no sentido que estudar números é estudar as relações dos elementos presentes em ω-sequências arbitrárias.

Tradicionalmente, o ambiente natural de articulação das análises estruturais é a teoria de conjuntos. Entretanto, Hellman defende que este paradigma seja abandonado e uma nova direção seja seguida. Isto porque o paradigma tradicional torna a aritmética dependente da teoria de conjuntos em um nível que, do ponto de vista matemático, seria desejável evitar. Além disto, é salutar evitar que problemas presentes na teoria de conjuntos contaminem a aritmética.

Segundo o estruturalismo professado por Hellman, uma sentença aritmética verdadeira expressa não algo que é o caso em alguma estrutura infinita de um certo tipo mas, sim, o que seria o caso em qualquer estrutura de tipo apropriado que pudesse exis-

[19] Tradução nossa do original: *Mathematics is the free exploration of structural possibilities, pursued by (more or less) rigorous deductive means.* (1993; p. 6).

ISSN: 2317-9570

tir. Isto, contudo, sem que se comprometa com a existência de tal estrutura (ou classe estruturas). Esta postura exime Hellman de se comprometer com a existência de objetos abstratos. Em particular, não há o comprometimento com a existência de estruturas de um certo tipo, embora haja o compromisso com a possibilidade lógica de existência de estruturas de tipo apropriado. Uma vez que há apenas o comprometimento com a possibilidade de existência de estruturas, esta posição é conhecida *estruturalismo modal*.

Visando à articulação da aritmética dentro dessa visão estruturalista e original, Hellman inicialmente fornece um padrão de tradução (estrutural modal) para a aritmética com o intuito de eliminar toda referência literal aos números. Em seguida, organiza os pressupostos por trás do padrão de tradução para, então, defender a adequação de sua proposta com relação à prática matemática.[20] Por fim, seguindo o enredo estabelecido para a aritmética, Hellman estende sua visão estruturalista da matemática para a análise real e a teoria de conjuntos.

Concentraremos nossos comentários na primeira dessas etapas, o estabelecimento do padrão de tradução estrutural modal das sentenças aritméticas. A razão para isso é simples: julgamos que as exigências que devem ser cumpridas para o estabelecimento da interpretação almejada são um preço muito alto a ser pago pelo não comprometimento ontológico, e seguimos Hellman em (1993; §1, §2) e (2005; §1, §5) para explicar o porquê.

Quando φ é uma sentença aritmética, Hellman procura por um padrão de tradução que associe φ a um contrafactual da forma:

(A) Se X fosse qualquer ω-sequência, então φ seria o caso em X.

O padrão de tradução a ser estabelecido e que associa φ ao contrafactual (A) deve ser munido das seguintes características: (i) respeitar a objetividade da verdade das sentenças aritméticas e (ii) evitar quantificações sobre estruturas e mundos possíveis - e, consequentemente, evitar o comprometimento com a existência de objetos abstratos. A condição (i) é satisfeita pela adoção de restrições oriundas da matemática, que constituem o componente categó-

[20]A coordenação dos pressupostos subjacentes à interpretação abarca dois argumentos, um externo ao estruturalismo modal e outro, mais complexo, interno ao estruturalismo. A adequação da proposta para a prática é estabelecida pela descrição da forma como demonstrações aritméticas são recuperadas dentro do estruturalismo modal.

rico do padrão de tradução. Restrições na lógica subjacente fornecem as garantias exigidas por (ii) e constituem o componente hipotético do padrão de tradução.

A linguagem formal na qual o padrão de tradução é forjado emprega um operador primitivo de necessidade modal $\Box$. Esta linguagem não está comprometida com a semântica de mundos possíveis e nela a formalização do contrafactual (A) é dada por

$$\text{(A')} \quad \Box\forall X \Big(X \text{ é uma } \omega\text{-sequência} \rightarrow \varphi \text{ é o caso em } X \Big),$$

no qual a noção de possibilidade expressa por $\Box$ é codificada pelo sistema modal S_5.

Para que a proposta logre sucesso é preciso formalizar as noções de ω-sequência e satisfatibilidade na linguagem do padrão de tradução. O modo canônico de fazê-lo é via teoria de conjuntos, mas Hellman rejeita essa estratégia pois isso reduziria seu programa estruturalista a um fragmento das teorias de conjuntos modais. A alternativa por ele adotada é considerar que algumas noções de ordem superior serão necessárias, o que remete ao emprego de uma linguagem de segunda ordem.

De posse de uma lógica de segunda ordem modal, Hellman se opõe à formalização usual de "φ é verdadeira" como $\Box(\mathrm{T}_2 \rightarrow \varphi)$, que expressa "é necessário que φ seja consequência lógica de T_2", em que T_2 é a conjunção dos axiomas da aritmética de Peano em segunda ordem. Isto porque Hellman não endossa a forma como os símbolos não lógicos da linguagem de T_2 são interpretados na lógica modal. O caminho por ele seguido é via a quantificação sobre relações: se s é o único símbolo não lógico da linguagem de T_2, então "φ é verdadeira" deve ser formalizada por

$$\Box\forall f\Big(T_2 \rightarrow \varphi\Big)\,[s/f],$$

em que f é uma variável de relação binária e $[s/f]$ indica que f substitui toda ocorrência do símbolo não lógico s. A expressão acima é uma sentença que corresponde a asserção metamatemática de que φ é consequência lógica dos axiomas de T_2.

Hellman destaca que a sentença acima implica questionamentos sobre os domínios das estruturas envolvidas no estabelecimento da consequência lógica. Estes questionamentos são dirimidos pela quantificação sobre tais domínios, resultando em

$$\Box\forall X\forall f\Big(T_2 \rightarrow \varphi\Big)^X\;[s/f].$$

Se ψ é uma sentença da aritmética de primeira ordem, as operações de soma e multiplicação devem ser definidas de modo usual - por recursão a partir dos

ISSN: 2317-9570

axiomas de Peano em segunda ordem - e as sentenças que definem essas operações são acrescidas aos axiomas de T_2, resultando em $T_2^\star$. Com isso, o padrão de tradução (estrutural modal) proposto por Hellman para sentenças φ da aritmética de primeira ordem é

$$(B)\quad \Box\forall X\forall f\forall g\forall h\left(T_2^\star \to \varphi\right)^X,$$

em que g e h são variáveis de relação ternárias e todas as ocorrências de s, $+$ e $\times$ são substituídas por, respectivamente, f, g e h.

O esquema (B) acima corresponde ao componente hipotético da interpretação estrutural modal defendida por Hellman e expressa apenas o conteúdo estrutural modal das sentenças aritméticas. Mas, como reconhece Hellman (1993), *há muito mais na prática da matemática do que a mera declaração de sentenças. Existe a prática de provar teoremas*[21] e esse "muito mais" deve ser capturado pelo componente categórico.

Este último componente garante a possibilidade de existência de alguma ω-sequência através do axioma

$$\Diamond\exists X\exists f\left(T_2\right)^X\ [s/f].$$

Entretanto, enquanto Hellman se compromete com a possibilidade de existência de uma ω-sequência, ele prefere ser neutro quanto à existência de algo que possivelmente seja uma ω-sequência. Para que este seja o caso a fórmula de Barcan deve ser bloqueada. Isto porque, caso contrário, não há impedimentos à inferência de $\exists X\Diamond(X$ é uma ω-sequência) a partir de $\Diamond\exists X(X$ é uma ω-sequência).

Além da fórmula de Barcan, o componente categórico exige um cuidado adicional com o emprego da lógica de base: o axioma da compreensão da lógica de segunda ordem deve ser instanciado, unicamente, por fórmulas $\psi(x_1,...,x_n)$ em que não há ocorrência de operadores modais. O esquema de compreensão, em que R não ocorre em ψ, é formulado do seguinte modo:

$$\Diamond\exists R\forall x_1....\forall x_n\left(R(x_1,...,x_n)\leftrightarrow\psi\right).$$

Adicionalmente, são incorporados à lógica subjacente ao padrão de tradução alguns elementos mereológicos, que por sua vez conduzem a uma reformulação do esquema de compreensão restrito descrito acima. Mas não precisamos entrar nesses detalhes. O que apresentamos até o momento é suficiente para o ponto ao qual queremos chamar a atenção.

Consideramos extremamente feliz o movimento de Hellman em

[21] Tradução nossa do original: *"But there is much more to mathematical practice than the mere utterance of sentences. There is the deductive practice of theorem proving."* (1993; p. 10).

direção a posição filosófica que endossa a objetividade do valor de verdade das proposições aritméticas sem, em contrapartida, requerer a existência de objetos abstratos para fixar o valor de verdade das proposições. Entretanto, discordamos dele com relação ao *assunto* ao qual a aritmética corresponde. Enquanto Hellman vê a aritmética como estudo de relações que ocorrem no interior de possíveis estruturas arbitrárias, para as quais só é oferecida a possibilidade lógica de existência, vemos o assunto aritmética como a atividade performada em obediência aos princípios diretivos. Uma característica marcante da análise desenvolvida por Hellman no contexto aritmético é sua direta adaptação para outros ramos da matemática, em especial a análise real e a teoria de conjuntos. Uma crítica à proposta de Hellman é a falta de clareza quanto em qual medida, se em alguma, a noção de possibilidade presente no trabalho de Hellman é vantajosa, para a aritmética e análise, em comparação àquela presente no estruturalismo que endossa o realismo de estruturas.

Mesmo que Hellman obtenha sucesso e sua proposta se revele superior as outras tradições estruturalistas, se o custo de fixação da verdade de sentenças da aritmética, de modo objetivo e não comprometida com o realismo, é a adoção de uma estratégia escorada em lógica de segunda ordem acrescida tanto do aparato modal S_5 sem fórmula de Barcan quanto de cuidados especiais no emprego do axioma esquema da compreensão, além de elementos de cunho mereológicos e axiomas de Peano em segunda ordem, então somos bastante reticentes em aceitar a posição defendida por Hellman.

Estamos convencidos que o quadro normativo viabiliza a fixação do valor de verdade das sentenças aritméticas, de modo muito mais natural e sem o comprometimento com objetos abstratos, do que o oferecido por Hellman. Para a lógica de base basta uma linguagem infinitária que, de acordo com as métricas expostas na seção anterior, é bastante próxima da linguagem de primeira ordem. Em vez de elementos categóricos que impõem condições artificiais sobre a lógica de base, são considerados princípios diretivos que são harmônicos com a história da aritmética, são de natureza algorítmica e facilmente aprendidos, além de todas as demais características elencadas nas seções anteriores.

4. Considerações finais

Este trabalho consiste em uma análise, para a aritmética, da pro-

ISSN: 2317-9570

posta presente em Freire (2018) e também abordada por Ramos e Freire neste volume da Revista de Filosofia Moderna e Contemporânea. Tal proposta contempla o problema da fixação do valor de verdade de sentenças de disciplinas matemáticas interpretadas e advoga que as condições de verdade das sentenças aritméticas são determinadas por um esquema em dupla camada. Uma das camadas é composta pelos princípios diretivos e, a outra, por sistemas formais.

Neste artigo os princípios diretivos foram submetidos à rigoroso escrutínio, o que redundou em uma caracterização do modelo padrão para aritmética (através da noção de conformidade aos princípios). A análise também revelou que é factível adotar tais princípios diretivos para a fixação do modelo padrão e que a ferramenta lógica que viabiliza a análise da verdade aritmética a partir dos princípios requer recursos linguísticos, semânticos e matemáticos muito modestos.

Uma vez verificada a viabilidade de fixação da verdade de sentenças aritméticas *via* princípios diretivos, foram delineadas três propostas bastante populares na filosofia da matemática contemporânea. Isso permitiu ao autor do presente trabalho apresentar, por contraste, algumas de suas impressões acerca da proposta de Freire (2018) e, principalmente, explicitar premissas e comprometimentos assumidos por cada uma das propostas. O resultado dessa comparação é bastante favorável à proposta de cunho normativo e fundada em princípios diretivos.

Há ao menos dois modos naturais de aprofundar a análise da verdade presente neste trabalho. Uma delas diz respeito a compatibilização entre o quadro normativo para a aritmética e uma posição estruturalista modal. Um ponto de partida para análise dessa possível compatibilização pode ser descrito do seguinte modo: se, *à la* Hellman, a avaliação da verdade de sentenças aritméticas deve ser consequência de um padrão de tradução da sentença com relação a todas as estruturas possíveis, então os princípios diretivos podem ser vistos como um critério que seleciona, dentre todas as estruturas possíveis, aquelas que são as estruturas aritméticas apropriadas - as estruturas que se conformam aos princípios. Outra abordagem, certamente interessante e filosoficamente relevante, é a investigação rigorosa dos componentes do pressuposto adotado na investigação das condições de verdade em aritmética. Em particular, uma análise minuciosa e criteriosa da noção de princípio diretivo. Uma tal investigação se faz necessária porque, embora a literatura dedicada à discussão e classificação de normas seja extensa, não dispomos de uma caracterização dessa noção que seja harmônica com o quadro geral das investigações aqui apresentadas. O suprimento de tal demanda, no âmbito aritmético, tem potencial para servir de base para a investigação da noção de princípio diretivo para disciplinas matemáticas interpretadas em geral.

Esses tópicos, filosoficamente intrigantes e intelectualmente desafiadores, são temas para trabalhos vindouros.

Referências

ALMEIDA, Edgar L. B. *Análise das Condições de Verdade e dos Requerimentos Existenciais em Axiomatizações da Aritmética*. 2017. Tese (Doutorado em Filosofia) Instituto de Filosofia e Ciências Humanas, Univesidade Estadual de Campinas (IFCH/Unicamp). Campinas: 2017.

BAYS, Timoty. On Putnam and his Models. *Journal of Philosophy*. v. 98, n. 7, p. 331-350. 2001.

BUTTON, T.; WALSH, S. Structure and Categoricity: Determinacy of Reference and Truth-Value in the Philosophy of Mathematics. *ArXiv e-prints*. 2015. Disponível em <http://adsabs.harvard.edu/abs/ 2015arXiv150100472B>. Acesso em: 24 dezembro, 2018.

EBBINGHAUS, H. -D. *Extended Logics: The General Framework*. In: FEFERMAN, S.; BARWISE, J. (Eds.) Model-Theoretic Logics. Springer. New York: *Perspectives in Mathematical Logic*, p. 25-76. 1985.

EBBINGHAUS, H. -D.; FLUM, J.; THOMAS, W. *Mathematical logic*. New York: Springer. 1996.

DRAKE, Frank. R. *Set Theory, an introduction to large cardinals*. Amsterdam: North Holland. 1974.

FREIRE, Rodrigo A. *Interpretation and Truth in Set Theory*. In: CARNIELLI, W.; MALINOWSKI, J. (Eds.) Contradictions, from Consistency to Inconsistency. Cambridge Univesity Press, p. 183-205. 2018.

FREIRE, Rodrigo A. *Os fundamentos do pensamento matemático no século XX e a relevância fundacional da teoria de modelos*. 2009. Tese

(Doutorado em Filosofia) Instituto de Filosofia e Ciências Humanas, Univesidade Estadual de Campinas (IFCH/Unicamp). Campinas: 2009.

GAIFMAN, Haim. *Non-Standard Models in a Broader Perspective*. In: ENAYAT, A.; KOSSAK, R. (Eds.) Non-standard models of arithmetic and set theory. American Mathematical Society, p. 1-22. 2003.

HELLMAN, Geoffrey. *Structuralism*. In: SHAPIRO, Stewart. (Ed.) The Oxford Handbook of Phylosopy of Mathematics and Logic. Oxford University Press, p. 536-562. 2005.

HELLMAN, Geoffrey. *Mathematics Without Numbers: Towards a Modal-Structural Interpretation*. UK: Oxford University Press. 1993.

KEISLER, Jerome. *Model theory for Infinitary Logic: Logic with countable conjunctions and finite quantifiers*. Amsterdam: North-Holland. 1971.

KLEENE, Stephen Cole. *Mathematical Logic*. Dover Publications. 1967.

KREISEL, Georg. *Informal Rigour and Completeness Proofs*. In: LAKATOS, Imre. (Ed.) Problems in the Philosophy of Mathematics. North-Holland, p. 138-157. 1967.

KUNNEN, K. *Set Theory*. London: College Publications. 2011.

KUNEN, K. *Set Theory: an introduction to independence proofs*. Amsterdam: North-Holland. 1992.

LINDSTRÖM, Per. On Extensions of Elementary Logic. *Theoria*. v. 35, n. 1, p. 1-11. 1969.

MCGEE, Vann. How We Learn Mathematical Language. *The Philosophical Review*. v 106, n. 1. 1997.

PUTNAM, Hillary. Models and Reality. *The Journal of Symbolic Logic*. v. 45, p. 464-482. 1980.

RAMOS, Luiza S. P.; FREIRE, Rodrigo A. Da Semântica para Demonstrações de Consistência e a Volta. *Revista de Filosofia Moderna e Contemporânea*. Brasília, 2018.

SHAPIRO, Stewart. *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford University Press. 1991.

THARP, Leslie H. Which Logic is the Right Logic? *Synthese*. v. 31, p. 1-21. 1975.

Agradecimentos

Agradeço a Rodrigo Freire por apresentar-me ao tema desse artigo e pelas muitas discussões acerca dos assuntos aqui tratados. Agradeço também a Alfredo Roque Freire pela leitura de uma versão prévia do texto e por apontar diversos e valorosos aprimoramentos. Por fim, reconheço que parte significativa do presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, durante meu estágio pós-doutoral no Programa de Pós-Graduação em Filosofia da Universidade de Brasília (PPGFIL/UnB).

O Problema das Justificações Parciais

[The Problem of Partial Justifications]

Alexandre Costa-Leite*

Resumo: É comum que algumas proposições e suas respectivas negações sejam justificadas. Isso dá origem ao problema das justificações parciais. Este é apresentado e três soluções são avaliadas, discutidas e examinadas.
Palavras-chave: justificações parciais, possibilidade, contingência e indeterminação

Abstract: It is common that some propositions and their respective negations be both justified. This gives rise to the problem of partial justification. It is introduced and three solutions are evaluated, discussed and examined.
Keywords: partial justifications, possibility, contingency, indetermination

1 - Introdução

Processos justificativos, ou seja, tentativas de justificar proposições e teorias caracterizam a substância do conhecimento. Justificar significa fundamentar por via de premissas e suportes de qualquer espécie para que certas proposições possam ser aceitas como verdadeiras (ou falsas) por um agente específico ou por uma comunidade. Algumas vezes, as justificações se apresentam na forma de demonstrações e, como tais, podem ser consideradas justificações fortes e totais, uma vez que, se corretas, asseguram que a proposição justificada é, necessariamente, verdadeira. Outras vezes, as justificações não são suficientes para certificar a existência de uma proposição verdadeira, certa e segura. Nesse caso, elas servem apenas para estabelecer que se trata de uma proposição contingente. Tais justificações são chamadas de justificações fracas, parciais, debilitadas ou incompletas.[1]

Muito já foi dito e pensado sobre a noção de *justificação* na história da filosofia. O mesmo não pode ser afirmado sobre as lógicas da justificação. Como exemplos de investigações no contexto de sistemas formais para a justificação, temos as pesquisas e abordagens realizadas, bem recente-

*Professor adjunto do Departamento de Filosofia da Universidade de Brasília (UnB). E-mail: costaleite@unb.br.
[1] Cf. [4], [2] e [13].

ISSN: 2317-9570

DOI: https://doi.org/10.26512/rfmc.v6i2.22104

mente, por Newton da Costa em [4] e Sergei Artemov em [1]. Por um lado, a abordagem de da Costa é solitária. O autor desenvolve no livro supracitado as suas lógicas da justificação como lógicas modais. Por outro lado, o trabalho de Artemov tem atualmente atraído grande atenção por parte de pesquisadores, sobretudo, no domínio da ciência da computação. Artemov enriquece as lógicas da justificação com termos de tal modo que é possível fazer operações entre justificações. [2]

As duas abordagens mencionadas acima formalizam, cada uma a sua maneira, dois tipos distintos de justificação.[3] As justificações fortes (totais, completas e perfeitas) são as demonstrações efetivadas no interior de sistemas formais, e elas têm algumas propriedades essenciais:[4] em geral, (i) a demonstração de uma proposição implica que essa proposição é verdadeira, (ii) se uma proposição é fortemente justificada, então a sua negação não é fortemente justificada e, finalmente, (iii) se ocorre a justificação de uma proposição e a justificação de sua negação, então qualquer proposição pode ser justificada e derivada. As justificações fracas (parciais, incompletas e imperfeitas) são as mais comuns e ocorrem nas ciências, na filosofia, na política, na história, nos debates etc. Elas não possuem as propriedades (i)-(iii).

O problema avaliado neste texto pode ser formulado nas duas abordagens formais para a justificação. Concentro a atenção, entretanto, nas lógicas da justificação propostas por da Costa e mostro como o problema pode ser formulado, *mutatis mutandis*, para os sistemas de Artemov. Todavia, não apresento soluções para estes.

2 - O problema das justificações parciais

Considere uma proposição φ e também a sua negação $\neg\varphi$. Lembrando que Newton da Costa utiliza J em [4] para formalizar a ideia de *justificação forte*. Para as justificações fortes, não é possível que ocorra algo do tipo $J\varphi \wedge J\neg\varphi$ já que isso implicaria $\neg J\neg\varphi \wedge J\neg\varphi$, ou seja, uma contradição e, a partir desse ponto tudo pode ser derivado.[5] A questão central deste artigo surge no contexto das lógi-

[2]Com a finalidade de evitar redundância na exposição, remeto o leitor ao artigo [2], o qual contém material suficiente para o entendimento deste texto. Para um estudo aprofundado sobre as diferentes lógicas da justificação, no segundo sentido, o leitor pode consultar [1].

[3]Detalhes em [2].

[4]Algumas dessas propriedades aparecem em forma axiomática no sistema de lógica da justificação proposto por da Costa em [4].

[5]O argumento é bem simples. Detalhes podem ser encontrados em [4], p.61.

ISSN: 2317-9570

cas da justificação propostas por Newton da Costa:

> [...] se nosso objetivo for a edificação de uma lógica da justificação fraca, na qual $J\alpha \wedge J\neg(\alpha)$ não cause problemas de trivialização, quer levando à demonstração de qualquer fórmula do tipo $J\beta$ ou à trivialização pura e simples, deve-se enfraquecer CJ. (Newton da Costa, p.61, in [4])

Vou chamar tal questão de *o problema das justificações parciais* (PJP): como é possível que proposições contraditórias sejam simultaneamente justificadas sem que isso destrua o sistema de inferências? Quais caminhos seguir para a construção de sistemas lógicos capazes de modelar o comportamento das justificações parciais? Avaliar, entender e apresentar três abordagens para esse problema é o interesse aqui.

3 - Perspectivas e sistemas de referência

Antes de abordar (PJP), é conveniente elucidar um ponto. Uma justificação total parece ser mais potente que as demais. As demonstrações são justificações fortes e, como tais, são fatuais, coerentes e não explosivas. Entretanto, surge um problema metodológico.

Considere o princípio da não contradição formulado do seguinte modo: $\neg(\varphi \wedge \neg\varphi)$. Na lógica clássica, estamos falando de uma tautologia, isto é, uma fórmula verdadeira para toda valoração. E isso pode ser verificado semanticamente e também por via de métodos de prova. Assim, podemos dizer que há uma justificação forte para tal princípio e, com isso, podemos afirmar que vale o princípio da não contradição. Considere agora a lógica trivalente $Ł_3$ de Łukasiewicz tal como definida em [10] e [6]. Usando as $Ł_3$-valorações é fácil ver que o princípio da não contradição obtém valor não designado para alguma valoração e, como tal, é inválido. Consequentemente, podemos também dizer que há uma justificação forte para a invalidade do princípio da não contradição.[6]

Diante do exposto, pode-se pensar então que há uma justificação para a validade do princípio da não contradição e que há, simultaneamente, uma justifica-

[6]Esse exemplo peculiar foi aqui selecionado por ser de fácil compreensão. Todavia, muitos outros argumentos semelhantes podem ser encontrados tanto na lógica quanto na matemática, como é o caso de algumas proposições da geometria euclidiana e suas rivais como as geometrias elíptica e hiperbólica.

ção para a sua invalidade. Mas justificar uma proposição e justificar também a sua negação é uma característica das justificações fracas. No entanto, no caso acima, estamos diante de justificações fortes que se manifestam no interior de sistemas formais. O que está acontecendo?

A explicação é a seguinte: os sistemas de referência que são usados como parâmetros para a especificação da justificação são distintos, dado que, no primeiro caso, estamos jogando com as regras da lógica clássica e, no segundo, com as regras da lógica trivalente. Notamos, com isso, que não existem justificações totais que sejam absolutas e que justifiquem independentemente de qualquer perspectiva ou sistema de referência. Por conseguinte, uma justificação forte deve ser entendida como forte somente em relação a um dado conjunto determinado de regras. De um ponto de vista universal e geral, *toda justificação pode ser vista como justificação parcial*. Isso retira a hegemonia das justificações fortes. Mas é em um sistema de referência específico que uma justificação é total ou não. Por isso, para julgar uma justificação como total ou parcial é preciso saber, antes, qual o sistema de referência que está sendo pressuposto.

4 - Justificações parciais e a possibilidade lógica

Em [2] (pág. 178), sugeri modos para a construção de sistemas de inferência capazes de modelar justificações fracas: um modo, digamos, fundado na paraconsistência e outro no tratamento da justificação fraca como um tipo de operador de possibilidade ($\Diamond$), mas não apresentei os detalhes. Faço isso agora, começando pela abordagem da justificação parcial como possibilidade. A ideia é a de que podemos pensar a justificação parcial como um tipo de possibilidade lógica (esse tipo de estratégia redutiva utilizei também no tratamento do conceito de *imaginação* como possibilidade lógica em [3]).

Considere um operador qualquer $*$ e as seguintes definições (estas formulam (i)-(iii) acima com mais precisão):

1. $*$ é fatual se, e somente se, $*\varphi \rightarrow \varphi$;
2. $*$ é coerente se, e somente se, $*\varphi \rightarrow \neg * \neg\varphi$;
3. $*$ é explosivo[7] se, e somente se, $*\varphi, *\neg\varphi \vdash \psi$.

[7] O conceito de *explosão* é, em geral, usado em outro sentido na literatura (cf. [5]). Aqui a noção é definida em um sentido peculiar. Se $*\neg$ fosse equivalente a $\neg*$, teríamos o sentido comum de explosão.

Para fornecer uma resposta ao (PJP), deve-se notar que um operador para modelar a justificação parcial não pode ser explosivo (no sentido acima de explosão) e não pode ser, idealmente, nem fatual e nem coerente. Em algumas situações, operadores fatuais implicam operadores coerentes.

Considere ainda a lógica modal $\boldsymbol{S5}$ e seus respectivos $\Box$ e $\Diamond$.[8] Assim, por um lado, levando em conta algumas das propriedades de $\Box$ nessa lógica, podemos concluir que $\Box$ de $\boldsymbol{S5}$ é fatual, uma vez que $\Box\varphi \to \varphi$. Além disso, $\Box$ é também coerente, pois $\Box\varphi \to \neg\Box\neg\varphi$ (isto é, $\Box\varphi \to \Diamond\varphi$), e é explosivo, já que $\Box\varphi, \Box\neg\varphi \vdash \psi$. Daí, se a base da lógica for clássica, o operador $\Box$ de $\boldsymbol{S5}$ não pode ser naturalmente visto como um operador capaz de modelar a justificação fraca, embora em algumas de suas variações ele seja utilizado para pensar a justificação forte (e esse é o caminho escolhido por da Costa em [4].). Por outro lado, com o uso da semântica dos mundos possíveis, o operador $\Diamond$ de $\boldsymbol{S5}$ não é fatual, dado que não vale $\Diamond\varphi \to \varphi$. Do mesmo modo, não há qualquer compromisso com a coerência, pois $\Diamond\varphi \to \neg\Diamond\neg\varphi$ não é válida. Igualmente, o operador $\Diamond$ não é explosivo, já que $\Diamond\varphi, \Diamond\neg\varphi \nvdash \psi$. Como consequência, o operador $\Diamond$ de $\boldsymbol{S5}$ parece ser um candidato razoável para modelar a justificação fraca.

De fato, quando estamos perante uma justificação fraca, não temos elementos suficientes para garantir que estamos lidando com uma proposição seguramente verdadeira. Se a justificação for parcial, então não há certeza sobre a verdade da proposição que é parcialmente justificada. Ela *apenas* pode ser verdadeira. A relação existente entre justificação parcial e possibilidade lógica parece evidente, visto que uma proposição parcialmente justificada é simplesmente possível. E a possibilidade lógica, não sendo global, exige, portanto, apenas verdade em alguma situação. Se uma proposição está fracamente justificada, então ela pode ser falsa. Esse fato está imediatamente conectado com a invalidade de $\Diamond\varphi, \Diamond\neg\varphi \vdash \psi$. Daí, o tratamento da justificação fraca como possibilidade lógica parece fornecer uma solução ao (PJP).[9]

Ainda conectado ao tema da possibilidade lógica, é interes-

[8] Para um estudo recente em lógica modal, o leitor pode verificar [15].

[9] Certamente, se a justificação fraca pode ser vista como a possibilidade em $\boldsymbol{S5}$, então deve existir uma ligação natural entre as justificações fracas e o sistema de lógica discussiva de Jaśkowski (Cf. [8]), ou seja, conjecturo que podemos pensar a lógica discussiva como uma lógica das justificações parciais. Fundamentar essa conjectura é algo que somente poderá ser feito em outra ocasião, contudo.

sante considerar quando o operador de contingência ∇ é tomado como primitivo em $S5$(tal como desenvolvido em [11]).[10] Esse operador engendra uma noção de justificação parcial. Novamente, por via da semântica dos mundos possíveis, é fácil ver que o operador ∇ não é fatual já que $\nabla\varphi \rightarrow \varphi$ não é fórmula válida. Também não é coerente pois $\nabla\varphi \rightarrow \neg\nabla\neg\varphi$ não é o caso. E, por fim, $\nabla\varphi, \nabla\neg\varphi \nvdash \psi$. Segue-se que é natural tratar a justificação parcial como contingência, já que esse operador satisfaz a exigência que permite uma abordagem formal ao conceito de *justificação fraca*.

O fato de a justificação ser parcial implica que a proposição justificada não é uma verdade necessária, dado que a justificação pode estar errada ou ser insuficiente uma vez que não justifica plenamente. Notemos, entretanto, que algumas proposições possíveis são também necessárias. Logo, pode ser mais útil caracterizar a justificação fraca como contingência, pois contingência não implica necessidade. Isso revela a dimensão vaga, incerta e imprecisa desse tipo de justificação.

5 - Justificações parciais e a indeterminação

As justificações parciais e o respectivo (PJP) encontram uma abordagem interessante quando a justificação fraca é vista como um tipo de operador trivalente que pode ser definido no interior da lógica $Ł_3$ de Łukasiewicz.[11] Łukasiewicz em [9] inicialmente sugere que esse valor extra capturado por 1/2 seja visto como *possibilidade*. Leituras mais tardias mostram que tal valor pode também ser visto como *indeterminação*.[12] Esse terceiro valor serve como base para que seja definido na linguagem de $Ł_3$ um operador trivalente I. Malinowski o interpreta como *é contingente que* e ainda como *é modalmente indiferente*. Epstein indica que tal operador é o próprio *indeterminado*. Tal conceito serve aqui para o objetivo de pensar a justificação fraca como o operador de indeterminação em $Ł_3$.

No ambiente trivalente de $Ł_3$, são conhecidas as definições dadas para o $\Diamond$ e o $\Box$. A partir de tais definições, o operador I é estabelecido por via da seguinte valoração: $v(\varphi) = 1/2$ se, e somente se, $v(I\varphi) = 1$. Esse operador reflete

[10] Propriedade básica nas lógicas da contingência é que $\nabla\varphi \leftrightarrow \nabla\neg\varphi$.

[11] Para um estudo sobre a hierarquia de lógicas de Łukasiewicz, o leitor pode consultar [10] e [6].

[12] Ver [10], pág. 18.

ISSN: 2317-9570

o funcionamento interno das justificações parciais, considerando que ele não é fatual, coerente e nem explosivo. Em $Ł_3$, o único valor que corresponde à verdade, ou seja, o único valor que é designado é 1. Para ver que I não é fatual, ou seja, que não vale $I\varphi \to \varphi$ considere $v(\varphi) = 1/2$. O operador I também não é coerente, já que $I\varphi \to \neg I\neg\varphi$ tem valor 0 se $v(\varphi) = 1/2$. Do mesmo modo, I não é explosivo pois $I\varphi, I\neg\varphi \nvdash \psi$ quando $v(\varphi) = 1/2$ e $v(\psi) = 1/2$. Por conseguinte, parece razoável que a justificação parcial seja pensada e concebida como um tipo de indeterminação (e essa leitura é compatível com as anteriores que tomam justificações parciais como operadores de possibilidade e contingência, respectivamente).

Com efeito, se a justificação for fraca, então não temos razões suficientes para defender que a proposição parcialmente justificada seja completamente determinada, isto é, estamos perante uma proposição indeterminada (meramente possível, portanto). A verdade de uma proposição apenas parcialmente justificada é incerta, indeterminada. Essa leitura do terceiro valor e do operador de indeterminação no âmbito das justificações fracas pode ser vista como mais uma motivação para as lógicas polivalentes em geral.[13]

O (PJP) pode ser formulado e estudado também no horizonte das famosas lógicas da justificação ao estilo Artemov e sua escola. De fato, recentemente, Che-Ping Su aborda em [14] uma versão bastante parecida com o problema apresentado por da Costa em [4], mas levando em conta apenas justificações fortes. O objetivo de Su é encontrar lógicas que não satisfazem $(\Box\varphi \wedge \Box\neg\varphi) \to \Box\psi$. Isso é o (PJP) formulado por da Costa para as justificações fortes, considerando que o operador J é exatamente o operador $\Box$. Mas, no contexto das lógicas de Artemov, o problema vai ser formulado com outra cara: não deve valer $(j : \varphi \wedge j' : \neg\varphi) \to j'' : \psi$.[14] Su então elabora uma lógica paraconsistente trivalente da justificação para responder ao problema e nessa lógica o terceiro valor utilizado é o b que representa *ambos verdadeiro e falso*.

[13] Um caminho que pode ser investigado nessas relações entre polivalência e justificação é o de estabelecer níveis de justificação. Nesse sentido, (PJP) pode também encontrar uma resposta no trabalho iniciado em [2]. Outro caminho para o entendimento das justificações parciais pode ser elaborado no estilo desenvolvido em [13] quando semânticas tetravalentes são utilizadas.

[14] cf. [14]. Além disso, é importante notar que Artemov em [1] não faz uso daquilo que ele chama de *factividade* $j : \varphi \to \varphi$ e de regras de inferência para polinômios de prova. Assim, consegue pensar justificações em geral. Uma investigação completa do problema das justificações parciais no contexto de outras lógicas (e também na abordagem de Artemov) é tema fértil que pretendo desenvolver detalhadamente em outra ocasião.

Há, de fato, muitas outras maneiras de resolver o problema das justificações parciais no interior da lógicas da justificação. Se se pretende manter uma leitura universal do operador ou termo de justificação, então nenhuma das vias apontadas acima é pertinente (possibilidade, contingência e indeterminação). Pelo contrário, deve-se alterar a lógica subjacente. Assim, em direção similar a de Su, o funtor de paraconsistentização desenvolvido em [5] poderia ser aplicado às hierarquias de lógicas da justificação encontradas em [1] para uma produção de lógicas paraconsistentes da justificação. Um caminho mais simples poderia ser seguido substituindo, por exemplo, a base clássica da lógica da justificação forte de da Costa por uma lógica paraconsistente.

6 - Conclusão

Vimos qual é o problema das justificações parciais (PJP) e três maneiras de abordá-lo, todas intrinsecamente relacionadas. Mesmo sendo possível fornecer abordagens para (PJP), não há como evitar o fato de que justificações fracas, parciais, debilitadas e incompletas existem. Elas estão por todos os lados e são frequentes em domínios como a filosofia, ciências acerca da realidade e em todos os aspectos de nossas vidas. Tais justificações não são fatuais e nem são coerentes e, muito provavelmente, são as responsáveis por gerar teorias conflitantes e divergências de toda ordem. Nesse contexto, não há qualquer possibilidade de um conhecimento estável e claro, e não estamos em posição teórica razoável para legislar acerca do mundo em seus variados aspectos. O que temos, sim, é apenas um conjunto de afirmações parcialmente justificadas. Estamos, portanto, no âmbito das justificações de sentenças contraditórias.

Isso leva céticos ao ato de suspensão do juízo:

> According to the mode deriving from dispute, we find that undecidable dissension about the matter proposed has come about both in ordinary life and among philosophers. Because of this we are not able either to choose or to rule out anything, and we end up with suspension of judgement. (Sexto Empírico, p.41, in [7])

A citação acima nos conduz até Oswaldo Porchat Pereira:

> Na leitura de Sexto Empírico, encontrei a ocasião de confirmar minha experiência do conflito insupe-

ISSN: 2317-9570

> rável dos dogmatismos, de sua perpétua *diaphonía*. Quem longamente meditou sobre as *Hipotiposes* não mais ousará cometer-se à edição do Discurso derradeiro. (Oswaldo Porchat Pereira, p.31, in [12])

Essa *diaphonía* pode ser entendida como a existência da pluralidade de justificações parciais para proposições contraditórias. Tal postura se apresenta como caminho natural se nossas teorias não são justificadas totalmente. Como o dispositivo epistêmico usado para doar verdade às proposições é vago e impreciso, como é o caso das justificações fracas, não temos elementos para saber qual é o caminho correto. Igualmente, uma vez que não há justificação total que seja absoluta, o caminho é impreciso até mesmo no âmbito das ciências formais.

Nesse sentido, do ponto de vista de agentes precários como são os agentes humanos, é curioso tentar saber se é possível a existência de uma ciência livre da pluralidade das justificações parciais.

Referências

[1] ARTEMOV, S. (2008). The logic of justification. In *The Review of Symbolic Logic*, 1(4), pp. 477-513.

[2] COSTA-LEITE, A. (2014). Lógicas da justificação e quase-verdade. *Principia: an international journal of epistemology*, 18 (2), pp.175-186.

[3] COSTA-LEITE, A. (2010). Logical properties of imagination. *Abstracta*, 6(1), pp. 103-116.

[4] DA COSTA, N. (1999). *O conhecimento científico*, 2a. edição, São Paulo: Discurso Editorial.

[5] DE SOUZA, E. G.; COSTA-LEITE, A; DIAS, D.H.B. (2016). On a paraconsistentization functor in the category of consequence structures. *Journal of Applied Non-Classical Logics*, 26(3), pp.240-250;

[6] EPSTEIN, R. (2001). *Propositional Logics: The Semantic Foundations of Logic*. Belmont: Wadsworth/Thomson Learning.

[7] EMPÍRICO, S. (2005). *Outlines of scepticism*. Editado por Julia Annas e Jonathan Barnes. Cambridge: Cambridge University Press.

[8] JAŚKOWSKI, S. (1948). Propositional calculus for contradictory deductive systems. *Logic and Logical Philosophy*, 7, 35-56.

[9] ŁUKASIEWICZ, J. (1970). On three-valued logic. In: *Jan Łukasiewicz: Selected Works*, editado por L. Borkowski. Warsawa: Noth-Holland Publishing. Original: 1920.

[10] MALINOWSKI, G. (1993). *Many-Valued Logics*. New York: Oxford University Press.

[11] MONTGOMERY, H; ROUTLEY, R. (1966). Contingency and non-contingency bases for normal modal logics. In: *Logique et Analyse*, vol.9, No 35-36, 318-328.

[12] PORTCHAT PEREIRA, O. (2006). Rumo ao ceticismo. São Paulo: Editora da UNESP.

[13] SCHANG, F; COSTA-LEITE, A. (2016). Une sémantique générale des croyances justifiées. *CLE e-prints*, 16(3), pp.1-24.

[14] SU, C-P. (2014). Paraconsistent Justification Logic: a Starting Point. *Advances in Modal Logic*. London: College Publications, pp. 513-532.

[15] VAN BENTHEM, J. (2010). Modal logics for open minds. Stanford: CSLI Publications.

Sobre a Noção Categorial de Proto-topos

[On the Categorical Notion of Proto-topos]

Edelcio G. de Souza*

Resumo: O objetivo da presente trabalho é mostrar como é possível fazer semântica para linguagens proposicionais em ambientes categoriais que não sejam *topos*. Proponho a definição de dois tipos de categorias denominadas *categorias com morfismos verdade* (CTM) e *proto-topos*. Em categorias com morfismos verdade, pode-se definir as "funções de verdade" que correspondem aos conectivos lógicos de negação, conjunção, implicação e disjunção. Em proto-topos, pode-se mostrar que as "funções de verdade" assim definidas satisfazem certas propriedades desejáveis com respeito aos valores de verdade *verdadeiro* e *falso*.
Palavras-chave: linguagens proposicionais, semântica categorial, morfismos verdade, proto-topos

Abstract: The aim of this paper is to show how is possible to do semantic for propositional languages in a categorical setting different from topos. I propose the definition of two kinds of categories called *categories with truth morphisms* (CTM) and *proto-topos*. In categories with truth morphisms, it can be defined the truth functions that correspond to the logical connectives of negation, conjunction, implication and disjunction. In proto-topos, I show that the truth functions defined in CTM's satisfy certain desirable properties with respect to the truth values true and false.
Keywords: propositional languages, categorical semantics, truth morphisms, proto-topos

1. Introdução

Na segunda metade do século XX, com os trabalhos de Lawvere, Tierney e outros, foi definida um importante tipo de categoria denominada *topos* (elementar)[1]. Em topos, pode-se desenvolver completamente a semântica para as linguagens de primeira ordem, junto com os conceitos relacionados de consequência e validade lógica. Descobriu-se, então, que as sentenças dessas linguagens que são válidas em topos são precisamente as teses da lógica intuicionista. Um resultado importante e, de certo modo, surpreendente e filosoficamente relevante[2].

*Edelcio Gonçalves de Souza, Departamento de Filosofia - FFLCH - USP, edelcio.souza@usp.br.

[1] Ver Lawvere 1964, Lawvere 1970 e Tierney 1972.

[2] Ver Boileau-Joyal 1981. Para a lógica intuicionista ver Heyting 1966.

ISSN: 2317-9570

DOI: https://doi.org/10.26512/rfmc.v6i2.22103

Todavia, quando nos restringimos apenas às linguagens proposicionais standard e, portanto, à definição dos morfismos verdade usuais, percebi que não era necessário operar em um ambiente categorial de topos. Procurei, então, responder à seguinte questão: *Qual é o ambiente categorial mais simples que seja capaz de definir os morfismos verdade usuais que correspondem à negação, conjunção, implicação e disjunção?*

A resposta à questão acima formulada é a definição que proponho de dois tipos de categorias que constituem generalizações da noção de topos. Estas são as *categorias com morfismos verdade* (CTM) e os *proto-topos*.

2. Topos

Uma categoria $\mathcal{C}$ é dita *finitamente bicompleta* (completa e cocompleta) quando todo diagrama finito em $\mathcal{C}$ possui limite e colimite. Nesse caso, $\mathcal{C}$ possui objetos terminais e iniciais e, em $\mathcal{C}$, pode-se fazer uma série de construções úteis: produtos e coprodutos de objetos e morfismos, equalizadores e coequalizadores, pullbacks e pushouts, etc[3].

Diz-se que uma categoria $\mathcal{C}$ *possui exponenciação* se, para cada $\mathcal{C}$-objeto a, temos um funtor exponencial $(-)^a : \mathcal{C} \to \mathcal{C}$ que é adjunto à direita do funtor $(-) \times a$. Em termos elementares, isso significa que $\mathcal{C}$ tem produto para cada par de $\mathcal{C}$-objetos e, para quaisquer $\mathcal{C}$-objetos dados a e b, existe um $\mathcal{C}$-objeto, dito a *exponencial* de b por a, denotado por b^a, e um $\mathcal{C}$-morfismo $ev_{ab} : b^a \times a \to b$, denominado *morfismo avaliação*, tal que vale a seguinte propriedade universal: Para todo $\mathcal{C}$-objeto c e $\mathcal{C}$-morfismo $g : c \times a \to b$, existe um único $\mathcal{C}$-morfismo $\hat{g} : c \to b^a$ tal que o seguinte diagrama comuta:

$$\begin{tikzcd} b^a \times a \arrow[dr, "ev_{ab}"] & \\ & \mathrm{b} \\ c \times a \arrow[uu, dashed, "\hat{g} \times 1_a"] \arrow[ur, "g"'] & \end{tikzcd}$$

isto é, $ev_{ab} \circ (\hat{g} \times 1_a) = g$. Segue-se, imediatamente, que existe uma correspondência biunívoca entre as classes $\mathcal{C}(c \times a, b)$ e $\mathcal{C}(c, b^a)$[4]. Quando uma categoria $\mathcal{C}$ é finitamente completa e possui exponenciação, então $\mathcal{C}$ é denominada uma *categoria cartesiana fechada*.

Considere um objeto d em uma categoria $\mathcal{C}$. Um subobjeto de d é uma certa classe de equivalência de monomorfismos com codomínio d. Sejam os $\mathcal{C}$-monomorfismos $f : a \rightarrowtail d$ e $g : b \rightarrowtail d$. Dizemos

[3] Vou utilizar a notação de Goldblatt 2006. Para textos de teoria de categorias em geral, pode-se consultar Arbib-Manes 1975, Awodey 2010, Herrlich-Strecker 1973 e MacLane 1971.

[4] Se $\mathcal{C}$ é uma categoria e a e b são dois $\mathcal{C}$-objetos, então $\mathcal{C}(a,b)$ denota a classe de todos os $\mathcal{C}$-morfismos de a em b.

ISSN: 2317-9570

que f é equivalente a g, em símbolos $f \equiv g$ se existe um isomorfismo $h : a \to b$ tal que o diagrama

$$\begin{tikzcd} a \arrow[d, "h"'] \arrow[dr, "f"] & \\ b \arrow[r, "g"'] & \mathrm{d} \end{tikzcd}$$

comuta, isto é, $f = g \circ h$.

É fácil ver que $\equiv$ é uma relação de equivalência e, se $Mon(d)$ é a classe dos monomorfismos com domínio d, então um *subobjeto* de d é qualquer elemento do conjunto quociente:

$$Mon(d)/\equiv := \{[f]_\equiv : f \in Mon(d)\}.$$

Se $\mathcal{C}$ é uma categoria com objeto terminal 1, então um *classificador de subobjetos* para $\mathcal{C}$ é um $\mathcal{C}$-objeto Ω junto com um $\mathcal{C}$-morfismo $\top : 1 \to \Omega$ tal que satisfaz a seguinte propriedade universal:

(Ω-Axioma). Para cada monomorfismo $f : a \rightarrowtail b$ existe um único $\mathcal{C}$-morfismo $\chi_f : d \to \Omega$, denominado o *morfismo característico* ou o *caráter* de f, tal que o diagrama

$$\begin{tikzcd} a \arrow[r, "f"] \arrow[d, "!"'] & d \arrow[d, "\chi_f"] \\ 1 \arrow[r, "\top"'] & \Omega \end{tikzcd}$$

é um quadrado pullback.

Com os ingredientes introduzidos acima pode-se definir a noção categorial de topos (elementar).

Definição 1 (Lawvere e Tierney) *Um **topos** (elementar)*[5] *é uma categoria $\mathcal{E}$ tal que:*

i. $\mathcal{E}$ é finitamente completa;

ii. $\mathcal{E}$ é finitamente cocompleta;

iii. $\mathcal{E}$ tem classificador de subobjetos;

iv. $\mathcal{E}$ possui exponenciação.

Depois que essa definição de topos foi estabelecida, C. J. Mikkelson descobriu que a condição *ii.* é implicada pelas outras condições[6]. Assim, um topos pode ser definido como uma categoria cartesiana fechada com classificador de subobjetos[7].

Um aspecto que vai se revelar importante para o presente trabalho é uma propriedade fundamental de categorias cartesianas fechadas que provém, utilizando-se exponenciação, da correspondência biunívoca entre as classes $\mathcal{C}(c \times a, b)$ e $\mathcal{C}(c, b^a)$, para todo $\mathcal{C}$-objeto a e b.

Seja $\mathcal{C}$ uma categoria cartesiana fechada e 0 é um objeto inicial em $\mathcal{C}$, então se existe um $\mathcal{C}$-morfismo $f : a \to 0$, tem-se que f é

[5]A palavra "elementar" é introduzida para distinguir essa noção da noção de topos original devida a A. Grothendieck (ver SGA4). No que se segue, a palavra "elementar" sempre será omitida.

[6]Ver Pare 1974.

[7]Para uma introdução à teoria de topos, ver Bell 2008, Freyd 1972, Johnstone 1977, Johnstone 2002, Kock-Wraith 1971, MacLane-Moerdijk 2012 e MacLarty 1995.

um isomorfismo (e, assim, a é inicial em $\mathcal{C}$).

Segue-se da propriedade fundamental acima que: *Em uma categoria cartesiana fechada* $\mathcal{C}$*, todo* $\mathcal{C}$*-morfismo* $0 \to a$*, cujo domínio é um objeto inicial em* $\mathcal{C}$*, é um monomorfismo.*

Quando a categoria $\mathcal{E}$ é um topos, várias propriedades podem ser estabelecidas. Mencionarei algumas que serão relevantes nas construções que se seguem[8].

Teorema fundamental da teoria de topos - Freyd. Se $\mathcal{E}$ é um topos e a é um $\mathcal{E}$-objeto, então a categoria slice $\mathcal{E} \downarrow a$ é também um topos.

Segue-se do teorema fundamental dois fatos básicos enunciados a seguir.

Fato 1. Em $\mathcal{E}$, pullbacks preservam epimorfismos. Isto significa que se o diagrama

$$\begin{tikzcd} a \arrow[r] \arrow[d, "g"'] & b \arrow[d, "f"] \\ c \arrow[r] & d \end{tikzcd}$$

é um quadrado pullback e f é um epimorfismo, então g é também um epimorfismo.

Fato 2. Em $\mathcal{E}$, coprodutos preservam pullbacks. Isto é, se os quadrados do diagrama

$$\begin{tikzcd} a \arrow[r, "\mathrm{f}"] \arrow[d, "\mathrm{g}"'] & d \arrow[d, "k"] & a' \arrow[l, "f'"'] \arrow[d, "g'"] \\ b \arrow[r, "h"'] & e & b' \arrow[l, "h'"] \end{tikzcd}$$

são pullbacks, então o quadrado

$$\begin{tikzcd} a + a' \arrow[r, "{[f, f']}"] \arrow[d, "g + g'"'] & d \arrow[d, "k"] \\ b + b' \arrow[r, "{[h, h']}"'] & e \end{tikzcd}$$

é um pullback também.

3. Linguagens proposicionais

Considere uma *linguagem proposicional standard* $\mathcal{L}$ construída de maneira usual. $\mathcal{L}$ possui, entre seus símbolos, os elementos de um conjunto de *variáveis proposicionais* $Var = \{p_0, p_1, p_2, ...\}$; os *conectivos*: $\neg$ (negação), $\wedge$ (conjunção), $\supset$ (implicação) e $\vee$ (disjunção); e símbolos de *parênteses*: (e). O conjunto For das *fórmulas* de $\mathcal{L}$ é o menor conjunto de expressões de $\mathcal{L}$ que contém Var e tal que valem as seguintes condições:

i. se $\alpha \in For$, então $\neg\alpha \in For$;

ii. se $\alpha, \beta \in For$, então $(\alpha \wedge \beta) \in For$;

iii. se $\alpha, \beta \in For$, então $(\alpha \supset \beta) \in For$;

iv. se $\alpha, \beta \in For$, então $(\alpha \vee \beta) \in For$

[8] Ver Goldblatt 2006 páginas 96 e 114-115.

 Revista de Filosofia Moderna e Contemporânea, Brasília, v.6, n.2, dez. 2018, p. 105-114
ISSN: 2317-9570

Um modo usual de se fazer semântica para $\mathcal{L}$ é por meio do conceito de matriz e valoração. Apresento uma versão simplificada desses conceitos que servirá para contextualizar as construções que se seguem.

Uma *matrix* para $\mathcal{L}$ é uma tripla

$$M = \langle Val, t, \{f_\neg, f_\wedge, f_\supset, f_\vee\}\rangle$$

em que Val é um conjunto, $t \in Val$, $f_\neg : Val \to Val$ e $f_\wedge, f_\supset, f_\vee : Val \times Val \to Val$.

Dado uma matriz M para $\mathcal{L}$ dizemos que $v : For \to Val$ é uma *valoração* para $\mathcal{L}$ se valem as seguintes condições:

i. $v(\neg\alpha) = f_\neg(v(\alpha))$;
ii. $v(\alpha \wedge \beta) = f_\wedge(v(\alpha), v(\beta))$;
iii. $v(\alpha \supset \beta) = f_\supset(v(\alpha), v(\beta))$;
iv. $v(\alpha \vee \beta) = f_\vee(v(\alpha), v(\beta))$.

Se α é uma fórmula de $\mathcal{L}$, dizemos que α é *M-válida*, em símbolos $M \models \alpha$, se $v(\alpha) = t$ para toda M-valoração v.

Em um ambiente categorial $\mathcal{E}$ (digamos um topos), uma matriz com base em $\mathcal{E}$ é uma tripla

$$M = \langle \mathcal{E}(1,\Omega), \top : 1 \to \Omega, \{f_\neg, f_\wedge, f_\supset, f_\vee\}\rangle$$

em que $\mathcal{E}(1,\Omega)$ é a classe dos morfismos de 1 no classificador de subobjetos, $\top : 1 \to \Omega$ é seu morfismos associado e $f_\neg : \Omega \to \Omega$ e $f_\wedge, f_\supset, f_\vee : \Omega \times \Omega \to \Omega$ são morfismos a serem definidos.

Vou mostrar que esses morfismos podem ser definidos em categorias $\mathcal{E}$ que são mais gerais que topos de tal modo que temos uma noção de $\mathcal{E}$-validade que implica a existência de uma lógica associada à $\mathcal{E}$.

4. Categorias com morfismos verdade

Passo, então, a introduzir um tipo de categoria em que os morfismos $f_\neg$, $f_\wedge$, $f_\supset$ e $f_\vee$ podem ser definidos. (Na realidade, as definições são as mesmas que encontramos para topos. A diferença é que a categoria de base é mais geral.)

Definição 2 *Uma categoria $\mathcal{E}$ é denominada uma **categoria com morfismos verdade** (CTM) se as seguintes condições forem satisfeitas:*

CTM1. $\mathcal{E}$ é finitamente completa;

CTM2. $\mathcal{E}$ é finitamente cocompleta;

CTM3. $\mathcal{E}$ tem classificador de subobjetos;

CTM4. (0-AXIOMA). Em $\mathcal{E}$, se existe um morfismo $a \xrightarrow{f} 0$ (0 é inicial em $\mathcal{E}$), então f é um isomorfismo[9].

Vou introduzir uma nomenclatura padrão. Se a é um $\mathcal{E}$-objeto,

[9] A propriedade fundamental das categorias cartesianas fechadas mencionada acima torna-se o 0-Axioma das CTM's.

então um morfismo $1 \to a$ (1 é terminal em $\mathcal{E}$) é dito um *elemento* de a. Nessa direção, o classificador de subobjetos Ω será denominado o *objeto valor de verdade* e $\top : 1 \to \Omega$ é um elemento de Ω dito o valor de verdade *verdadeiro*.

Como dito acima, uma consequência do 0-Axioma é que, em uma CTM $\mathcal{E}$, um morfismo do tipo $0 \to a$ é um monomorfismo, Assim, usando o Ω-Axioma, fica definido um $\mathcal{E}$-morfismo $\perp : 1 \to \Omega$ que é o caráter de $0_1 : 0 \rightarrowtail 1$. (Ver diagrama.)

$$\begin{array}{ccc} 0 & \xrightarrow{0_1} & 1 \\ {\scriptstyle 0_1}\downarrow & & \downarrow{\scriptstyle \chi_{0_1} = \perp} \\ 1 & \xrightarrow[\top]{} & \Omega \end{array}$$

Observe que nada impede que $\top = \perp$. Mas, nesse caso, temos que $\mathcal{E}$ é uma categoria degenerada (todos os objetos de $\mathcal{E}$ são isomorfos entre si). O morfismo $\perp : 1 \to \Omega$ é um elemento de Ω dito o valor de verdade *falso*.

Se a é um $\mathcal{E}$-objeto, denotamos por

$$\underbrace{a \times a \times \ldots \times a}_{n \text{ vezes}} = a^n$$

o produto de n cópias de a (n é um número natural positivo). Dado uma CTM $\mathcal{E}$, um *morfismo verdade* em $\mathcal{E}$ é qualquer morfismo do tipo $\Omega^n \to \Omega$ para algum natural positivo n. (Note que Ω^n não é um exponencial em $\mathcal{E}$.)

Vou mostrar que, em uma CTM $\mathcal{E}$, os morfismos verdade que correspondem à negação, conjunção, implicação e disjunção podem ser definidos de modo usual (em topos).

Seja $\mathcal{E}$ uma CTM com classificador $\top : 1 \to \Omega$. Definimos os seguintes morfismos verdade:

1) $f_\neg : \Omega \to \Omega$ é o caráter de $\perp : 1 \to \Omega$. O morfismo $f_\neg$ é dito o morfismo *negação*.

2) $f_\wedge : \Omega \times \Omega \to \Omega$ é o caráter do morfismo produto $\langle \top, \top \rangle : 1 \to \Omega \times \Omega$. O morfismo $f_\wedge$ é dito o morfismo *conjunção*.

3) $f_\supset : \Omega \times \Omega \to \Omega$ é o caráter do equalizador $e : \rightarrowtail \Omega \times \Omega$ dos morfismos $f_\wedge$ e pr_1 (a primeira projeção do produto). Note que, em qualquer categoria, equalizadores são monomorfismos. O morfismo $f_\supset$ é dito o morfismo *implicação*.

4) $f_\vee : \Omega \times \Omega \to \Omega$ é o caráter da imagem do morfismo coproduto $f = [\langle \top_\Omega, 1_\Omega \rangle, \langle 1_\Omega, \top_\Omega \rangle] : \Omega + \Omega \to \Omega \times \Omega$. O morfismo $f_\vee$ é dito o morfismo *disjunção*.

Cabe mencionar nesse ponto que em CTM's é possível definir para cada morfismo $f : a \to b$ o morfismo imagem de f dado por $f(a) \stackrel{imf}{\to} b$ construído por meio de um equalizador de um pushout de f por f, isto é, um equalizador de um colimite do diagrama $b \stackrel{f}{\leftarrow} a \stackrel{f}{\to} b$.

Com os morfismos verdade assim definidos, podemos mostrar

ISSN: 2317-9570

como os mesmos agem quando aplicados aos valores de verdade $\perp$ e $\top$. Os resultados dessas ações são reunidos nas tabelas usuais de verdade com uma única exceção.

Para o morfismo negação, temos a tabela:

x	$f_\neg \circ x$
$\top$	$\perp$
$\perp$	$\top$

Para os outros morfismos verdade temos:

Para a conjunção

x	y	$f_\wedge \circ \langle x, y \rangle$
$\top$	$\top$	$\top$
$\top$	$\perp$	$\perp$
$\perp$	$\top$	$\perp$
$\perp$	$\perp$	$\perp$

Para a implicação

x	y	$f_\supset \circ \langle x, y \rangle$
$\top$	$\top$	$\top$
$\top$	$\perp$	$\perp$
$\perp$	$\top$	$\top$
$\perp$	$\perp$	$\top$

Para a disjunção

x	y	$f_\vee \circ \langle x, y \rangle$
$\top$	$\top$	$\top$
$\top$	$\perp$	$\top$
$\perp$	$\top$	$\top$
$\perp$	$\perp$	?

Assim, restaria a prova de que $f_\vee \circ \langle \perp, \perp \rangle = \perp$.

A fim de se obter esse resultado, tive que fortalecer a noção de categoria com morfismos verdade adicionando mais dois axiomas.

5. Proto-topos

Passo, então, a definir um tipo de categoria que é uma CTM com axiomas adicionais, mas que não se constitui em um topos. Ainda é matéria de investigação, se esses axiomas adicionais podem ser obtidos em categorias com morfismos verdade.

Definição 3 *Uma categoria $\mathcal{E}$ é denominada um **proto-topos** se as seguintes condições forem satisfeitas*[10]*:*

PT1. $\mathcal{E}$ é uma CTM;

PT2. Em $\mathcal{E}$, pullbacks preservam epimorfismos;

PT3. Em $\mathcal{E}$, coprodutos preservam pullbacks.

Usando (PT2), pode-se demonstrar o seguinte resultado[11]:

Em qualquer proto-topos $\mathcal{E}$, se o diagrama

$$\begin{tikzcd} a \arrow[r, "\mathrm{f}"] \arrow[d, "u"'] & b \arrow[d, "v"] \\ c \arrow[r, "\mathrm{g}"'] & d \end{tikzcd}$$

é um quadrado pullback, então existe um morfismo $h : f(a) \to g(c)$

[10]Os axiomas adicionais correspondem aos Fatos 1 e 2 da seção *Topos*.

[11]Ver Goldblatt 2006, páginas 149-150.

que torna o quadrado da direita do diagrama

$$\begin{tikzcd} a \arrow[r, "f^*"] \arrow[d, "u"'] & f(a) \arrow[r, "im(f)"] \arrow[d, "h"] & b \arrow[d, "v"] \\ c \arrow[r, "g^*"'] & g(c) \arrow[r, "\mathrm{im(g)}"'] & d \end{tikzcd}$$

um quadrado pullback.

Com o lema e o axioma (PT3), pode-se demonstrar o resultado que falta para o morfismo disjunção.

Assim, em um proto-topos $\mathcal{E}$, o comportamento de $f_\vee$ com respeito aos valores de verdade $\top$ e $\bot$ é dado pela seguinte tabela.

x	y	$f_\vee \circ \langle x, y \rangle$
$\top$	$\top$	$\top$
$\top$	$\bot$	$\top$
$\bot$	$\top$	$\top$
$\bot$	$\bot$	$\bot$

que corresponde à tabela usual da disjunção na lógica proposicional clássica.

6. Algumas questões

Concluo o texto com um levantamento de questões que têm o objetivo de conduzir o trabalho na direção de um estudo sobre os ambientes categoriais mais gerais que são capazes de fornecer semântica para linguagens proposicionais usuais, nas linhas apontadas acima.

Questão 1 *O primeiro ponto a ser investigado é se os axiomas de proto-topos já são válidos nas categorias com morfismos verdade. Se esse for o caso, é claro que as CTM's já constituiriam o ambiente categorial suficiente para a definição dos morfismos verdade com as propriedades desejadas.*

Questão 2 *Outro problema relacionado com a questão acima é se, nas CTM's, vale um teorema análogo ao teorema fundamental da teoria de topos. Se $\mathcal{E}$ é uma categoria com morfismos verdade e a é um $\mathcal{E}$-objeto, então a categoria slice $\mathcal{E} \downarrow a$ é também uma categoria com morfismos verdade. A mesma questão também para os proto-topos.*

Questão 3 *Já é um resultado bastante conhecido que as sentenças válidas em todos os topos são exatamente as teses da lógica intuicionista. Como todo topos é uma CTM ou um proto-topos, será que o resultado continua valendo ou teríamos uma lógica intermediária entre a lógica clássica e a lógica intuicionista?*

Questão 4 *Encontrar exemplos de CTM's ou de proto-topos que não são topos e também não são bivalentes. A categoria $SET^{\leq\aleph_0}$ dos conjuntos no máximo enumeráveis é um exemplo de CTM que não é um topos mas que é bivalente, seu classificador de subobjetos é o classificador usual de SET que possui dois elementos.*

ISSN: 2317-9570

Questão 5 *Desenvolver em categorias com morfismos verdade ou proto-topos o conceito de álgebra de subojetos.*

Referências

ARBIB, M. A. e MANES, E. G. (1975). *Arrows, structures, and functors: The categorical imperative*. Academic Press.

ARTIN, M., GROTHENDIECK, A. e VERDIER, J. L. (1972). Théorie des topos et cohomologie étale des schémas (SGA 4). *Lecture Notes in Mathematics*, vol 269.

AWODEY, S. (2010). *Category theory*. Oxford University Press.

BELL, J. L. (2008). *Toposes and local set theories: an introduction*. Courier Corporation.

BOILEAU, A. e JOYAL, A. (1981). La logique des topos. *The Journal of Symbolic Logic*, vol 46(1), pp 6-16.

FREYD, P. (1972). Aspects of topoi. *Bulletin of the Australian Mathematical Society*, vol 7, pp 1-76.

GOLDBLATT, R. (2006). *Topoi: the categorical analysis of logic*. Dover Publications, Inc.

HEYTING, A. (1966). *Intuitionism*. 2nd revised edition. North-Holland.

HERRLICH, H. e STRECKER, G. E. (1973). *Category theory*. Allyn and Bacon Inc.

JOHNSTONE, P. T. (1977). *Topos theory*. Academic Press.

JOHNSTONE, P. T. (2002). *Sketches of an elephant: A topos theory compendium*, 2 vols. Oxford University Press.

KOCK, A. e WRAITH, G. C. (1971). *Elementary toposes*. Aarhaus Lecture Note Series, 30.

LAWVERE, F. W. (1964). An elementary theory of the category of sets. *Proceedings of the national academy of sciences*, vol 52(6), pp 1506-1511.

LAWVERE, F. W. (1970). Quantifiers and sheaves. *Actes du congres international des mathematiciens, Nice*, vol 1, pp 329-334.

MACLANE, S. (1971). *Categories for the working mathematician*. Springer-Verlag.

MACLANE, S. e MOERDIJK, I. (2012). *Sheaves in geometry and logic: A first introduction to topos theory*. Springer-Verlag.

MACLARTY, C. (1995). *Elementary categories, elementary toposes*. Clarendon Press.

PARÉ, R. (1974) Co-limits in topoi. *Bulletin of the American Mathematical Society*, vol 80(3), pp 556-561.

TIERNEY, M. (1972). Sheaf theory and the continuum hypothesis. *Toposes, algebraic geometry and logic*. Lecture Notes in Mathematics. Springer-Verlag. Pp 13-42.

ISSN: 2317-9570

Sobre as Limitações do Dilema do Bonde para a Avaliação dos Riscos Impostos por Veículos Autônomos

[On the Limitations of the Trolley Problem for Evaluating the Risks Imposed by Autonomous Vehicles]

Renato Rodrigues Kinouchi*

Resumo: Neste trabalho discuto a adequação da versão original do dilema do bonde no que concerne a questão dos potenciais acidentes com veículos autônomos. A principal crítica se refere à impossibilidade de incluir questões relativas aos riscos e incertezas envolvidos em tais acidentes. A seguir, apresento um modelo abstrato que satisfaz os requisitos conceituais da noção padrão de risco, o qual assumirá a forma de um dilema do bonde de natureza probabilística. Mediante a análise de dois exemplos, procuro mostrar que o novo modelo é mais adequado para se avaliar a questão dos riscos impostos por veículos autônomos. Embora não pretenda dar uma solução para o problema colocado, procuro esclarecer como tais veículos pode impor riscos distintos a pedestres e passageiros, uma questão ética não contemplada na versão original do dilema.
Palavras-chave: dilema do Bonde, probabilidade, risco, valores, veículos autônomos.

Abstract: In this paper I discuss the adequacy of the trolley problem original version for the issue of autonomous vehicles accidents. The main criticism concerns the impossibility of including aspects related to risks and uncertainties involved in such accidents. Next, I present an abstract model which satisfies the conceptual requirements of the standard notion of risk and which will take the form of a probabilistic trolley problem. By way of two examples, I defend that this new model is more adequate to evaluate the risks imposed by autonomous vehicles. Although not intending to solve the problem, I try to clarify how such vehicles may impose distinct risks to pedestrians and passengers, an ethical question that the original version of the trolley problem has overlooked.
Keywords: autonomous vehicles, probability, risk, trolley problem, values.

Introdução

Formulado por Philippa Foot e mais tarde modificado por Judith Thomson, o *dilema do bonde* admite várias versões mas a "estrutura básica de todos os dilemas é a mesma: se voce nao agir, cinco pessoas irão morrer; se voce agir, uma pessoa será morta mas

*Professor Adjunto da Universidade Federal do ABC. E-mail: renato.kinouchi@ufabc.edu.br.

ISSN: 2317-9570

DOI: https://doi.org/10.26512/rfmc.v6i2.22105

as cinco serão salvas"(BRUERS e BRAECKMAN, 2014, p. 251)[1]. Originalmente, o dilema do bonde era um dos vários argumentos que Foot (1967) dirigiu à doutrina do *duplo efeito*, usualmente invocada por pensadores católicos para dar sustentação as suas visões sobre o aborto [2]. O duplo efeito em questão refere-se a uma distinção entre os resultados intencionalmente visados por uma ação e os resultados colaterais previstos mas não intencionalmente visados. No que tange a discussão do aborto naquela época, os adeptos de tal noção afirmavam que a cirurgia de histerectomia tinha como efeito previsto, mas não intencional, a morte de eventuais embriões; no entanto, outras intervenções cirúrgicas que levassem ao óbito da criança eram consideradas intencionais, de tal maneira que, nesses casos, os médicos estariam atentando contra vidas inocentes. Convém enfatizar que o dilema do bonde era um dos argumentos mobilizados por Foot em sua discussão da doutrina do duplo efeito, mas, com efeito, havia outros argumentos adicionais, dentre eles a possibilidade de um cirurgião retalhar uma pessoa saudável tendo em vista transplantar seus órgaos para outros cinco pacientes (cf. FOOT, 1967, p. 11). No geral, Foot chega a conclusão de que o *dever negativo* de não fazer o mal prevalece sobre o *dever positivo* de oferecer auxílio; de tal maneira que, para o caso do cirurgião, seria reprovável promover o bem de cinco pessoas à custa da morte de uma outra; mas no caso do bonde, a decisão diz respeito a evitar dois males, e assim desviar a trajetória do bonde seria justificável por minimizar as mortes. Para Foot, a distinção entre deveres positivos e negativos precisava ser levada em conta na discussão do aborto, todavia a autora não se posiciona definitivamente sobre a questão propriamente dita: " Eu não estou argumentando em prol ou contra tais pontos de vista, mas somente tentando discernir algumas das correntes de pensamento que nos movem para frente e para trás. A leviandade dos exemplos nao tem por intenção ofender"(FOOT, 1967, p.18).

[1] Todas as traduções de citações foram realizadas pelo autor deste ensaio.

[2] Segundo a autora: "A doutrina do duplo efeito é baseada na distinção entre aquilo que alguém pode prever como um resultado de sua ação voluntária e aquilo que, em sentido estrito, ele intencionalmente visa (...) As palavras "duplo efeito" se referem aos dois efeitos que uma ação pode produzir: aquilo que é visado e aquilo que é previsto mas de modo algum desejado (...) Diz-se, por exemplo, que a operação de histerectomia envolve a morte de fetos como uma consequência prevista mas não diretamente visada do ato cirúrgico, enquanto outras operações matam a criança e têm por propósito direto eliminar uma vida inocente, uma distinção que tem evocado reações particularmente amargas por parte de pessoas que não são católicas" (Foot, 1967, p. 5-6).

ISSN: 2317-9570

Anos mais tarde, Judith Thomson retomou a discussão iniciada por Foot mas deu novos contornos ao dilema. Para além da configuração original, Thomson (1976, 1985) propôs outros dois casos, a saber: a versão do bonde com *loop* e a versão do *homem gordo* sobre a ponte. Na versão do bonde com *loop*, o trilho lateral onde há uma pessoa isolada se liga novamente ao ramal principal, de tal maneira que o bonde retorna por detrás das cinco pessoas. O resultado disso é a possibilidade de se afirmar que a morte da pessoa isolada, devido ao desvio de rota, poderia ser entendida como um *meio* para se evitar a morte das outras cinco pessoas. Já na versão do *homem gordo*, tal pessoa, cujo tamanho seria suficiente para obstruir o bonde, é empurrada de cima de uma ponte também como um *meio* de evitar a morte das cinco pessoas. Na realidade, deve-se a Thomson as linhas gerais das discussões subsequentes sobre o dilema do bonde (cf. OTSUKA, 2008) e desde a publicação de seus trabalhos uma quantidade enorme de diferentes versões do dilema foram propostos. O principal ponto do debate envolve uma bem documentada inconsistência nas respostas das pessoas entrevistadas quando se comparam a versão padrão do dilema e a versão do homem gordo. Na primeira versão, a maioria dos entrevistados julga permissível a ação de acionar a alavanca que desvia o bonde resultando na morte da pessoa no trilho lateral ao invés das cinco pessoas no trilho principal; entretanto, apenas uma minoria julga permissível a ação de empurrar um homem gordo que se encontra sobre uma ponte acima dos trilhos com o intuito de bloquear a passagem do bonde (HAUSER et al. 2008). Em um trabalho de revisão, Bruers e Braeckman examinaram 35 anos de literatura, classificaram dezessete diferentes grandes versões do problema e mostraram que "a maioria das intuições morais das pessoas não seguem as éticas consequencialistas", um resultado que torna "o dilema do bonde um experimento mental interessante para o estudo das éticas deontológicas" (BRUERS e BRAECKMAN, 2014, p. 252).

Recentemente, o dilema do bonde passou a ser debatido em um contexto muito diferente. Com o desenvolvimento de veículos autônomos, não demorou muito tempo para se estabelecer uma analogia com o experimento moral clássico. Com efeito, suponhamos que um veículo autônomo precise *escolher* entre as opções de atropelar cinco pessoas ou desviar resultando na morte do passageiro: qual deve ser a decisão a ser tomada? Sem pretender oferecer uma resolução do problema, este ensaio visa analisar a adequação dessa analogia com o dilema do bonde. Na próxima seção, procuro mostrar que o dilema original é incapaz de apreender uma das características essenciais da problemática dos carros autônomos, a saber, seu caráter probabilístico. Isso faz com que a maior parte da literatura produzida sobre o dilema do bonde não seja pertinente ao que realmente se procura determinar na questão dos acidentes com veículos autônomos. Não obstante, acredito que se o dilema do bonde original for transformado em um modelo de caráter probabilístico, então uma pertinente nova questão filosófica aparece, a saber: o que irá definir a distribuição dos riscos de acidentes com veículos autônomos? Essa é a pergunta que tenho a intenção de discutir ao longo deste trabalho.

A analogia entre o dilema do bonde e os veículos autônomos

Em 2015, a prestigiosa revista *Nature* noticiou o desenvolvimento de algoritmos computacionais para o gerenciamento de colisões em situações aparentemente semelhantes ao do dilema do bonde (DENG, 2015). Nesse novo contexto, o dilema original sofre algumas adaptações: o bonde é substituído por um automóvel, a decisão passa a ser feita pelo algoritmo de controle do veículo e as opções oferecidas são entre atropelar cinco pedestres ou desviar em uma manobra fatal para o passageiro. Em 2016, a não menos prestigiosa *Science* divulgou um outro estudo bastante interessante (BONNEFON et al, 2016: cf. SHARIFF et al, 2017), entretanto pouco surpreendente, segundo o qual muito embora os participantes da pesquisa apoiem a tese de que os veículos autônomos devem evitar atropelamentos fatais a pedestres, esses mesmos participantes prefeririam não se locomover em veículos governados por essa diretriz, os chamados *veículos utilitaristas*, preferindo utilizar *veículos autoprotetores*, que não salvariam os pedestres às custas da morte do passageiro:

> Embora as pessoas tendam a concordar que seria melhor para todos que os veículos autônomos fossem utilitaristas (no sentido de minimizar o número de vítimas no trânsito), essas mesmas pessoas têm um incentivo pessoal de se locomoverem em veículos que as protegerão a qualquer custo. Assim, se fossem colocados no mer-

ISSN: 2317-9570

> cado ambos os tipos de veículos, poucas pessoas desejariam utilizar veículos utilitaristas, apesar de preferirem que os outros assim o fizessem (BONNEFON et al, 2016, p. 1575).

Geralmente se reconhece que o dilema do bonde original é um experimento mental muito pouco realista, e por isso o debate também inclui boa dose de ceticismo sobre a adequação de modelos de tipo dilema do bonde para o caso de veículos autônomos. De um ponto de vista jurídico, Casey (2017) afirma que as abstrações filosóficas presentes no dilema do bonde pouco contribuem para os problemas enfrentados pelos engenheiros, sendo mais importante o estudo e o desenvolvimento do aparato legal que irá balizar o uso dessas tecnologias. De um ponto de vista mais filosófico, Nyholm & Smids (2016) assinalam que a analogia carece de sustentação em pelo menos três frentes, a saber: "com respeito às características gerais da situação de decisão, com respeito ao papel da responsabilidade moral e legal, e com respeito à situação epistêmica dos tomadores de decisão" (NYHOLM e SMIDS, 2016, p. 1287). Uma linha de argumentação influente tem sido defendida por Goodall (2014, 2016), segundo a qual a problemática dos veículos autônomos precisa levar em consideração o conhecimento já fartamente produzido na área de análise de risco. Segundo essa abordagem, embora seja razoável supor que a frequência de casos nos quais veículos autônomos terão que enfrentar *escolhas* difíceis – tais como entre matar cinco pedestres ou o ocupante do veículo – será extremamente baixa, por outro lado é praticamente certo que haverá acidentes dos mais variados tipos e com as mais diversas expectativas de mortalidade. Por sinal, as taxas de mortalidade dependem, entre outras coisas, de se o condutor consumiu álcool ou não (duas vezes maiores para condutores alcoolizados), se é homem ou mulher (28% a mais se for mulher), jovem ou idoso (acidentes são quase três vezes mais letais para pessoas acima dos 70 anos em comparação com jovens de 20 anos; cf. EVANS, 2008). Embora os especialistas consigam predizer o resultado dos acidentes mais catastróficos (KOCKELMAN e KWEN, 2002; O´DONNELL e CONNOR, 1996), praticamente tudo o mais que envolve acidentes de trânsito é de natureza probabilística. Ocorre que o dilema do bonde original não serve para tratar de tal questão, dado que o resultado das ações de desviar ou não desviar a trajetória resultam em consequências letais dadas como certas. A seguir, apresento uma passagem relativamente longa de Nyholm e Smids (2016) que sumariza esse ponto:

> Tão logo começamos a considerar esses vários detalhes adicionais, torna-se claro que aquilo com que estamos lidando aqui não são resultados cujas características são conhecidas com certeza. Na realidade, estamos lidando com situações repletas de incertezas e com numerosas avaliações de risco mais ou menos confiáveis (Cf. GOODALL, 2014, p. 96). Isso significa que precisamos abordar a ética dos carros autônomos usando um tipo de raciocínio moral que não temos razão ou ocasião de usar quando pensamos sobre os casos usuais discutidos na literatura do dilema do bonde. No primeiro caso, precisamos nos engajar em um raciocínio moral sobre riscos e gerenciamento de riscos. Precisamos também de um raciocínio a respeito de decisões sob incerteza. Em contraste, o raciocínio moral empregado por quem lida com um dilema do bonde não é sobre riscos e sobre como responder a diferentes riscos. Nem sobre como realizar decisões que envolvem incerteza. Essa é a diferença categórica entre a ética do dilema do bonde e a ética dos algoritmos e dos acidentes com carros autônomos. Raciocinar sobre riscos e incerteza é categoricamente diferente de raciocinar sobre fatos conhecidos e resultados certos. Os conceitos-chave usados diferem drasticamente no que diz respeito às inferências que eles garantem. E aquilo que selecionamos usando tais conceitos são coisas pertencentes a categorias metafísicas e condições modais diferentes (e.g., risco de dano, de um lado, e dano real, de outro). Portanto, as questões éticas difíceis e específicas que riscos e incertezas engendram não estão em jogo nos dilemas do bonde. Todavia, elas certamente aparecem na ética dos carros autônomos. Permitam-me dar um exemplo. Um número significativo de pessoas pode julgar moralmente inaceitável atropelar um

ISSN: 2317-9570

> pedestre se isso necessariamente matá-lo (Cf. THOMSON, 2008). Não obstante, e se a probabilidade estimada de colisão fatal for de 10%? Ou de apenas 1%? Para muitas pessoas, impor uma probabilidade de morte de 1% a um pedestre inocente tendo em vista salvar cinco ocupantes de um veículo pode parecer uma escolha correta. Os dilemas do bonde não exigem tais juízos. Nos cenários envolvidos nos dilemas do bonde, todos os resultados são assumidos como 100% certos, e desse modo não se precisa refletir sobre a maneira de se confrontar diferentes resultados incertos e/ou arriscados uns contra os outros (NYHOLM e SMIDS, 2016, p. 1286).

Como visto, há razões para considerar o dilema do bonde original como sendo insuficiente para lidar com a problemática dos veículos autônomos. É necessária uma abordagem que seja capaz de estimar os riscos impostos pela utilização de tal tecnologia. Com efeito, abordagem baseada no conceito de risco envolve aspectos valorativos para os quais análise filosófica pode contribuir significativamente, para além dos estreitos limites impostos pelo dilema original (cf. LIN, 2015). A seguir discuto uma nova versão que, acredito, pode lançar alguma luz adicional sobre a questão dos veículos autônomos. Daqui em diante essa nova versão do dilema do bonde original será chamada de *modelo do bonde probabilístico*.

O modelo do bonde probabilístico

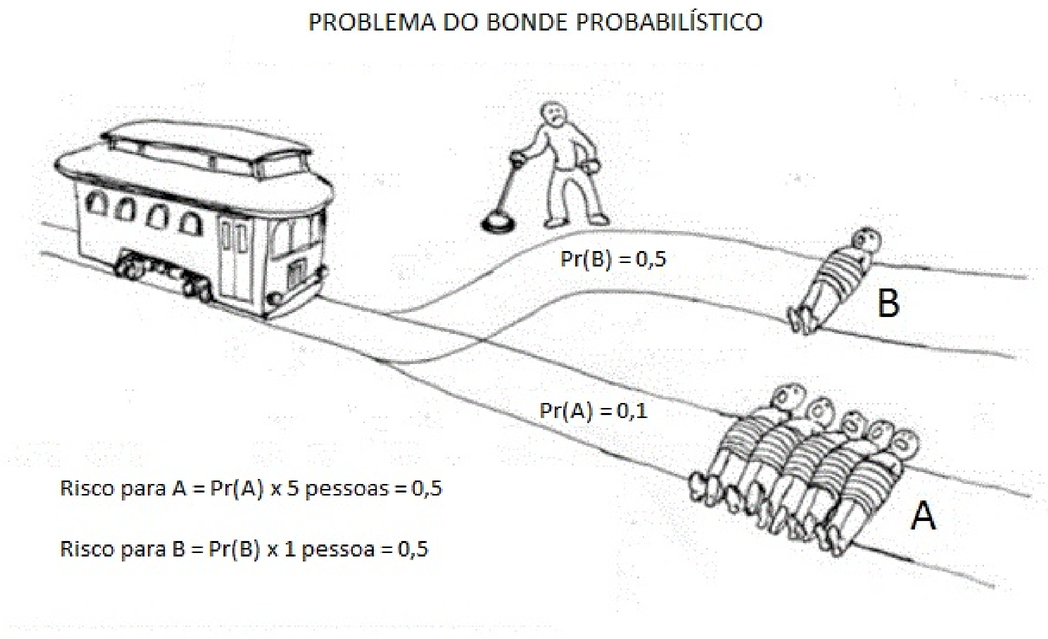
Faz-se necessária uma adaptação no dilema do bonde original para transformá-lo em um modelo mais apropriado às situações que realmente serão enfrentadas por veículos autônomos; a saber, é necessário estipular uma probabilidade entre ato de acionar a alavanca e a consequência indesejável do atropelamento das pessoas atadas aos trilhos. Em outras palavras, acionar a alavanca não resultará inescapavelmente em mortes, mas ensejará uma probabilidade para as mortes. Assim, o dilema do bonde original se transforma em um modelo que inclui a noção de risco se considerarmos a relação entre o acionamento da alavanca e a consequência indesejável como sendo de natureza probabilística ao invés de determinista. Para o filósofo sueco Sven Hansson, renomado filósofo da tecnologia:

> A exclusão da questão de exposição a riscos nas considerações da maior parte da teoria moral pode ser claramente vista nas suposições deterministas comumente feitas nos exemplos usuais de tipo vida-ou-morte empregados para explorar as implicações das teorias morais. No famoso dilema do bonde, você assume saber que se acionar a alavanca então uma pessoa será morta, enquanto que se você não acionar, então cinco pessoas serão mortas (...). Isso claramente contrasta com os dilemas da vida real, onde os problemas das nossas ações que envolvem vidas humanas raramente são acompanhadas por um conhecimento certo sobre as consequências de cursos de ação alternativos (HANSSON, 2012, p. 44).

Segundo a noção padrão usada em estatística, um *risco* é o *valor esperado* de um evento indesejado, o qual pode ser calculado multiplicando-se a probabilidade do evento – expressa por um número dentro do intervalo [0, 1] – por uma estimativa de sua severidade (HANSSON, 2013)[3] . Quando aplicado ao dilema do bonde, o valor esperado do risco de atropelamento consiste na probabilidade de óbito vezes o número de pessoas em cada trilho. Na verdade, a versão proposta é uma generalização da versão original, isto é, o problema do bonde original torna-se o caso especial em que a probabilidade de morte é igual a 1. No dilema original, acionar ou não a alavanca implica que o evento indesejado irá ocorrer inescapavelmente para uma das alternativas. Mas ao trans-

[3]É importante notar que, de acordo com tal definição, a magnitude do risco varia em função tanto da probabilidade do evento quanto da quantidade de dano que o evento pode provocar. Por exemplo, supondo-se que as probabilidades de terremotos no Alasca e na Califórnia fossem iguais, o risco no segundo caso seria muito maior em virtude da grande diferença em termos de perdas materiais e humanas que tal um evento geológico poderia provocar. Cumpre assinalar que o conceito de risco é valorativo (*value-laden*) até mesmo para essa definição técnica, dado que um dos termos da equação deve conter algum tipo de estimativa de um valor (vidas humanas, bens materiais, etc.) sob risco. A respeito desse ponto, Hanson (2004, 2005, 2009, 2012, 2013) costuma enfatizar que "o risco sempre se refere à possibilidade de algo ruim acontecer. Em virtude de ser indesejável, risco envolve valores" (HANSSON, 2013, p. 10). Todavia, ser valorativo não significa deixar de ser factual; em outras palavras, o conceito de risco não se resume exclusivamente a juízos de valor (cf. MÖLLER, 2012). Na verdade, as análises de risco podem ser vistas como socialmente construídas no sentido trivial de que qualquer atividade humana depende de cooperação social, convenções linguísticas, etc., mas, embora seja verdade que as pessoas exibem diferentes percepções de risco acerca de, por exemplo, terremotos, parece estar além de qualquer dúvida razoável que eventos sísmicos de grande magnitude são muito mais frequentes nas áreas de contato entre as placas tectônicas, o que justifica investimentos em projetos de edifícios à prova de terremotos, sistemas de alarmes e outros dispositivos de segurança socialmente adotados em cidades nas proximidades de falhas geológicas.

formar aquela certeza de morte em risco de morte, o valor esperado desse risco pode ser representado pela multiplicação da probabilidade do evento indesejado – e ser atropelado certamente pode ser considerado desagradável – pela estimativa da severidade do acontecimento. Para ilustrar a nova situação, consideremos o enunciado a seguir: "Um bonde trafega sem freios no trilho A e irá atropelar cinco pessoas com probabilidade 10% de que todas elas morram; um transeunte encontra-se ao lado de uma alavanca cujo acionamento desviará o bonde para um trilho B, mas nesse caso o bonde irá atropelar uma pessoa com probabilidade de 50% de que ela morra". A figura abaixo pode facilitar a compreensão do modelo:



Nesta nova versão do dilema, pode-se variar as probabilidades de cada opção, sendo assim também possível criar situações de indiferença de risco onde as opções resultam em um mesmo valor esperado. Para os valores do enunciado acima, os riscos para os ramais resultam em 0,1 x 5 = 0,5 para o trilho A e 0,5 x 1 = 0,5 para o trilho B; ou seja, no tocante ao valor esperado, as opções A e B se equivalem. Em termos técnicos, as opções são indiferentes. O cálculo utilitário, nesse caso, pouco ajuda, não porque falha em fornecer uma estimativa numérica precisa do risco, mas porque o risco é o mesmo para ambas opções. Esse cenário, com efeito, assemelha-se a um caso comum de escolha de imposição de riscos: por um lado, ao acionar a manivela (opção B) concentramos o risco na pessoa isolada, e, por extrapolação, essa escolha penalizaria minorias em geral; por outro lado, quanta exposição a risco uma maioria deve aceitar (opção A) para proteger indivíduos desafortunados e minorias?

Como já discutido, a analogia entre veículos autônomos e o dilema do bonde original costuma ser considerada insuficiente e pouco proveitosa, todavia o modelo do bonde probabilístico consegue incluir uma questão com implicações filosóficas importantes, a saber, a da distribuição dos riscos impostos pelos veículos autônomos. Cumpre assinalar que embora o bonde probabilístico não seja um modelo estritamente *realista* acerca dos possíveis acidentes, ainda assim inclui, mesmo que de maneira abstrata, um componente de incerteza completamente ausente no dilema original.

Imposição de riscos por veículos autônomos: dois exemplos

A situação descrita no dilema original é certamente pouco verossímil, e quando transposta para a questão dos veículos autônomos costuma tomar a forma de um modelo também pouco verossímil, a saber: um carro autônomo com um passageiro trafega por uma avenida; subitamente cinco pedestres cruzam à frente do veículo; o atropelamento resultará na morte dos cinco pedestres as pode ser evitado se o algoritmo que controla o carro desviar o veículo, que nesse caso colidirá resultando na morte do passageiro. Assim colocada, a analogia entre o dilema do bonde original e os veículos autônomos não costuma receber muito crédito, pois o dilema original acaba restringindo a discussão a situações cujas consequências são dadas como certas.

O problema dos potenciais acidentes com veículos autônomos requer um modelo probabilístico para recobrir as mais variadas situações que implicam em imposição de riscos. Todavia, em se tratando de riscos, também é necessário levar em consideração a severidade deles, isto é, a quanti-

[4]Na realidade, dever-se-ia também considerar aspectos qualitativos – tais como gênero, idade, estado de saúde geral. etc. – o que torna qualquer modelo quantitativo apenas um retrato parcial das situações de risco. No entanto, para os propósitos presentes, confinamos discussão ao aspecto quantitativo.

ISSN: 2317-9570

dade de pessoas afetadas[4]. Com efeito, faz diferença se há apenas um ocupante do veículo e cinco pedestres, ou o caso inverso de cinco ocupantes do veículo e apenas um pedestre. Para ilustrar os riscos impostos por essas variadas condições, consideremos os dois exemplos a seguir.

Exemplo 1: Cinco pedestres subitamente se colocam à frente de veículo autônomo transportando um passageiro. Suponhamos que devido à proteção oferecida pelo veículo ao seu ocupante, o atropelamento não irá causar dano físico ao passageiro. Suponhamos ainda que o acidente não necessariamente causa a morte dos cinco pedestres, mas sim engendra a probabilidade de 50% de que todos eles morram. Nesse caso, o valor esperado do risco consiste na probabilidade de morte (0,5) multiplicada pela severidade do acidente (5 vidas), totalizando 2,5. Por outro lado, se o veículo desviar, no máximo haverá apenas uma morte (a do passageiro), de modo que o valor esperado do risco não pode ultrapassar 1,0; portanto bem menor que em caso de atropelamento.

Se o veículo em questão for autoprotetor, ele nunca se desvia visto que o atropelamento dos pedestres não resulta em ameaça à integridade física do passageiro; em razão disso, o risco de eventuais mortes recai sempre sobre os pedestres. Por outro lado, um veículo utilitarista procuraria minimizar os óbitos, e nas circunstâncias acima descritas sempre desviaria, pois o valor esperado do risco da manobra de desviar (1,0) nunca supera o valor esperado do risco de atropelamento (2,5). Com efeito, é possível que muitas manobras de desviar não resultem na morte do passageiro, mas todas as possíveis mortes virão dessa população. Em resumo, agora o risco recai sobre os ocupantes desses veículos.

Não obstante, pode-se colocar a questão adicional: considerando-se a situação do passageiro, não seria razoável manter a trajetória e "torcer" para que os pedestres saiam vivos do acidente (o que, afinal de contas, pode acontecer com probabilidade de 50%)? Com efeito, há pesquisas que propõem veículos "rawlsianos" (LEBEN, 2017) que optam por desviar exceto em caso de maior ameaça ao passageiro. Um veículo desse tipo não seria exatamente um carro autoprotetor, o qual sempre protege seu ocupante; um carro rawlsiano calcularia os riscos envolvidos para as opções e procuraria evitar uma manobra que causasse a morte do passageiro, mas nos casos em que a manobra significa um risco aceitável ao passageiro, tal veículo desviaria. Todavia, assoma-se uma questão adicional: qual o nível mínimo de risco aceitável que faria um carro rawlsiano desviar? Ao tratar desse problema, Leben (2017) sugere a adoção de um critério de *maximin* (cf. LEBEN, 2017, p. 108-112) segundo o qual, independentemente do valor esperado total do risco aos pedestres (principal critério usado por veículos utilitaristas), um veículo rawlsiano mantém seu curso e protege seu passageiro a partir do ponto em que a probabilidade de morte para o passageiro, em virtude da manobra, for maior que a probabilidade de morte dos pedestres. No entender de Leben: "a diferença conceitual chave é que o carro rawlsiano evita uma alternativa ruim para o sujeito com pior ganho, mesmo se isso resultar em maiores oportunidades para todos os outros" (LEBEN, 2017, p. 112). Todavia, a alternativa proposta por Leben também acarreta problemas dignos de menção. Ao procurar evitar o pior desfecho, um veículo rawlsiano tende a "mirar" veículos/motoristas mais seguros: por exemplo, entre atropelar um motociclista sem capacete (mais inseguro) ou um motociclista com capacete (mais seguro), o carro sempre escolheria a segunda opção, impondo riscos exatamente a quem preza pela segurança.

Exemplo 2: Apenas um pedestre subitamente se coloca a frente de veículo autônomo transportando cinco passageiros; isto é, agora há mais passageiros do que pedestres ameaçados. Suponhamos que devido à proteção oferecida pelo veículo aos seus ocupantes, o atropelamento não irá causar dano físico aos passageiros. Suponhamos ainda que o acidente não irá necessariamente causar a morte do único pedestre, mas sim engendra a probabilidade de 50% de que ele morra.

Nesse caso, o valor esperado do risco consiste na probabilidade de morte do pedestre (0,5) multiplicada pela severidade do acidente (1 vida), totalizando 0,5. Por outro lado, se o veículo desviar, ele pode colocar em risco os cinco passageiros; cabe então indagar: qual probabilidade aceitável para manobra de desviar? Tal como colocado no final da seção anterior deste trabalho: quanto risco uma maioria deve aceitar em prol de uma minoria? O cálculo utilitário leva ao seguinte resultado: quando a probabilidade de morte ao desviar for de 10%, o valor esperado do risco aos passageiros será equivalente ao do pedestre (0,1 x 5 = 0,5). No caso de um veículo utilitarista, o carro deve desviar para valores abaixo de 10% de probabilidade de fatalidade da manobra

ISSN: 2317-9570

para os passageiros; mas para valores acima disso, o veículo deve manter a trajetória. Não obstante, no caso de um veículo rawlsiano, tal como o defendido por Leben (2017), o veículo mantém a trajetória se a probabilidade de morte dos passageiros em virtude da manobra for maior do que a do pedestre (que é de 50%); mas para qualquer valor abaixo de 50%, o veículo desvia, diferindo tanto de veículos utilitaristas – que desviam somente se a fatalidade para os passageiros for menor que 10% – quanto de veículos autoprotetores – que invariavelmente não desviam pois o atropelamento não implica em ameaça à integridade dos passageiros.

As relações de imposição de risco acima descritas, com os valores de probabilidade e número de pessoas utilizados nos dois exemplos, podem ser resumidas da seguinte maneira. Para o exemplo 1, onde há cinco pedestres e um passageiro, veículos utilitaristas sempre desviam e o risco é imposto ao passageiro; já veículos autoprotetores nunca desviam, de modo que o risco é imposto aos pedestres; finalmente, veículos rawlsianos impõe riscos à alternativa cuja probabilidade de mortalidade for menor, o que pode ser interpretado como uma tendência do veículo "mirar" a alternativa mais segura. Para o exemplo 2, onde há apenas um pedestre e cinco passageiros, um veículo autoprotetor novamente impõe risco somente aos pedestres; um veículo utilitarista, por sua vez, desvia se a probabilidade de fatalidade da manobra for menor que 10%, tendendo a impor risco ao pedestre acima desse patamar (entretanto, vale notar que veículos utilitaristas podem ser uma boa alternativa se forem usados prioritariamente para transporte coletivo, tais como ônibus e trens); finalmente, um veículo rawlsiano desvia se probabilidade de fatalidade da manobra para os passageiros for menor que para o pedestre (que é de 50%), e assim impõe risco aos passageiros se a probabilidade de fatalidade da manobra estiver na faixa acima de 10% mas abaixo de 50%.

Tais relações de imposição de risco não podem ser capturadas pela analogia entre veículos autônomos e o dilema do bonde original, pois em tais cenários não há "qualquer incerteza sobre os resultados das decisões, mas uma discussão abrangente sobre algoritmos morais precisará englobar os conceitos de risco, valor esperado e atribuição de culpa" (BONNEFON et al, 2016, p. 1576). Para que tais aspectos sejam levados em consideração é necessário, no mínimo, utilizar alguma adaptação na direção apontada pelo modelo do bonde probabilístico.

Cumpre assinalar que, na verdade, este não se trata de um modelo mais realista, pois ainda recorre a abstrações que não levam em conta os detalhes específicos relativos ao real comportamento de veículos autônomos. Não obstante, por ser de natureza probabilística, ele amplia o escopo da análise e abarca os problemas de imposição de riscos que eram até então negligenciados.

Conclusão

Entusiastas dos veículos autônomos acreditam que sua utilização, em um futuro não muito distante, tornará o trânsito muito mais seguro. Ocorre que tal tecnologia não será perfeitamente segura, em virtude dos inúmeros fatores que contribuem para a ocorrência de acidentes. O dilema do bonde, em sua versão original, coloca em pauta situações extremas, onde qualquer das alternativas disponíveis acarretarão em vítimas fatais. Mas ao focar apenas nessas situações extremas, perde-se de vista que a grande maioria dos futuros acidentes com veículos autônomos, inclusive os acidentes fatais, irão ensejar uma probabilidade de dano à integridade física de passageiros e/ou pedestres. Como visto nos exemplos apresentados, veículos autônomos autoprotetores impõem riscos a pedestres e outros veículos; por sua vez, veículos autônomos utilitaristas podem ser uma boa alternativa para o transporte público, mas para transporte individual tenderiam impor riscos ao passageiro (o que do ponto de vista mercadológico não é nada atraente); finalmente, veículos autônomos rawlsianos (isto é, aqueles que evitam o pior desfecho por meio de algum tipo de cálculo maximin) tendem a "mirar" pedestres ou outros veículos considerados mais seguros – por exemplo, entre colidir com um motociclista usando capacete e outro sem capacete, tal veículo escolheria a primeira opção dado que a probabilidade de morte seria ligeiramente menor.

Tal como alertado no início deste artigo, aqui não se pretendeu resolver esses vários problemas éticos ensejados pela adoção dos veículos autônomos. Mas isso não quer dizer que a filosofia e a ética sejam irrelevantes para o debate, tal como às vezes alguns tecnólogos parecem pensar. Significa dizer, na realidade, que é necessário compreender o novo fenômeno a partir dos conceitos de risco e de incerteza, de modo a conceber modelos mais eficazes para essa tarefa. A proposta do bonde probabilístico é um passo nessa direção. O resultado mais significativo, em meu entender, consiste em se reconhecer que veí-

ISSN: 2317-9570

culos autônomos irão impor riscos de maneira distinta a pedestres e passageiros: uma questão ética para o qual famoso dilema do bonde original é insuficiente.

Agradecimentos: Este trabalho foi escrito durante visita *ao Centre for Philosophy of Natural and Social Sciences* (LSE) com financiamento da FAPESP, bolsa 2017/17081-4. O autor agradece a Roman Frigg por supervisionar a visita e fazer valiosas sugestões à pesquisa.

Referências

BONNEFON, J-F.; SHARIFF, A.; RAHWAN, I. The social dilemma of autonomous vehicles. *Science*, 352 (6293), p. 1573-1576, 2016.

BRUERS, S.; BRAEKCMAN, J. A review and systematization of the Trolley Problem. *Philosophia*, 42, p. 251-269, 2014.

CASEY, B. Amoral machines, or: How roboticists can learn to stop worrying and love the law. *Northwestern University Law Review*, 111, p. 1347, 2017.

DENG, B. The robot's dilemma. *Nature*, 523 (7558), p. 24-25, 2015.

EVANS, L. Death in traffic: Why are the ethical issues ignored? *Studies in Ethics, Law, and Technology*, 2 (1), 2008. Disponível em: https://doi.org/10.2202/1941-6008.1014. Acessado em 30 de agosto de 2018.

FOOT, P. The problem of abortion and the doctrine of the double effect. *Oxford Review, 5*, p. 5-15, 1967.

GOODALL, N. J. Machine ethics and automated vehicles. In: G. Meyer and S. Beiker (Eds), *Road vehicle automation.* Dordrecht: Springer, p. 93-10, 2014.

_____. Away from Trolley Problems and towards risk management. *Applied Artificial Intelligence*, 30 (8), p. 810-821, 2016.

HANSSON, S. O. Philosophical perspectives on risk. *Techné*, 8 (1), p. 10-35, 2004.

_____. The epistemology of technological risk. *Techné*, 9 (2), p. 68-80, 2005.

_____. Technology, prosperity and risk. In: J. K. B. Olsen, S. A. Pedersen and V. F. Hendricks (Eds.), *A Companion to the Philosophy of Technology*. Oxford: Wiley-Blackwell, 2009.

_____. A panorama of the philosophy of risk. In: S. Roeser, R. Hillerbrand, P. Sandin and M. Peterson (Eds.), *Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk*, p. 27-54. Dordrecht: Springer, 2012.

_____. *The ethics of risk: Ethical analysis in an uncertain world*. New York: Palgrave MacMillan, 2013.

KOCKELMAN, K. M.; KWEN, Y-J. Driver injury severity: An application of ordered probit models. *Accident Analysis & Prevention*, 34 (3), p. 313-321, 2002.

LEBEN, D. A Rawlsian algorithm for autonomous vehicles. *Ethics and Information Technology*, 19 (2), p. 107-115, 2017.

LIN, P. Why ethics matters for autonomous cars. In: J. Markus et al, *Autonomes fahren*, p. 69-85. Berlin/Heidelberg: Springer, 2015.

MÖLLER, N. The concepts of risk and safety. In: S. Roeser, R. Hillerbrand, P. Sandin and M. Peterson (Eds.), *Handbook of Risk Theory: Epistemology, Decision Theory, Ethics, and Social Implications of Risk*, p. 55-85. Dordrecht: Springer, 2012.

NYHOLM, S. & SMIDS, J. The ethics of accident-algorithms for self-driving cars: an applied trolley problem? *Ethical Theory and Moral Practice*, 19 (5), p. 1275-1289, 2016.

O´DONNELL, C. J. & CONNOR, D. H. Predicting severity of motor vehicle accident injuries using models of ordered multiple choice. *Accident Analysis & Prevention*, 28 (6), p. 739-753, 1996.

SHARIFF, A., BONNEFON, J-F., RAHWAN, I. Psychological roadblocks to the adoption of self-driving vehicles. *Nature Human Behaviour*, 1 (10), p. 694-696, 2017.

THOMSON, J. J. Killing, letting die, and the trolley problem. *The Monist, 59*, p. 204-217, 1976.

_____. The trolley problem. *The Yale Law Journal*, 94 , p. 1395-1415, 1985.

_____. Turning the trolley. *Philosophy & Public Affairs*, 36 (4), p. 359-374, 2008

Uma Crítica Popperiana ao Raciocínio Sociológico de Jean-Claude Passeron

[A Popperian Analisys of Jean-Claude Passeron's Sociological Reasoning]

Sergio Tarbes*

Resumo: No livro *O Raciocínio Sociológico: o espaço não popperiano do raciocínio natural*, o sociólogo francês Jean-Claude Passeron afirma que a utilização estrutural de linguagem natural nas ciências sociais cria o que denomina "raciocínio sociológico", um método baseado no raciocínio natural onde não seriam possíveis a "demarcação" e "falseabilidade" de Popper e os "paradigmas" e a "ciência normal" de Kuhn. O presente artigo se propõe a analisar os fundamentos empíricos da afirmação de Passeron por meio de um método inspirado nas ideias de Karl Popper, consistente, no caso, na apresentação de exemplos empíricos que a contradigam. Os resultados permitem concluir que se existem dificuldades nas ciências sociais para a realização de experimentos e para a consolidação de paradigmas, tais dificuldades não podem ser atribuídas à utilização da linguagem natural.
Palavras-chave: Passeron, raciocínio sociológico, ciências sociais, Popper, Kuhn.

Abstract: In the book *Le raisonnement sociologique: L'espace non-poppérien du raisonnement naturel*, the French sociologist Jean-Claude Passeron argues that the structural use of natural language in the social sciences creates what he calls "sociological reasoning," a method based on natural reasoning where Popper's "demarcation" and "falsifiability" and Kuhn's "paradigms" and "normal science" would not be possible. This article analyzes the empirical foundations of Passeron's assertion by means of a method inspired by the ideas of Karl Popper, consisting, in this case, in the presentation of empirical examples that contradict it. The results allow us to conclude that if there are difficulties in the social sciences to perform experiments and to consolidate paradigms, such difficulties cannot be attributed to the use of natural language.
Keywords: Passeron, sociological reasoning, social sciences, Popper, Kuhn.

Introdução

O livro *O Raciocínio Sociológico – o espaço não-popperiano do raciocínio natural*, de Jean-Claude Passeron, é, antes de tudo, um libelo em defesa do valor epistemológico das

*Mestre em Filosofia pela Universidade de Brasília. E-mail: sebrat01@gmail.com.

DOI: https://doi.org/10.26512/rfmc.v6i2.22106

ciências sociais. Mas a obra também é uma teoria sobre as possibilidades epistemológicas e metodológicas das ciências sociais, em especial as que o autor denomina "históricas" e "gerais": a sociologia, a antropologia e a própria história.

Desenvolvendo a defesa, Passeron afirma que as ciências sociais já são realizadas na única forma em que podem ser realizadas[1], que quaisquer avanços teóricos ou tecnológicos não alterarão a essência de sua abordagem atual, e que as características de que se revestem as ciências sociais já são finais, e não parte de um processo evolutivo que pode vir a alterar seus métodos e seu alcance.

Desenvolvendo a teoria, o autor expõe sua visão sobre as características essenciais das ciências sociais, em virtude das quais é possível identificar suas diferenças frente às ciências naturais e estabelecer seu método específico. Passeron afirma que as ciências sociais utilizam a linguagem natural, diferentemente das ciências naturais que utilizam extensamente linguagens artificiais[2], e que o objeto das ciências sociais insere-se sempre em um contexto histórico, delimitado no tempo e no espaço, que jamais se repete com todas suas características e variáveis em outro local ou outro momento, em tudo diferente do objeto das ciências naturais. Essas duas características – linguagem natural e contexto histórico – combinam-se para criar o que Passeron denomina de "raciocínio sociológico" ou "raciocínio natural": uma forma de explicar a realidade histórica e social que se utiliza de conceitos contextualizados, metáforas e comparações metodologicamente controladas, e vulnerabilidade empírica fundada em observações amplas e sistematizadas, mas válidas apenas em determinados contextos espaço-temporais.

Com isso Passeron afirma existir nas ciências sociais um espaço não-popperiano, ou seja, uma área do conhecimento humano onde os pontos centrais da filosofia das ci-

[1] Em relação ao aspecto que interessa ao presente artigo, essa única forma diz respeito à utilização privilegiada de linguagens naturais nas ciências sociais.

[2] Ainda que seja questionável a utilização das expressões "língua artificial" ou "linguagem artificial", em referência aos sistemas formais, ou sistemas simbólicos, como a matemática, a geometria e a lógica, elas foram mantidas por duas razões principais: primeiro, porque são as expressões utilizadas por Passeron no texto sob análise (vide PASSERON, 1995, p. 470), e, segundo, porque estabelecem uma dicotomia, bastante interessante para os objetivos do presente trabalho, com as expressões "língua natural" e "linguagem natural", também utilizadas por Passeron. As línguas ou linguagens naturais seriam aquelas utilizadas pelo ser humano em sua comunicação usual com outros seres humanos e por meio da qual são eles capazes de realizar descrições sobre o mundo que os cerca e narrativas sobre as experiências vividas ou observadas. Essa explicação sucinta é suficiente para que o leitor possa compreender a argumentação apresentada no presente artigo.

ências de Karl Popper – mas ao final, e por outras razões, também de Thomas Kuhn[3] – não teriam aplicação. Nesse sentido, não se aplicariam às ciências sociais a "demarcação" e a "falseabilidade" de Popper[4] e os "paradigmas" e a "ciência normal" de Kuhn[5], ideias essas que ambos os filósofos afirmaram ser aplicáveis a todas as ciências[6].

Caso Passeron esteja certo, a sociologia já teria atingido sua maturidade com uma abordagem específica das ciências sociais históricas – consistente no "raciocínio natural" ou "raciocínio sociológico" -, o que afastaria em definitivo a hipótese de unicidade do método científico e posicionaria o âmbito de validade das ideias de Popper e Kuhn apenas nas ciências naturais.

Não se pode afirmar que a posição defendida pelo autor francês – por ele construída ao longo da década de 1980 -, tenha prevalecido. Conforme apontado por Roberto M. Ferreira "(...) desde os anos 90 vem crescendo o número de autores que, talvez descontentes com os resultados do *interpretive turn*, voltaram a defender a possibilidade do unitarismo metodológico", o que "implica recusar a suposta especificidade epistemológica do objeto das ciências sociais" (FERREIRA, 2008: 19).

O presente artigo insere-se nesse debate, qual seja, em termos amplos, sobre as especificidades dos métodos a serem utilizados nas ciências sociais em relação àqueles utilizáveis nas ciências naturais, e, em termos específicos, sobre a validade das referidas afirmações e Popper e Kuhn para as ciências sociais.

Estabelecida a oposição específica entre as ideias de Passeron e as de Popper e Kuhn, este trabalho se propõe a analisar um dos pilares da tese de Passeron. Extrai-se do texto de Passeron que o "raciocínio sociológico" é moldado a partir de duas características prin-

[3]O debate entre Karl Popper e Thomas Kuhn é amplo, não cabendo aqui revisitá-lo. As razões pelas quais as teses de Passeron se opõem a ambos será esclarecida ao longo do presente artigo.

[4]O próprio Popper parece concordar com essa afirmação de Passeron, uma vez que propôs dois métodos específicos para as ciências sociais, a análise situacional e a engenharia social de ação gradual, bastante diferentes do método por ele proposto para as ciências naturais, decorrente da demarcação. Não obstante, as razões popperianas para tanto são bem diferentes daquelas apresentadas por Passeron. O presente artigo analisa um conjunto específico de razões apresentado por Passeron.

[5]Kuhn também parece concordar com essa afirmação de Passeron, mas a razão para tanto advém de uma discordância entre ambos: Kuhn acreditava que as ciências sociais não tinham se desenvolvido integralmente, e apenas por essa razão não tinham ainda atingido o estágio paradigmático e de ciência normal. Como já afirmado na nota anterior, o presente artigo analisa um conjunto específico de razões apresentado por Passeron.

[6]Exemplos dessas afirmações podem ser encontrados em POPPER, 2007, p. 35 ("denomino problema de demarcação o problema de estabelecer um critério que nos habilite a distinguir entre as ciências empíricas, de uma parte, e a Matemática e a Lógica, bem como os sistemas 'metafísicos', de outra") e KUHN, 2009, p. 31 ("a aquisição de um paradigma (...) é um sinal de maturidade de qualquer campo científico").

cipais. A primeira, a de necessariamente ser construído em linguagem natural. A segunda, a de necessariamente referir-se a objetos inseridos em contextos delimitados espacial e temporalmente (históricos). A argumentação aqui desenvolvida restringe-se a analisar o primeiro desses aspectos, em especial algumas das consequências que Passeron afirma advirem da utilização da linguagem natural nas ciências sociais.

Sobre a possibilidade de crítica das proposições de Passeron por meio da refutação empírica

Uma vez que Passeron pretendeu excluir Popper e Kuhn do jogo das ciências sociais, nada mais justo que se utilize os argumentos de pelo menos um deles para tentar fazer com que ambos retornem. Para tanto, a opção metodológica de análise e crítica das teses de Passeron, relativas às consequências da utilização da linguagem natural pelas ciências sociais, será inspirada nas ideias popperianas de refutação empírica, conforme sucintamente descrito adiante.

A construção do autor francês é filosófica - um exercício crítico a ser examinado e criticado e não a ser refutado -, e, portanto, talvez não seja de todo correto utilizar, a título de instrumento de crítica, um método criado para teste de teorias científicas. Por outro lado, e ao construir suas proposições, Passeron aparentemente atravessou a tênue fronteira entre asserções metafísicas e asserções empíricas. Se isso é verdade, como se tentará demonstrar mais adiante, suas proposições passam a ser empiricamente criticáveis.

Em outras palavras, e como será detalhadamente descrito a seguir, é a forma lógica sob a qual foram redigidas grande parte das proposições apresentadas por Passeron que permite que o método de refutação empírica aqui utilizado, inspirado nas ideias de Popper, seja a elas aplicável. Assim, o que torna possível criticar empiricamente as proposições apresentadas por Passeron, não reside no entendimento de serem "científicas" ou de o autor tê-las construído com essa intenção, e também não depende dos conceitos de teoria científica ou de explicação científica construídos por Popper, mas sim do fato de serem logica e empiricamente refutáveis.

O método de refutação empírica popperiano pode ser logicamente descrito (ou encontra-se logicamente fundamentado) por meio do *modus tollens* da seguinte forma (POPPER, 2007, p. 80):

> - se uma conclusão "p" é dedutível do sistema de enunciados (proposições) "t",

ISSN: 2317-9570

- e se "p" revela-se falsa,
- então "t" será necessariamente falso.

Para avançar em nossa análise crítica sobre as proposições de Passeron torna-se necessário, então, deduzir consequências de suas proposições. Pela restrição estabelecida ao escopo do presente trabalho, devem ser analisadas consequências relacionadas exclusivamente à utilização de linguagem natural pelas ciências sociais.

A primeira dessas consequências pode ser extraída diretamente da Proposição 1: "As ciências empíricas são linguagens de descrição do mundo que devem produzir um tipo particular de conhecimento com as provas empíricas que a estrutura lógica dessas linguagens torna possíveis e necessárias" (PASSERON, 1995, p. 400). Se, conforme afirma Passeron, "as ciências empíricas são linguagens de descrição do mundo" e se essas linguagens produzem "um tipo particular de conhecimento" advindo "da estrutura lógica dessas linguagens", o que determinaria, inclusive, as provas empíricas possíveis em seu âmbito, então é porque *cada ciência tem suas provas empíricas, sua linguagem e seu tipo de conhecimento específicos* (primeira consequência).

A segunda dessas consequências pode ser extraída da combinação de subproposições das Proposições 1, já referida, e 2: "Não existe e não pode existir linguagem protocolar unificada da descrição empírica do mundo histórico" (PASSERON, 1995, p. 405), quais sejam, a proposição 2.4: "A sociologia, assim como a história ou a antropologia, em seus enunciados finais, só se pode falar em língua natural" (PASSERON, 1995, p. 422), a proposição 1.2.1: "Um alto grau de consenso realizado num grupo de especialistas e relacionado a um alto grau de estabilização de uma linguagem de descrição do mundo define um 'paradigma' científico" (PASSERON, 1995, p. 403) e a proposição 2.1: "A sociologia não toma e nem pode tomar a forma de um saber cumulativo, isso é, um saber cujos conhecimentos acumulados fossem organizados por um paradigma teórico" (PASSERON, 1995, p. 407). Assim, se cada ciência tem sua linguagem específica (já referida proposição 1), se a linguagem específica da sociologia é a linguagem natural (proposição 2.4), se apenas um "alto grau de consenso realizado num grupo de especialistas", consenso esse dependente de "um alto grau de estabilização de uma linguagem", torna possível a definição de um paradigma científico (proposição 1.2.1), e se a "sociologia não (...)

pode tomar a forma de um saber cumulativo" organizado "por um paradigma teórico" (proposição 2.1), então é porque *a linguagem natural, sempre "instável", impede a estabilização da linguagem científica e, consequentemente, a constituição de "paradigmas", a realização de ciência normal e a acumulação de conhecimentos científicos* (segunda consequência).

A terceira e última consequência pode ser extraída da combinação das Proposições 1, já referida, e 3: "A prova empírica de uma proposição teórica jamais pode em sociologia revestir-se da forma lógica da 'refutação' ('falsificação') no sentido popperiano" (PASSERON, 1195, p. 426) com subproposições da proposição 2, quais sejam a já referida proposição 2.4 e a proposição 2.2: "A vulnerabilidade e, portanto, a pertinência empíricas dos enunciados sociológicos não podem ser definidos a não ser numa situação de levantamento parcial da informação sobre o mundo que é o da observação histórica – jamais a da experimentação" (PASSERON, 1995, p. 409). Assim, se os tipos de prova empírica a serem utilizados em uma dada ciência são determinados pela linguagem nela utilizada e respectiva estrutura lógica (proposição 1), se "a sociologia (...) só se pode falar em língua natural" (proposição 2.4), se "a pertinência empírica dos enunciados sociológicos" só pode ser realizada por meio de "observação (...) jamais da experimentação" (proposição 2.2), e se "a prova empírica de uma proposição teórica(...) em sociologia" não pode "revestir-se da forma lógica da 'refutação' ('falsificação') no sentido popperiano" (proposição 3), então é porque *a utilização de linguagem natural em uma dada ciência impede a realização de observação e de experimentos falseadores no sentido popperiano* (terceira consequência).

Esclarecido como se dá a refutação empírica em sua estrutura lógica e a forma como foram deduzidas da teoria de Passeron as três consequências que são aqui submetidas à crítica, resta esclarecer como essas consequências serão efetivamente submetidas ao confronto empírico[7].

Quase todas as proposições de Passeron expressam impossibilidades existenciais, uma vez que

[7] Por confronto empírico quer-se aqui dizer a busca de uma ocorrência no mundo real declarada impossível pela consequência decorrente das proposições de Passeron. Caso encontrada essa ocorrência, então terá ocorrido a refutação. O método lógico-empírico aqui utilizado funda-se, como já salientado, nas ideias de Popper, cujos aspectos essenciais estão aqui sendo observados: a dedução de consequências a partir de um sistema de proposições; a possibilidade de descrição prévia, em abstrato, de uma ou mais ocorrências que refutem essas consequências; a observação ou verificação, em concreto, da ocorrência refutadora prevista em abstrato; e a possibilidade de a observação dessa ocorrência ser reproduzível.

ISSN: 2317-9570

construídas com expressões como "jamais" (proposição 1.1, 2.2, 3, 3.1.1), "nenhum" ou "nenhuma" (proposição 1.1.1, 3.1), "não existe e não pode existir" (proposição 2), "não toma nem pode tomar" (proposição 2.1), "não podem" e "não pode" (proposição 2.2, 2.2.1, 2.2.2), "só pode" – no sentido de "não pode ser diferente de" (proposição 2.2.3, 2.4), "não passível de" (proposição 2.3), ou "devem ser sempre" – no sentido de "não podem deixar de ser" (proposição 2.4.1) (PASSERON, 1995, pp. 400 a 461). Pode-se afirmar, portanto, que as proposições de Passeron são construídas sob a forma de negativas de enunciados estritamente existenciais, na forma genérica "não há e não haverá ...". A já referida proposição 2.1, por exemplo, pode ser escrita na forma "não há e não haverá saber acumulativo na sociologia organizado por um paradigma teórico".

Analisando logicamente os enunciados universais, Popper havia concluído que são equivalentes a negações de enunciados estritamente existenciais, ou seja, que a todo enunciado universal corresponde uma negativa de enunciado estritamente existencial de sentido lógico equivalente, e vice-versa (POPPER, 2009, p. 72). Com efeito, é fácil perceber que expressões do tipo "não toma nem pode tomar", extraída da referida proposição 2.1, expressam um enunciado universal. O enunciado "todos os cisnes são brancos" pode ser escrito sob a forma "um cisne não toma e não pode tomar uma cor diferente de branco".

Popper explica que ambos os tipos de enunciado, estritamente universais e negativas de estritamente existenciais, "precisamente por agirem" no sentido da "não existência de certas coisas ou estados de coisas, proscrevendo ou proibindo, por assim dizer, essas coisas ou estados de coisas", tornam-se logicamente refutáveis por um enunciado existencial singular (POPPER, 2009, p. 72). Empiricamente, a proibição de que existam cisnes de cores diferentes de branco pode ser falseada pela simples apresentação de um cisne de qualquer outra cor, apresentação essa que pode ser expressa sob a forma de enunciados existenciais singulares, como, por exemplo, "eis aqui um cisne negro"[8].

Assim como as proposições, as consequências delas deduzidas também são negativas de enunciados estritamente existenciais,

[8]O método de refutação empírica aqui utilizado não requer a realização de experimentos dirigidos por proposições que se pretenda refutar. Observações dirigidas pelas proposições que se pretende refutar detém igual valor, tanto lógico quanto empírico. Isso é totalmente compatível com as ideias de Popper que inspiram a forma de análise aqui utilizada.

de onde decorre que para refutar uma consequência se faz necessário contrapor um enunciado existencial singular. Desse modo, no intuito de testar a pertinência empírica das consequências das proposições de Passeron, que proíbem determinadas ocorrências, serão apresentados exemplos que demonstrem já terem efetivamente ocorrido consequências que as proposições afirmam serem impossíveis.

Sobre a utilização de linguagens nas diversas ciências

Passeron afirma que cada ciência tem uma linguagem que lhe é característica, da qual decorrem necessariamente o tipo de conhecimento produzido naquela ciência e o tipo de prova empírica nela utilizado, ou seja, que cada ciência tem sua linguagem, seu tipo de conhecimento e suas provas empíricas específicas.

Ao longo de todo o seu texto, o autor atribui às ciências naturais características muito específicas das ciências formais e da física, segundo ele consequências necessárias da utilização de linguagens artificiais, e atribui às ciências sociais características específicas da história, segundo ele consequências necessárias da utilização de linguagem natural[9], e nesse contexto dicotômico desenvolve sua argumentação.

Objetivando apreciar a pertinência empírica da afirmação do autor francês faz-se necessário escapar desse posicionamento dicotômico estremado, que cria um viés tendente a retirar do campo de análise todas as demais ciências, e apresentar e analisar exemplos de outras ciências, que não as duas referidas, como, por exemplo, a biologia (como ciência natural) e a economia (como ciência social) analisadas a seguir, e a psicologia cognitiva (como ciência humana), analisada mais adiante em outra seção deste artigo.

Com relação às Ciências Biológicas é possível observar que as áreas de conhecimento que tenham relação com a físico-química, o metabolismo e a genética detenham avançado grau de utilização de linguagens artificiais na construção do conhecimento a elas relacionado. Por outro lado, o conhecimento relacionado, por exemplo, à patologia e à fisiologia, conforme será visto mais adiante, encontra-se quase que totalmente construído em linguagem

[9]Vide nota de rodapé nº 2, retro.

[10]As afirmações contidas no parágrafo podem ser constatadas, por exemplo, em ROBBINS & COTRAN, 2005, e BERNE et al., 2004.

ISSN: 2317-9570

natural[10]. Essas afirmações são válidas para descrições de explicações a respeito de constituição e funcionamento dos seres vivos tanto em zoologia, quanto em botânica, quanto em medicina.

Um aspecto interessante a ser suscitado refere-se ao fato de a Ecologia utilizar a linguagem natural e linguagens artificiais (estatísticas e grafismos geométricos) na descrição dos mais variados meio-ambientes - que dizem respeito ao mundo natural e que, apesar disso, dizem respeito a contextos espaciais diferenciados –, demonstrando a extrema flexibilidade existente na utilização das linguagens pelos cientistas.

O exemplo mais contundente da utilização da linguagem natural nas ciências naturais talvez esteja na Teoria da Evolução – teoria estruturante de toda a biologia contemporânea, que tem ramificações e consequências em praticamente todas as suas subáreas -, que foi integralmente construída em linguagem natural, conforme pode ser facilmente observado no exame do livro *Da Origem das Espécies*, de Charles Darwin (DARWIN, 2006).

Dados esses exemplos de utilização de linguagem natural nas ciências naturais, cabe dar exemplo de utilização intensiva, e até mesmo estrutural, de linguagens artificiais nas ciências humanas. Entre essas, a Economia é provavelmente aquela que atingiu maior grau de formalização. São famosas na Microeconomia as curvas que demonstram a análise gráfica das correlações entre diversos conceitos, como por exemplo, as curvas de oferta e demanda e as respectivas regras para o deslocamento dessas curvas, ou para de elasticidade-preço da demanda, ou para demonstrar conceitos como o de receita marginal (FRANK, 2012, pp. 82/82, 103 a 105, 243/244, respectivamente). Não obstante, toda essa análise pode ser também realizada de forma algébrica, evidenciando de forma ainda mais clara proporções e relações quantitativas (FRANK, 2012, pp. 93 e 94). Apesar de essas relações serem mais utilizadas em Microeconomia, esse tipo de formalização e a utilização de gráficos para demonstrar correlações também são largamente utilizados em Macroeconomia, como, por exemplo, curvas de demanda por mão de obra em relação a preços de produtos (FRANK, 2012, pp. 488 e 489).

O melhor exemplo de utilização estrutural de linguagens artificiais nas ciências humanas talvez seja encontrado na Econometria. As várias definições possíveis dessa área do conhecimento remetem à necessidade de “dar suporte empírico aos modelos construídos pela economia matemática", ou de analisar quantitativamente "os fenômenos econômicos concretos, baseada no desenvolvimento simultâneo de teoria e observação, relacionadas por métodos de inferência adequados", de forma que a "econometria pode ser definida como a ciência social na qual as ferramentas da teoria econômica, matemática e inferência estatística são aplicadas à análise dos fenômenos econômicos" (GUJARATI, 2000, Introdução, p. XXVI).

É possível observar, a partir desses exemplos, que as ciências não são linguagens específicas de descrição do mundo, como afirmou Passeron em sua já referida Proposição 1, ou seja, que cada ciência tenha sua linguagem específica ou característica. Modo contrário, os exemplos examinados indicam que as ciências são descrições do mundo realizadas nas diversas linguagens disponíveis, artificiais e natural, que são utilizadas aparentemente a partir das necessidades dos cientistas, da disponibilidade de ferramentas e da compatibilidade entre linguagem e o objeto específico de estudo, entre outros possíveis aspectos.

Com relação à afirmação no sentido de que a predominância de uma ou outra linguagem determina o tipo de conhecimento a ser produzido e o tipo de prova a ser utilizada em uma dada ciência, as duas seções seguintes apresentam exemplos de que é possível a constituição de paradigmas, a realização de ciência normal e o acúmulo de conhecimento científico com a utilização exclusiva de linguagem natural e de que é possível, utilizando-se exclusivamente a linguagem natural, construir experimentos diretos ou buscar observações diretas que objetivem promover o já referido confronto empírico.

Sobre a possibilidade de consolidação de paradigmas em linguagem natural

Conforme já apontado, Passeron afirma que a linguagem natural, sempre "instável", impede a estabilização da linguagem científica e, consequentemente, a constituição de "paradigmas", a realização de ciência normal e a acumulação de conhecimentos científicos.

Demonstra-se aqui, com pelo menos dois exemplos extraídos da biologia (o segundo deles da medicina, aqui considerada como subárea da biologia), que a utilização de linguagem natural não impede a consolidação de paradigmas.

Em seu livro, A Origem das Espécies, Darwin não utilizou

ISSN: 2317-9570

linguagens artificiais (DARWIN, 2006, pp. 453 a 760). Toda a teoria foi construída, demonstrada e fundamentada exclusivamente em linguagem natural, e transformou-se no que talvez seja o exemplo de paradigma científico contemporâneo de maior estabilidade e mais ampla aceitação.

Darwin apoiou sua teoria em três pilares empíricos, quais sejam (a) observações por ele realizadas sobre a variedade de animais domésticos provocada intencionalmente pelos criadores por meio da seleção artificial de características desejáveis (DARWIN, 2006, cap. I, pp. 453 a 476), (b) observações por ele realizadas sobre a variedade de animais selvagens, especialmente em sua famosa viagem às Ilhas Galápagos a bordo do Beagle, que ele atribuiu à seleção natural (DARWIN, 2006, caps. II a VIII, pp. 477 a 626), e (c) observações sobre os registros geológicos de fósseis (DARWIN, 2006, caps. IX e X, pp. 627 a 669). A base empírica observacional coletada por Darwin foi suficientemente forte para dar sustentação a sua teoria, a qual foi objeto de reforços posteriores, quer seja por novas observações e pesquisas, quer seja pela evolução de áreas de conhecimento correlatas, como a genética. De qualquer forma, parece não haver dúvidas de que o paradigma criado por Darwin teve influência direta, ou mesmo orientou, boa parte da ciência normal realizada no campo da biologia e, com isso, propiciou um enorme acúmulo de conhecimento, a maior parte do qual produzido em linguagem natural.

Concluindo o raciocínio, sobre a Teoria da Evolução pode-se afirmar duas coisas: primeiro, que nos dias atuais não mais existem cientistas, especialmente biólogos, que questionem seu *status* científico[11]; e, segundo, que não existem dúvidas no sentido de que a Teoria da Evolução tornou-se um paradigma científico em sentido kuhniano (e é isso o que aqui se discute até esse ponto).

Especificamente sobre a possibilidade de realização de experimentos ou observações em linguagem natural, que atendam o critério popperiano de falseabilidade, é de se reconhecer que experimentos engenhosamente construídos, tanto em laboratórios – como o de Richard Lenski, e outros, acompanhando 45 mil gerações de bactéria Escherichia coli, ao longo de 20 anos, e sua evo-

[11] Vide, por exemplo, DENNETT, 1995, pp. 18-21

[12] Os resultados dos experimentos de Lenski *et al* estão descritos em DAWKINS, 2009, pp. 114-130, e foram originalmente publicados em *Procedings of the National Academy of Science*, nº 91, 1994, pp. 6808-14, *apud* DAWKINS,

lução frente a contingências ambientais intencionalmente provocadas[12] – como no próprio ambiente natural (ou ambiente natural restrito) – como os de John Endler com peixes *Poecilia reticulata* e sua evolução frente a contingências naturais (predadores e alterações ambientais)[13], sucintamente descritos adiante - vem sistematicamente corroborando consequências dedutíveis da Teoria da Evolução. Com relação a observações falseadoras, o próprio Popper referiu-se ao "melanismo industrial"[14] por meio do qual pode-se observar a seleção natural ocorrendo "sob nossos olhos" (POPPER, 1995, p. 242).

Outras áreas da biologia também podem ser consideradas paradigmas kuhnianos. Em medicina, a Patologia, enquanto "estudo das alterações estruturais e funcionais que ocorrem nas células, tecidos e órgãos decorrentes de doenças", ao tempo em que "tenta explicar os motivos dos sinais e sintomas que os pacientes manifestam", fornecendo "uma base racional para a abordagem clínica e o tratamento" (ROBBINS & COTRAN, 2005, p. 4); e a Fisiologia, enquanto descrição "da função dos organismos, nos vários estágios da organização, do nível subcelular ao organismo intacto" (BERNE, 2004, p. xiii); encontram-se quase que totalmente construídas em linguagem natural, o que pode ser observado a partir de manuais utilizados no ensino nas universidades de medicina a respeito da fisiologia e patologias humanas[15]. Ainda que, em termos de Fisiologia, a elucidação dos mecanismos homeostáticos requeira um aprofundamento em bioquímica, e, portanto, a utilização parcial de linguagens artificiais, isso não invalida a afirmação de extensa utilização da linguagem natural nesse campo de conhecimento.

Tais exemplos atendem aos requisitos necessários para serem tidos como "paradigmas" kuhnianos, quais sejam, (a) constituírem um corpo de conhecimento científico coerente, (b) que goze da aceitação da maior parte da comunidade científica, não sendo exigida a unanimidade, (c) que oriente a pesquisa posterior, estabelecendo

2009, p. 417.

[13]Os experimentos de Endler estão descritos em DAWKINS, 2009: 130-136, e foram originalmente publicados nas revistas *Evolution*, nº 34, 1980, pp. 76-91 e *Environmental Biology of Fishes*, nº 9, 1983, e no livro *A Natural Selection in the Wild*, Princeton, Princeton University Press, 1986, apud DAWKINS, 2009, pp. 415-416.

[14]Popper refere-se aqui às famosas observações no sentido de que mariposas cor-de-fuligem estariam sendo naturalmente selecionadas, por gozarem de proteção contra predadores fornecida por mimetismo, em áreas urbanas atingidas pela fuligem industrial.

[15]Por exemplo em ROBBINS & COTRAN, 2005, e BERNE *et al.*, 2004.

o "quebra-cabeças" a ser resolvido, e permitindo assim a ocorrência da "ciência normal", e (d) que, com isso, possibilite o acúmulo de conhecimento científico e o progresso da ciência (KUHN, 2009, Capítulos 1, 2 e 3, e Posfácio: 221 a 227).

A existência desses exemplos, entre outros, leva-nos à conclusão de que a utilização exclusiva de linguagem natural na construção de um corpo de conhecimentos científicos não impede o surgimento de paradigmas, a realização de ciência normal e o acúmulo de conhecimento científico, em sentido kuhniano, contrariando o afirmado por Passeron em suas já referidas proposições.

Sobre a possibilidade de realização de experimentos em linguagem natural

Conforme apontado, Passeron afirma que a vulnerabilidade empírica de uma ciência que utiliza linguagem natural só pode basear-se em observações, enquetes, investigações, etc. e nunca em experimentos intencionalmente realizados.

A contestação empírica de tal afirmação exige que sejam apontados exemplos que demonstrem a possibilidade de idealizar, realizar e descrever experimentos, e posteriormente registrar e publicar seus resultados, exclusivamente em linguagem natural, e que isso já é realizado em diversas ciências. Ainda que Passeron indique a existência de outros fatores que, aliados à utilização de linguagem natural, possam impedir ou dificultar a experimentação direta nas ciências sociais, o que é aqui analisado é a afirmação de que o elemento "linguagem natural", isoladamente, é suficiente para a produção de tal resultado.

Tomemos como primeiro exemplo o artigo Julgamento sob incerteza: heurísticas e vieses[16], de autoria de Kahneman e Amos Tversky, ganhadores do prêmio Nobel de Economia de 2012 pelo desenvolvimento de teorias de decisão. Logo ao início do referido artigo os autores apontam que "muitas decisões estão baseadas em crenças relativas à probabilidade de eventos incertos, tais como o resultado de uma eleição, a culpa de um réu ou a futura cotação do dólar", e estabelecem o problema a ser tratado no artigo: "o que determina essas crenças?". Suscintamente respondem logo a seguir, "as pessoas se apoiam em um número limitado de princí-

[16] Artigo originalmente publicado na revista *Science*, vol. 185, 1974, integralmente incluído em KAHNEMAN, 2012, pp. 524 a 539

pios heurísticos que reduzem as tarefas complexas de avaliar probabilidades e predizer valores a operações mais simples de juízo. De um modo geral essas heurísticas são bastante úteis, mas às vezes levam a erros graves e sistemáticos". No corpo do artigo os autores descrevem "três heurísticas que são empregadas para avaliar probabilidades e prever valores", descrições por meio das quais "os vieses aos quais essas heurísticas conduzem são enumerados e as implicações aplicadas e teóricas dessas observações são discutidas" (KAHNEMAN, 2012, pp. 524 e 525).

As três heurísticas analisadas pelos autores são por eles denominadas "representatividade", "disponibilidade" e "ajuste e ancoragem", cujos respectivos vieses são analisados a partir de um ou mais experimentos diretos. Relativamente à primeira heurística, vez que não nos interessa aqui analisar todas elas, os autores analisam os vieses descritos como (a) a insensibilidade das pessoas à probabilidade à priori de resultados, (b) a insensibilidade ao tamanho amostral, (c) concepções errôneas de possibilidades, (d) insensibilidade à previsibilidade, (e) ilusão de validade e (f) concepções errôneas de regressão. Para cada um desses pontos analisados os autores descrevem pelo menos um experimento direto realizado por eles ou por terceiros. Descrever todos os experimentos fugiria ao escopo do presente trabalho, assim, e como exemplo, tomaremos um dos experimentos descritos pelos autores em relação ao viés (b) insensibilidade ao tamanho amostral, o qual envolveu a apresentação a 95 alunos de graduação da seguinte questão (KAHNEMAN, 2012, pp. 526 e 527):

> Uma determinada cidade é atendida por dois hospitais. No maior, cerca de 45 bebês nascem todo dia, e no hospital menor nascem cerca de 15 bebês por dia. Como você sabe, cerca de 50% dos bebês são meninos. Entretanto, a porcentagem exata varia no dia a dia. Às vezes pode ser mais elevada que 50%, às vezes menos. Pelo período de um ano os dois hospitais registraram os dias em que mais do que 60% dos bebês eram meninos. Qual hospital você acha que registrou mais dias desses?
>
> O hospital maior. (21)
>
> O hospital menor. (21)
>
> Mais ou menos iguais (ou seja, dentro de 5% um do outro) (53)

Considerando que os valores

ISSN: 2317-9570

entre parênteses são o número de alunos que escolheram cada resposta, os autores apontam que "a maioria dos participantes julgou a probabilidade de obter mais do que 60% como sendo a mesma para o hospital pequeno e o grande" a despeito de a teoria da amostragem exigir "que o número esperado de dias em que mais de 60% dos bebês são meninos é muito maior no pequeno hospital do que no maior, pois uma grande amostra tem menor probabilidade de se afastar de 50%", a partir do que concluem que "essa noção fundamental de estatística evidentemente não faz parte do repertório intuitivo das pessoas" (KAHNEMAN, 2012, p. 527).

Esse experimento, assim como todos os demais descritos no referido artigo, foi integralmente idealizado, realizado (com exceção de operações estatísticas elementares) e descrito, e posteriormente teve seus resultados registrados e publicados, integralmente em linguagem natural. Trata-se, sem qualquer dúvida, de um experimento direto, e não de uma mera observação sistemática, o qual pode ser reproduzido em qualquer parte do mundo por outros cientistas e por outras instituições utilizando-se de outros grupos amostrais. É fácil observar que a "teoria" dos autores – os seres humanos se apoiam em um número limitado de heurísticas para reduzir tarefas complexas de juízo a tarefas mais simples - e seu componente específico – algumas noções fundamentais de estatística, entre elas as decorrentes do tamanho da amostragem, não fazem parte do repertório intuitivo das pessoas –, são empiricamente confrontáveis por esse e por outros experimentos diretos facilmente imagináveis.

Poder-se-ia alegar que, nesse exemplo, a consequência a ser empiricamente refutada não decorre de um enunciado estritamente universal, mas sim de um enunciado numericamente universal, uma vez que o conjunto dos seres humanos seria finito, limitado a determinado espaço, e contável em determinado momento do tempo. Quanto a isso pode-se afirmar que a teoria exposta pelos autores no artigo, conforme descrita no parágrafo imediatamente anterior, tem nítida pretensão de ser universal, no sentido de ser aplicável a todos os seres huma-

[17] É o próprio Popper quem afirma que "a questão de saber se as leis da ciência são estritamente ou numericamente universais não pode ser resolvida através da argumentação. Trata-se dessas questões que só podem ser resolvidas por acordo ou convenção". Sua preferência deve-se apenas a ele considerar "útil e frutífero encarar as leis naturais como enunciados sintéticos e estritamente universais", tendo em vista aspectos metodológicos - em POPPER, 2009, cap. 13. Universalidade Estrita e Numérica, pp. 66.

nos, ainda que a totalidade dos seres humanos seja uma universalidade numérica[17] e que algumas pessoas, em razão de treinamento profissional ou experiência de vida, estejam preparadas para substituir determinadas heurísticas por conhecimento real e, em consequência, evitar determinados vieses. Nesse sentido, Kahnemam aponta que por meio do conjunto de experimentos foram documentados "erros sistemáticos na opinião das pessoas normais" e que ele e Tversky localizaram "esses erros no projeto do mecanismo cognitivo" (KAHNEMAN, 2012, p. 16), mecanismo cognitivo esse pertencente a todos os seres humanos.

Sobre a qualidade científica e impacto do artigo, o próprio autor (KAHNEMAN, 2012, p. 16) acentua que ele "continua sendo um dos trabalhos em ciência social mais amplamente citados (mais de trezentos artigos acadêmicos fizeram referência a ele em 2010)" e que:

> (...) estudiosos de outras disciplinas acharam-no útil, e as ideias de heurísticas e vieses tem sido utilizadas proveitosamente em inúmeros campos, incluindo diagnósticos médicos, análises judiciais, serviços de inteligência e espionagem, filosofia, finanças, estatísticas e estratégia militar. [18]

Como segundo exemplo, agora na área de biologia, o já referido experimento de John Endler a respeito dos efeitos da seleção natural sobre peixes popularmente conhecidos como "*guppy*", inicialmente publicado por meio do artigo *Natural Selection on Color Patterns in Poecilia reticulata* (ENDLER, 1980), mostra-se como excelente exemplo de experimento construído em linguagem natural com o objetivo de testar consequências pré-estabelecidas de uma teoria também construída integralmente em linguagem natural, a Teoria da Evolução por meio de Seleção Natural.

Os machos da espécie de peixe selecionada por Endler apresentam fortes e variados padrões de cores, e, apesar disso, foi sistematicamente observado que nos rios onde não havia predadores as cores dos machos eram intensas e onde havia predadores fortes as cores dos machos tendiam a se esmaecer e se tornarem mais próximas em padrão e tonalidade do cascalho existente no fundo. A explicação causal dada pela Teo-

[18] Como exemplos de utilização das pesquisas de Kahneman em microeconomia veja FRANK, 2012, Capítulo I, "Três Armadilhas de Decisões Importantes", pp. 8 a 15.

ISSN: 2317-9570

ria, e em consequência sua previsão, é a de que em um ambiente com predação os machos menos vistosos, que se confundam melhor com o cascalho do fundo, tem maior probabilidade de sobrevivência e, portanto, de reprodução, fazendo com que a característica "cores menos vistosas" seja transmitida predominantemente às novas gerações; já em um ambiente sem predadores, a preferência das fêmeas pelos machos mais coloridos, fenômeno observado em várias espécies, levaria a uma seleção para reprodução a partir desse critério, fazendo com que a característica "cores mais vistosas" seja predominantemente transmitida às novas gerações.

Endler concebeu o experimento de forma a confrontar os dois critérios de seleção natural e a testar integralmente as previsões teóricas. Dividiu dez populações de peixes em tanques que simulavam o rio de onde haviam sido retirados. Em cinco tanques preparou o fundo com cascalho fino e outros cinco com cascalho grosso. As populações habitaram esses tanques por seis meses sem a presença de predadores, período após o qual foi observado que o número de manchas coloridas nos machos havia aumentado explosivamente, exatamente conforme a teoria havia previsto.

A seguir, dois tanques em cada grupo diferenciado pelo tipo de material no fundo receberam predadores fortes, dois receberam predadores fracos e um não recebeu qualquer tipo de predador. Passados cinco e quatorze meses todos os tanques foram recenseados e os resultados apontaram que:

a) nos quatro tanques com predadores fortes o número de manchas coloridas nos machos despencou e a intensidade das cores reduziu-se, já sendo evidente no 5º mês e ainda mais no 14º mês;

b) nos outros seis tanques, sem predadores ou com predadores fracos, o número de manchas continuou a aumentar até estacionar em um platô, demonstrando que, como predadores fracos não impõem uma redução significativa no número de machos, a preferência das fêmeas por "cores mais vistosas" teve maior influência na transmissão dessa característica às novas gerações;

c) nos tanques com predadores, fortes ou fracos, o cascalho grosso do fundo promoveu manchas maiores enquanto o cascalho fino favoreceu manchas menores, indicando que, mesmo no caso da predação fraca, o mimetismo protegeu os animais que se confundissem melhor com o fundo do rio, permitindo-lhes que, sobrevivendo em maior número, transmitissem suas características às novas gerações;

d) nos tanques sem predadores o cascalho grosso promoveu o aparecimento de manchas menores e o cascalho fino de manchas maiores, indicando que não apenas as cores, mas também o contraste das mesmas contra o fundo, ou seja, a visibilidade dos padrões coloridos, favorecia a seleção para reprodução pela preferência das fêmeas.

Dawkins relata que, não satisfeito com os resultados, Endler repetiu o experimento isolando setores de um riacho em que era possível reproduzir as três situações retro descritas em condições naturais, e repetiu as medições de início e de evolução das populações, obtendo os mesmos resultados (DAWKINS, 2009, pp. 135 e 136). As duas séries de experimentos confirmaram todas as previsões teóricas.

Examinados esses dois exemplos de experimentos, ambos construídos em linguagem natural e destinados a testar consequências dedutíveis de teorias também construídas em linguagem natural, a conclusão é novamente contrária às afirmações de Passeron.

Conclusões

Conforme visto, foram deduzidas três consequências das proposições de Passeron, as quais foram analisadas a partir de um procedimento simplificado (apresentação de exemplos contrários) inspirado no método de confronto empírico proposto por Karl Popper.

A análise da primeira consequência demonstrou, por meio dos exemplos apresentados, que não existe uma correlação necessária entre ciência e linguagem específica. No campo das ciências naturais, domínio inconteste das linguagens artificiais segundo Passeron, encontramos as Ciências Biológicas que, apesar de apoiarem-se na física e na química e de utilizarem extensivamente métodos quantitativos em suas pesquisas, apresenta extensas áreas em que o conhecimento é construído estrutural e essencialmente em linguagem natural. No campo das ciências sociais, domínio também incontestável da linguagem natural segundo Passeron, a economia aparece como uma ciência que se apoia cada vez mais na matemática e na estatística, as quais tem uso predominante e estrutural na econometria.

Restou afastada, assim, a proposição de Passeron, que afirma

ISSN: 2317-9570

que as ciências têm, ou utilizam, linguagens que lhes são típicas, as quais determinam seu tipo particular de conhecimento e o tipo de prova empírica de que se valerão. Em verdade, os exemplos demonstram ampla utilização de linguagens naturais e artificiais pelos cientistas, talvez de acordo apenas com as necessidades de sua pesquisa, de seu experimento, de sua demonstração.

A análise da segunda consequência, especialmente por meio do exemplo da Teoria da Evolução por meio da Seleção Natural, de Darwin, demonstrou que a constituição de paradigmas e a realização de ciência normal, em sentido kuhniano, com o consequente acúmulo de conhecimento, exclusivamente em linguagem natural é totalmente possível. Os exemplos de outras áreas da biologia examinados apenas reforçaram tal demonstração.

Foram assim afastadas as proposições de Passeron que afirmavam que a utilização de linguagem natural – por suas ambiguidades e polissemias, por sua indexação a um contexto espaço-temporal sempre alterável, por sua baixa precisão quantitativa, etc - torna impossível acumular conhecimento em torno de um paradigma científico com ela estruturalmente construído.

A análise da terceira consequência foi realizada por meio de dois exemplos de experimentos, um realizado na área de psicologia cognitiva, outro na área de biologia evolutiva, que demonstram que a construção de experimentos diretos e clássicos, bem como o registro e publicação de seus resultados, em áreas do conhecimento constituídas predominantemente também em linguagem natural, é plenamente possível.

Assim, restaram também afastadas as proposições de Passeron que afirmavam que a utilização estrutural de linguagem natural por determinada ciência impede a realização de experimentos diretos em linguagem natural, restringindo a tipologia da prova empírica a serem nela utilizadas à observação sistemática.

Pode-se concluir, portanto, que a afirmação de Passeron, no sentido de que a utilização de linguagem natural nas ciências sociais cria um espaço no qual a demarcação e a refutação de uma teoria por experimentos ou observações empíricas, afirmados por Popper, e os paradigmas e a ciência normal, afirmados por Kuhn, não teriam lugar, não há de ser aceita.

As análises aqui procedidas não afirmam que o critério de demarcação de Popper ou o paradigma e a ciência normal de Kuhn são possíveis nas ciências sociais.

Essa análise implicaria um escopo extremamente amplo e estranho ao presente trabalho. Afirma-se aqui, apenas, que a utilização de linguagem natural, por si só, não excluí a possibilidade de as ideias de Popper e Kuhn serem válidas também para as ciências sociais.

No debate sobre a unicidade do método científico, portanto, o presente artigo contribui apenas no sentido de deixar tal possibilidade em aberto, afastando as restrições construídas por Passeron, relacionadas à utilização de linguagens específicas pelas ciências, que foram aqui analisadas.

Os fatos de um paradigma não ter ainda surgido na sociologia ou de experimentos falseadores não serem extensamente utilizados ou aceitos nesse campo do conhecimento, também não significam, por óbvio, que não são possíveis: o que não foi ainda realizado ou não ocorreu talvez ainda venha a ser ou ocorrer.

Por outro lado, paradigmas e experimentos talvez não sejam mesmo possíveis nas ciências sociais. A título de exemplo, apesar de sempre ter afirmado a unicidade do método científico, o próprio Popper, reconheceu importantes diferenças entre as ciências naturais e as ciências sociais, como, por exemplo, entre as possibilidades de realização de experimentos diretos (POPPER, 1980, caps. 24 e 25), ou de aplicação de métodos quantitativos (POPPER, 1980, p. 111), ou, ainda, ao conceber a análise situacional como método típico das ciências sociais (POPPER, 1980, pp. 110 e 111 e 116 e 117).

Entretanto, sejam quais forem as eventuais razões para que a demarcação popperiana e o paradigma kuhniano não sejam eventualmente possíveis nas ciências sociais, se assim um dia se demonstrar, as conclusões do presente artigo são no sentido de que a utilização estrutural da linguagem natural nas ciências sociais não se incluirá entre elas.

Referências

BERNE, Robert M., LEVY, Mathew, KOEPPEN, Bruce, STANTON, Bruce. *Fisiologia*. Trad: Nephtali Segal Grinbaum e outros. Rio de Janeiro: Elsevier, 2004.

DARWIN, Charles. *On The Origin of Species*, em *From So Simple a Begining: the four great books of Charles Darwin*. Ed. by Edward O. Wilson. 1st ed. New York: W.W. Norton & Company, 2006.

ISSN: 2317-9570

DAWKINS, Richard. *O Maior Espetáculo da Terra: as evidências da evolução*. Trad: Laura Teixeira Motta. São Paulo: Companhia das Letras, 2009.

DENNETT, Daniel C. Darwins' Dangerous Idea. New York: Simon & Schuster, 1995.

ENDLER, John. *"Natural Selection on Color Patterns in Poecilia reticulata"*. In: *Evolution*, nº 34, 1980, pp. 76-91. Disponível em: https://www.jstor.org/stable/2408316?seq=1#page_scan_tab_contents. https://www.jstor.org/stable/2408316?seq=1#page_scan_tab_contents. Acessado em 22/11/2016.

FERREIRA, Roberto Martins. *Popper e os dilemas da sociologia*. São Paulo: Annablume; Fapesp, 2008.

FRANK, Robert H., BERNANKE, Ben S., JOHNSTON, Louis D. *Princípios de Economia.* Trad: Heloisa Fontoura e Monica Stefani. 4ª ed. Porto Alegre: AMGH, 2012.

KAHNEMAN, Daniel. TVERSKY, Amos. *Julgamento sob incerteza: heurísticas e vieses.* In *Science*, vol. 185, 1974, disponível em KAHNEMAN, 2012, pp. 524 a 539

KAHNEMAN, Daniel, *Rápido e Devagar – Duas Formas de Pensar*. Trad. Cássio Arantes Leite. 1ª ed. Rio de Janeiro: Objetiva, 2012.

KUHN, Thomas S. *A Estrutura das Revoluções Científicas.* Trad. Beatriz Vianna Boeira e Nelson Boeira. São Paulo: Perspectiva, 2009.

PASSERON, JEAN-CLAUDE. *O Raciocínio Sociológico: o espaço não-popperiano do raciocínio natural*. Trad. Beatriz Sidou. Petrópolis, RJ: Vozes, 1995.

POPPER, KARL R. *A Miséria do Historicismo*. Trad. Octanny Silveira da Mota e Leônidas Hegenberg. São Paulo: Cultrix, 1980.

__________. *A Lógica da Pesquisa Científica*. Trad. Leônidas Hegenberg e Octanny Silveira da Mota. São Paulo: Cultrix, 2009.

__________. *Popper Selections*. Edited by David Miller. Princeton, Princeton University Press: 1995.

ROBBINS & COTRAN, *Patologia – Bases Patológicas das doenças*. Ed: Kumar, Abbas, Fausto; Trad: Maria da Conceição Zacarias e outros. Rio de Janeiro: Elsevier, 2005.

Relendo *O Processo* de Kafka como a Manifestação de uma Patologia Social

[Rereading Kafka's *The Process* as a Social Pathology Manifestation]

Ronaldo Manzi*

Resumo: Esse texto busca retomar, primeiramente, a leitura de Agamben de *O processo* de Kafka. Sua proposta é que o personagem, Joseph K., realiza uma autocalúnia. Agamben sugere que tal autocalúnia é o que cada homem realiza de si – daí o interesse nesse tema. Seguindo a análise de Agamben, vemos que se trata de uma justificação do culpado perante à lei (como se o direito fosse a forma moderna de redenção de uma culpa). A proposta desse ensaio é interpretar a obra de Kafka por outra via. Pretende-se mostrar que é o social que faz com que K. assuma uma culpa (não havendo, portanto, uma autocalúnia). Nesse caso, haveria uma passagem de uma capacidade de o sujeito duvidar da legitimidade do processo à sua aceitação incondicional – como se o social o levasse a perder essa capacidade de dúvida. Daí porque se dirá de uma patologia do social: uma sociedade que "adoece" o indivíduo.
Palavras-chave: autocalúnia, confissão, tragédia, comédia, patologia social.

Abstract: First of all, this work seeks to resume Agamben's reading of Kafka's *The process*. He intends to show that the character Joseph K. performs an auto-slander. Agamben suggests that such self-slander is what each man does with himself – than, the interest in this theme. Following Agamben's analyses, we could sight that it is about of a justification of the guilty before the law (as if the law were the modern form of redemption a guilt). This essay proposes to interpret Kafka's work in a different way. It wants to show that it is the social whose conducts K. to assume a guilty (therefore there is not an auto-slander). At this point, it could have a passage from the capacity of the subject to doubt of the process' legitimacy to its unconditional acceptance of it – as if the social could lead him to lose his capacity of doubt. Hence, that is why it could be said that it is a social pathology which "sickens" the individual.
Keywords: self-slander, confession, tragedy, comedy, pathology of the social

*Doutor em Filosofia (USP/RUN); Pós-doutor em Filosofia (USP); e Psicanalista. E-mail: manzifilho@hotmail.com.

DOI: https://doi.org/10.26512/rfmc.v6i2.22107

Relendo *O Processo* de Kafka como a manifestação de uma patologia social

Édipo é uma figura marcante na história do pensamento. A peça *Édipo Rei* de Sófocles mostra como Édipo é culpado de algo e segue seu destino – o de ser culpado por algo que ele nem mesmo sabia. Esse é o trágico em Édipo: ser culpado sem saber. Aliás, sem que haja uma acusação formal – é-se culpado; e paga-se o preço dessa culpa: um destino a se cumprir. Não por acaso, em vários momentos, Sigmund Freud destaca como a figura de Édipo é comparável com uma questão bíblica: todos somos culpados desde o pecado original e paga-se o preço dessa culpa.

Giorgio Agamben afirma em um texto intitulado *Comédia* (em *Categorias italianas*, 1996) que "a pessoa é a 'máscara' que a criatura assume e abandona nas mãos do direito para purificar-se" (AGAMBEN, 2014, p. 42). Ou seja, o processo teria uma função de purificação (para livrar o sujeito de sua culpa). Agamben afirma também (já no texto denominado *Nudez* (2009)), que "a pessoa (tanto nas suas vestes trágicas como nas cômicas) é também a portadora da culpa e a ética que [essa] implica é necessariamente ascética, porque fundada numa cisão (entre o indivíduo e a sua máscara, entre a pessoa ética e a jurídica)" (AGAMBEN, 2010, p. 68). Esse ensaio busca mostrar que Joseph K. – personagem central da obra de Kafka – seria este sujeito portador de ética ascética.

É certo que no texto de Agamben, *Comédia*, não há uma referência a *O processo* de Kafka. Sua preocupação é com a literatura e a cultura italiana, mais especificamente com Dante Alighieri. Mas ao analisar Kafka em seu texto *K.*, Agamben diz explicitamente que o universo de *O processo* é cômico, e não trágico (*cf.* AGAMBEN, 2010, p. 32). Guardando a distância entre o trabalho de Dante e Kafka (e a distância temporal), não é "descabido" pensar a ideia de cômico tal como havia definido anos antes nas *Categorias italianas*.*O processo* é uma obra que descreve como K. acorda um dia sem qualquer aviso prévio sendo acusado de algo que não pode ser dito e que ele deve acatar. Trata-se de um processo que K. não tinha notícia nem mesmo dos representantes da justiça. A questão é como ele assume seu "caso": em uma forma de autocalúnia como sugere Agamben, ou, como pretendo mostrar, em um processo de passagem da dúvida à crença (como ficará detalhado mais à frente acompanhando o próprio desenrolar da

ISSN: 2317-9570

obra de Kafka: a capacidade de o sujeito duvidar da correspondência da fala/do anúncio do processo com a realidade, para uma completa crença no processo sem que seja necessário encontrar uma correspondência ou não com essa realidade). Mais especificamente: é K. que chama para si uma culpa ou é a sociedade que a impõe a ele?

Mas quem seria esse personagem? O que significaria esse sobrenome K.? Agamben diz em um texto denominado *K.* (em *Nudez*, 2009) que K. está associado à calúnia – algo vindo do direito romano: *Kalumniator* – aquele que calunia:

> que a calúnia represente a chave do romance – e, talvez, de todo o universo kafkiano, tão potentemente marcado pelas potências míticas do direito – torna-se, contudo, ainda mais esclarecedor se observarmos que, a partir do momento em que a letra K. não substitui simplesmente *kalumnia*, mas em se refere ao *kalumniator*, isto é, ao falso acusador, tal só pode significar que o falso acusador é o próprio protagonista do romance, o qual, por assim dizer, intentou um processo calunioso contra si próprio. O 'alguém' (*jemand*) que, com a sua calúnia, deu início ao processo, é o próprio Josef K. (AGAMBEN, 2010, pp. 31-32).

A calúnia seria a chave do romance que, aliás, descreveria nossa sociedade, uma vez que Agamben afirma que "cada homem intenta um processo calunioso contra si próprio" (AGAMBEN, 2010, p. 32). A calúnia séria a chave de interpretação do texto *desde a primeira linha da obra de Kafka*. Ou seja, o processo só existe porque K. acusa a si mesmo e reconhece a acusação sem saber do que é acusado – ele faria uma autocalúnia: se auto acusa de algo (como veremos). Afinal, o tribunal, nesse processo, não o acusa realmente, mas simplesmente acolhe a acusação que ele faz de si mesmo. Será?

Quem é K.?

Em uma passagem em que K. conhece um personagem que se identifica como "o comerciante Block", ele pergunta ao tal comerciante: "é esse mesmo o seu nome – perguntou K. – Certamente – foi a resposta. – Por que o senhor duvida? – Pensei que pudesse ter motivo para silenciar o seu nome – disse K." (KAFKA, 1997, p. 204). Por indução, podemos pensar que K. é o silêncio de alguém que não pode anunciar a si (porque era essa a desconfiança de K. com o comerciante). Não se trata exatamente de um anonimato. Em alemão, a letra "k" é o que se usa para integrar o artigo indefinido ("kein") que nega os nomes que vêm após ele. K. é um sujeito determinado: Joseph. Mas esse Joseph tem como sobrenome o K. que tem essa possibilidade de negar o determinado (o nome)[1]. Por outro lado, Agamben diz que K. vem de *Kalumniator* (como visto na citação acima). Penso que é mais plausível que K. seja *aquele que poderia negar* os outros personagens (quem poderia colocar todos em dúvida). Não se trata somente de uma especulação – essa outra possibilidade de leitura pode nos levar a repensar a obra de Kafka e como ela representa o que poderíamos denominar patologia social, como será definido à frente.

Poderíamos colocar em outros termos: o que está em jogo aqui é uma tragédia ou uma comédia? Mais especificamente: de algo que é um destino ou um acontecimento de redenção (ser redimido pela lei)?

A sugestão de Agamben é que se trata de uma comédia. No texto K., ele afirma:

> cada homem intenta um processo calunioso contra si próprio. É este o ponto de partida de Kafka. Por isso o seu universo não pode ser trágico, mas somente cómico: a culpa não existe – ou, antes, a única culpa é a auto-calúnia, que consiste no acusar-se de uma culpa inexistente (isto é, da sua própria inocência, e é este o gesto cômico por excelência) (AGAMBEN, 2010, pp. 32-33).

Mas o que se entende aqui por trágico e cômico?

Há um subcapítulo no texto *Comédia* que Agamben nomeia *Culpa trágica e culpa cômica*. A culpa, segundo seu estudo, estaria ligada ao destino final do sujeito em sua vida: o homem acaba se salvando ou há uma danação? Que se perceba que, de todo modo, há uma culpa. A questão é se ele se salva ou se dana. No cômico o sujeito aparece como culpado ou inocente perante a justiça divina; na tragédia se culpa um sujeito sem que haja redenção. De uma forma geral: "(...) *a tragédia aparece como a culpabilização do justo e a comédia como a justificação do culpado*" (AGAMBEN, 2014, p. 26).

O que se pensa como trágico fica mais interessante se levar-

[1] Devo essa observação à minha amiga Paula Alves Martins de Araújo.

mos em conta o pecado natural (bíblico, por exemplo) e o pecado pessoal (aquilo que cada um faz). Ou seja, haveria um pecado que nós herdamos simplesmente por sermos homens (algo independente de nossa vontade, como a vergonha da própria nudez); também o pecado que vem de nossa responsabilidade pessoal. De qualquer modo, o sujeito é "aparentemente" inocente e depois culpado, seja pelos seus atos, e/ou por ter nascido com essa culpa. Por outro lado, a análise de Agamben sobre Dante, nos mostra como no cômico há uma passagem de um início horrível, como a entrada no inferno na *Divina comédia*, e a ascensão até o céu – de culpado a inocente (e não da "inocência" à culpa, tal como Édipo).

Eis a conclusão de Agamben: "a contraposição entre o personagem 'cômico' Dante, que se purifica da culpa pessoal mostrando até o fim a sua vergonha, e Édipo, o herói trágico que, enquanto *pessoalmente* inocente, não pode nem confessar sua culpa nem aceitar a vergonha, não poderia ser mais clara" (AGAMBEN, 2014, p. 36). Assim, guardando a diferença com a obra de Dante, mas levando em conta que Agamben afirma que *O processo* de Kafka é uma obra cômica, podemos pensar que a pessoa se purifica, afinal, mesmo que seja sob o julgo do direito – um romance que mostra como, na modernidade, é o edifício jurídico que purifica a pessoa.

Minha questão é: a situação de K. é de fato cômica?

Agamben diz no texto *K.*:

> só há calúnia, com efeito, se o acusador estiver convencido da inocência do acusado, se acusar sem que haja uma culpa a demonstrar. No caso da autocalunia, esta convicção torna-se ao mesmo tempo necessária e impossível. O acusado, na medida em que se autocalunia, sabe perfeitamente estar inocente, mas, na medida em que se acusa, sabe igualmente bem ser culpado de calúnia, merecer o seu labéu. Tal é a situação kafkiana por excelência. Mas porque é que K. – porque é que cada homem – se autocalunia, se acusa falsamente? (AGAMBEN, 2010, p. 33).

Que se veja: o que era uma simples interpretação de uma obra se torna a referência para se pensar *cada homem*... Eis porque me volto a esse tema – Agamben generaliza a sua interpretação desse processo para se pensar o homem em geral: "K. (cada homem) autocalunia-se para se subtrair à lei, à acusação que aquela parece inexoravelmente dirigir-lhe e à qual não é possível fugir (...)" (AGAMBEN, 2010, p. 41).

Fazer com que o sujeito se sinta culpado – eis a força da lei segundo Agamben: "a estratégia da lei consiste em fazer o acusado crer que a acusação (a porta) lhe é destinada (talvez) a ele precisamente, que o tribunal exige dele (talvez) alguma coisa, que está (talvez) em curso um processo que lhe diz respeito" (AGAMBEN, 2010, p. 42). Esses "talvez" de Agamben me permite pensar de outra forma. É certo que, seguindo Agamben, por não haver uma acusação real, nem mesmo uma representação da lei senão por alusões (pois nenhum juiz é de fato anunciado), o que haveria seria somente uma autopunição em que o sujeito acusa a si mesmo e se deixa capturar pela lei. Afinal, se se deixa morrer por algo que nem anuncia seu nome, porque a porta da lei não prescreve nada. Entretanto, quero destacar mais uma força que vigora como lei e te pune sem que haja de fato um processo que se passa como imaginário e em que todos acreditam e sabem sobre ele. Diria, por isso, que se trata de uma loucura social que se instaura e nos in-forma (colocanos na forma de).

Portanto, a interpretação de Agamben sobre esse caso, desse processo de K., traz consequências. Afinal, ele diz que *cada homem* (ou seja, abstraindo-se da obra de Kafka) se autocalunia. Seria isso verdade? É isso que de fato ocorre? Os homens se acusam falsamente?

A questão da culpa

Segundo Agamben, quem se autocalunia põe a si como causa de algo: o que está em causa/ "a coisa" em causa está sendo julgada. Colocar algo em causa por estar sendo julgado: o acusar e o se submeter à causa – como se o sujeito acusasse a si de uma culpa que não sabe qual é. Diz Kafka: "o processo calunioso é uma causa na qual nada há que esteja em causa, em que é a própria causa a ser posta em causa, isto é a acusação como tal. E onde a culpa consiste no dar início ao processo, a sentença não pode ser senão o próprio processo" (KAFKA, 1997, p. 35). Sentencia-se alguém sem motivo declarado, que não é dito, que *não é anunciável* por nenhuma instância, por alguém a quem não se tem acesso, para que haja uma culpa qualquer que impere e faça com que o sujeito acredite *no que está sendo dito* à sua revelia e que, de fato, muda seu ser – sua forma-de-ser.

Ter uma forma-de-ser e ser con-

ISSN: 2317-9570

denado por ser dessa forma (naturalmente) – parece que não há defesa contra isso (ter uma natureza culpada, como biblicamente se é culpado desde Adão), "uma vez que o direito responde transformando em crime o próprio pôr em causa e fazendo da autocalunia o seu fundamento" (AGAMBEN, 2010, p. 36), diz Agamben. Diz também que, quando o direito se torna uma questão metafísica, em que está em questão o próprio ser, se chega a isso: "o direito não só pronuncia a condenação no preciso momento em que reconhece o infundado da acusação, como transforma também o subterfúgio do autocaluniador na sua própria justificação eterna" (AGAMBEN, 2010, p. 36). Na verdade, Agamben parece oscilar nesse momento, porque também diz de um destino trágico do homem (e não cômico): "e ainda que o homem fosse sempre inocente, se nenhum homem em geral pudesse ser declarado culpado, subsistiria sempre como pecado original a autocalunia, a acusação sem fundamento que ele dirige a si próprio" (AGAMBEN, 2010, p. 36).

Bem, é isso que fazemos conosco? Sofre-se por que não sabemos mais qual é a acusação e porque se sente acusado de algo?

Em uma de suas conferências, mais especificamente, na conferência 21 (*O desenvolvimento da libido e as organizações sexuais* (1916)), Sigmund Freud diz:

> e compreende a voz do poeta como se lhe dissesse: 'Tu te revoltas em vão contra a tua responsabilidade, e proclamas o que fizeste contra tais intenções criminosas. Mas és, sim, culpado, pois não foste capaz de aniquilá-los: eles permanecem ainda, inconscientes, dentro de ti'. E aí está a verdade psicológica. Ainda que o homem reprima seus impulsos maus, banindo-os para o inconsciente, e queira então dizer que não é responsável por eles, ele é obrigado a sentir essa responsabilidade na forma de um sentimento de culpa cujo fundamento desconhece (FREUD, 2014a, pp. 440-441).

Que culpa é essa? Trata-se de algo que se fez e se arrepende? Uma culpa por ter desejado algo que socialmente não é bem-vindo?

O trágico em Freud é que o homem tem um destino e o realiza culpando a si – um destino porque foi estruturado, socialmente, a agir de tal forma: a ser culpado. Mas Freud indica que a culpa vem de algum desejo reprimido socialmente – eis o preço de sermos seres sociais. Mas e se não houve nem mesmo um desejo que pudesse ser reprimido? E se a culpa for instaurada sem nem mesmo desejarmos algo objetivamente? E se somos julgados por algo que nem sabemos o que é?

Franz Kafka parece sugerir isso em sua obra *O processo*. Uma coisa é se sentir culpado por ter desejado algo (que podemos confessar); outra, por nem saber do que se trata. O trágico, num primeiro olhar, está no primeiro caso: é inevitável que, sendo um sujeito, por ser reconhecido socialmente, é-se culpado de algo; o cômico é quando sou culpado por algo sem saber do que se trata e, por isso, possível de ser redimido, uma vez que não se tem uma acusação explícita de um ato que o sujeito seria culpado – uma forma de resumir o argumento de Agamben.

Freud nos encaminha a uma ideia de confissão em muitos momentos de sua obra – durante a interpretação de uma fala parece ser necessário *confessar* que era isto que se desejou (um saber que não se sabe). Essa ideia de confissão fica ainda mais clara quando Freud, em 1925, escreve seu famoso texto *A negação*. Sobre a questão da negação, logo no começo escreve: "negar algo num juízo é dizer, no fundo: 'Isso é algo que eu gostaria de reprimir'. O juízo negativo é o substituto intelectual da repressão, seu 'Não' é um sinal distintivo, seu certificado de origem, como '*Made in Germany*', digamos" (FREUD, 2011, p. 278). No parágrafo seguinte já descreve o método analítico como uma tentativa de fazer com que o analisando confesse algo. Que se veja:

> às vezes é possível obter de forma cômoda o esclarecimento que buscamos acerca do material reprimido inconsciente. Perguntamos o seguinte: 'O que você considera o mais improvável naquela situação? O que acha que estava mais distante de sua mente então?'. Se o paciente cai na armadilha e fala aquilo em que menos pode acreditar, quase sempre está confessando a coisa certa (FREUD, 2011, p. 276).

Com este tipo de passagem clássica, Freud nos diz que a confissão de nossos desejos se dá pela negativa. Ao mesmo tempo, indica como é tão desprazeroso anunciar, em voz alta, o que desejamos. O fato é que desejamos algo mesmo sem anunciar. Eis o trágico: seguimos algo que, se anunciarmos, jamais nos perdo-

ISSN: 2317-9570

aríamos. É o que escreve em 1915 em *A repressão*. A pergunta é de Freud: "por que deveria um impulso instintual [pulsão] sucumbir a esse destino [repressão/recalque]"? (FREUD, 2010, p. 83).

Podemos nos remeter ao tema da tragédia desde Aristóteles, quando ele observa, na *Poética*, que uma tragédia é mais bela quando ela cumpre um propósito (como se o que poderia ser visto como um "acaso", no fundo, seja simplesmente o cumprimento de um destino) (*cf.* ARISTÓTELES, 1997, p. 29). Uma boa peça trágica seria, assim, exatamente o encontro com o destino. Um destino *anunciado*.

O acaso parece aqui perder sua força, porque *seria acaso somente para alguém que não quer "ouvir" o que já se sabe*. Lembrando que *acaso* em Aristóteles é aquilo que foge do necessário. Explico-me melhor: necessário é tudo aquilo que é e só pode ser dessa forma; contingente é aquilo que é necessário que seja, mas que pode ser de uma forma ou de outra; acaso é aquilo que não é necessário que seja. Por essa perspectiva, o trágico, talvez, seja cumprir sem saber o que já estava escrito (algo necessário, mas contingente a forma de sua realização). O que movimenta a tragédia é que sabemos que os fatos previstos irão acontecer independentes da vontade da nossa razão consciente – morrer é uma dessas tragédias.

Por outro lado, Giorgio Agamben, uma vez, em *Profanações*, em 2005, no texto *Desejar*, meio embasbacado com isto, se questiona como é possível que algo tão humano como o desejar se passe como inconfessável. Afinal, por que esconder algo que se deseja? (AGAMBEN, 2007, p. 49). Agamben parece aqui se questionar sobre o problema da confissão: por que é tão embaraçoso dizer o que se deseja? A resposta freudiana é certeira: há uma censura no contato entre o que se deseja e o que a consciência se permite saber. Mas o sentimento de culpa seria somente isso? Não se sabe mais do que isso?

Agamben nos lembra uma ideia de cripta: de uma imagem do que fazemos do inconfessável. E isto nos leva a uma "servidão voluntária" da consciência – uma forma de dizer que se é prisioneiro do seu próprio desejo: "é que o desejo inconfessado somos nós mesmos, para sempre prisioneiros na cripta" (AGAMBEN, 2007, p. 49). Há aqui uma ideia de que não quero dizer o que desejo e mesmo que não sei saber dizer em palavras o que desejo. A meu ver, nada mais freudiano. Mas a ideia de *confissão* fica em suspenso, porque Freud parece também oscilar com essa ideia.

Se seguirmos Freud, poderíamos dizer que nossos desejos são inconfessáveis porque não queremos saber sobre eles. É algo estranho porque não nomeável, não colocado em jogo. Entretanto, algo que reconhecemos tacitamente diante desta tragédia como o mais *familiar*.

Voltemos à tragédia. Édipo, enfim, cumpre seu destino. Aliás, cumpre-o tentando evitá-lo. O oráculo lhe *prenuncia* o espetáculo: irás matar teu pai e casar com tua mãe. Diante desse horror, Édipo foge. Mas foge para cumprir o que disse o oráculo. E, ao perceber o que ele fez, *arranca seus olhos*, literalmente, para não ver ou *tentar* desesperadamente não ver o que já sabia de algum modo. Há aqui, portanto, uma figura trágica do destino. Isto leva Jacques Lacan, por exemplo, a afirmar em seu *Seminário VIII – A transferência* (1960-1961) que o analisando busca a análise e se depara com "(...) o tropo dos tropos, o que denominamos seu destino" (LACAN, 1991, p. 372). Para ficar nele? Para o vingar?

Sei o quanto é contestada esta questão da confissão no meio analítico, principalmente depois da crítica de Michel Foucault (em a *História da sexualidade – A vontade de saber* (1976)) ter destacado um mecanismo de poder no método interpretativo de Freud. A interpretação de Foucault pode ser contestada se lembrarmos, por exemplo, dessa passagem de Freud (*A questão da análise leiga: diálogo com um interlocutor imparcial* (1926)) em que ele está discutindo com alguém que o acusaria por isso:

> 'entendo', diz nosso ouvinte imparcial. 'Você' supõe que todo neurótico tem algo que o oprime, um segredo, e, ao fazer com que ele o exponha, livra-o do peso e o faz sentir-se bem. Mas esse é o princípio da confissão, desde sempre utilizado pela Igreja católica para assegurar seu domínio sobre os espíritos'. Sim e não, [é a resposta de Freud a esse interlocutor imaginário] devemos responder. A confissão é parte da análise, como uma introdução, digamos. Mas está longe de constituir a essência da análise ou de explicar seu efeito. Na confissão o pecador diz o que sabe; na análise o neurótico deve dizer mais. E não temos notícia de que a confissão tenha desenvolvido a capacidade de eliminar sintomas patológicos (FREUD, 2014b, pp.

132-133).

Que se veja nessa outra passagem em que Freud diz:

> (...) a situação fica intolerável quando, de repente, ele [uma pessoa qualquer] não consegue afastar a ideia de que empurrou uma criança para debaixo das rodas de um carro, de que jogou um desconhecido na água, do alto de uma ponte, ou quando chega a se perguntar se não é o homicida que a polícia busca, como autor de um crime recém-descoberto. É um evidente absurdo, ele bem sabe, nunca fez mal a ninguém; mas, ainda que fosse realmente o assassino procurado, a sensação – o sentimento de culpa – não poderia ser mais forte (FREUD, 2014b, p. 128).

O fato de ter lembrado de Foucault aqui não é para confirmar ou negar a sua crítica a Freud. Foucault faz uma ligação entre poder, saber e sexualidade – uma forma de mostrar que há uma função de normalizar o homem e que ele deve confessar. Assim, falar sobre o sexo socialmente seria uma espécie de transgressão deliberada e não uma negação particular de algo. Se é assim, por que há censura? Por que há culpa?

Foucault nos mostra como não se trata somente de uma ação de intolerância coletiva em relação ao sexo, mas também uma questão judicial, médica, pedagógica etc. que exige uma elaboração teórica sobre a questão, devido a um dispositivo de controle do próprio discurso. Haveria, portanto, uma exigência confessional para que cada um se diga desejando algo muito bem claro e, por isso, de possível controle. Dito de outro modo, faz-se com que os segredos sejam, no fundo, controláveis e acessíveis ao dito (sem margem ao não-dito). A questão aqui é mais forte em relação ao poder de "analisar a formação de um certo tipo de saber sobre o sexo, em termos não de repressão ou de lei, mas de poder" (FOUCAULT, 2005, p. 121). Não parece ser isso que está em jogo quando lemos *O processo* e ao fazermos dessa obra uma constelação para pensarmos o social.

Talvez o interesse nesta discussão é mostrar uma contra ideia a este confessionário. Com Giorgio Agamben, podemos pensar que se é culpado mesmo não desejando (ou seja, não há o que confessar). A figura central aqui é o personagem K. Entretanto, Agamben, ao invés de tratar da *tragédia*, irá tratar da *comédia*. Há, portanto, uma mudança de centro da forma de se dispor a tonalidade da qual o sujeito lida com seu próprio destino. Tento articular esta questão – e, de uma forma aberta a outras leituras.

Agamben faz uma diferenciação entre autocalunia e confissão. Segundo sua leitura, o processo de K. visa que ele confesse algo e que ele faça uma autocondenação – uma acusação de si que seria *falsa* – algo somente que se diz sem ter uma crença na própria fala. Ter crença em algo não significa que esse algo seja real – significa simplesmente que você acredita nessa fala. Mas uma fala que não anuncia uma confissão de algo que se tenha realmente sentido ou vivido – uma fala que não tem a força de verificabilidade, mas simplesmente a instauração da fantasia social que se passa como real.

De fato, K. não faz uma confissão. K. não tem o que confessar em relação ao seu processo. Se tem algo estranho ao que se vê do processo de K. é o que Agamben mesmo diz sobre a confissão: "à medida que a prática da tortura se difunde, a confissão interioriza-se e, como a verdade arrancada à força pelo algoz, torna-se qualquer coisa que o sujeito é coagido pela sua consciência a declarar espontaneamente" (AGAMBEN, 2010, p. 38) – K. não declara nada espontaneamente, porque nada daquilo parece real – se passa *como real*. Parece tudo uma loucura que ele ainda não entendeu e está sendo introduzido – introduzido à loucura que todos acreditam e ele vive a sua acusação na carne. Na verdade, nem mesmo interessa a sentença, mas a tortura do processo. *Uma tortura, porque não há nada a confessar*.

O processo de aceitação do processo

Do que K. é acusado? Por que ele está sendo detido em sua casa?

A primeira resposta (sobre o porquê de ser detido) que K. recebe é de um homem, cujo nome é Franz, junto à janela do seu quarto, que revela a ele o que está ocorrendo: "não fomos incumbidos de dizê-lo. Vá para o seu quarto e espere. O procedimento acaba de ser iniciado e o senhor ficará sabendo de tudo no devido tempo" (KAFKA, 1997, p. 11). Há uma anunciação de algo aqui. De quê?

Não se pode dizer que K. está realizando uma calúnia de si neste momento, mesmo porque ele não esperava aquilo – estava de pijama e se sente incomodado por não seguir sua rotina devido a estranhos fazendo acusações que ele não tem/tinha a menor ideia do que seja. Não é K. que se autojulga neste momento, ao menos a meu ver, – ele está desorientado e é Franz que o "acalma" com essa fala: "o senhor ainda vai perceber como tudo isso é verdade" (KAFKA, 1997, p. 12). A questão é que o "tudo" não é anunciado...

As primeiras reações de K. são sempre as mesmas: deseja ter clareza da situação – ele não sabe o que está acontecendo e não entende a lógica do acontecimento. Diz K. a um dos guardas: "como posso estar detido? E desse modo?"; resposta: "lá vem o senhor de novo – disse o guarda, mergulhando um pão com manteiga no potinho de mel. – Não respondemos a perguntas como esta" (KAFKA, 1997, p. 14). E na insistência de K., respondem: "oh céus! – disse o guarda. – É incrível como o senhor não consegue se submeter à sua situação e parece empenhado em nos irritar inutilmente, a nós, que decerto somos neste momento os mais próximos de todos os seus semelhantes!" (KAFKA, 1997, pp. 14-15). Que se veja: K. não se submete à situação, como destaca o guarda. Ele só vai se submeter quando se convencer que todos estão envolvidos no "caso" e é só ele que ignora/ignorava a presença de um tribunal – nada de autocalúnia por enquanto.

Entretanto, os mesmos guardas dão uma dica do que está acontecendo: "aqui não há erro. Nossas autoridades, até onde as conheço, e só conheço seus níveis mais baixos, não buscam a culpa na população, mas, conforme consta na lei, são atraídas pela culpa e precisam nos enviar – a nós, guardas. Esta é a lei. Onde aí haveria erro?" (KAFKA, 1997, p. 15). Nesse momento fica ambíguo: é K. que sente e atrai a lei ou a lei que sente que K. poderá se sentir culpado? É um mistério. Trágico ou cômico?

Torna-se mais misteriosa a afirmação da lei e a força da culpa quando K. conversa pela primeira vez com alguém superior aos guarda: "tiro essa conclusão [de o caso não ter tanta importância] do fato de ser acusado e não conseguir descobrir a mínima culpa da qual me pudessem acusar. Isso também é secundário, a questão principal é: por quem sou acusado? Que autoridade conduz o processo?" (KAFKA, 1997, p. 21). Não me parece que aqui também haja autocalúnia, mas uma *imposição* de uma culpa que o sujeito deve passar. É nesse sentido que trata-se mais, a meu ver, de uma patologia social – de uma forma de inculcar uma culpa no sujeito sem que ele tenha ideia do que está em jogo.

Por exemplo, o inspetor diz ser um personagem secundário e quase nada saber do processo. Foi a K. que foi exigido que se vestisse bem para se apresentar ao inspetor. Surpreendentemente, o inspetor está malvestido! Isso nos diz que a lei não precisa de formalidades, mas somente, ser cumprida, seja ela qual for; o acusado, por sua vez, deve cumprir as formalidades e se apresentar com toda formalidade *diante* da lei.

Há também uma ideia de que o processo é importante e que K. não deve se distrair com qualquer coisa fora da lei – eis, ao meu ver, uma imposição! Um exemplo é o conselho de um dos guardas e do inspetor que parece sem sentido desde o começo: "nós recomendamos não se distrair com pensamentos inúteis, mas se concentrar, pois grandes exigências serão apresentadas ao senhor" (KAFKA, 1997, p. 16). Parece-me que aqui está se instaurando uma loucura: "(...) posso entretanto aconselhar o senhor a pensar menos em nós e no que vai acontecer e mais em si mesmo" (KAFKA, 1997, p. 22). Como se vê, impõe-se uma forma de ser sem que aja qualquer explicação do porquê. É isso que estou chamando de loucura: quando toda referência à verdade não tem qualquer forma de verificabilidade e tal é tratado *como se fosse* algo inquestionável. Afinal, a composição inicial de *O processo* é mais ou menos assim: uma fala que se impõe a K.: "não diga que é inocente – baste reparar melhor em seus atos e verás que é culpado". Pergunto-me: a lei farejou uma culpa ou está instaurando/gestando uma?

No começo do processo, vemos que K. tenta conciliar-se com o inspetor e os guardas, mas estes dizem que não há conciliação possível; dizem também que não é o caso de se desesperar... *é só um processo!* Eis como a detenção de um sujeito pela lei é absolutamente ambígua: "o senhor me entendeu mal. É claro que o senhor está detido, mas isso não deve impedi-lo de exercer sua profissão. Tampouco deve ficar tolhido no seu modo de vida habitual" (KAFKA, 1997, p. 25) – *somente um processo!*

A confusão mental de K. é ter confirmações desse processo por pessoas inesperadas. Por exemplo, a partir da senhoria Grubach (dona da pensão) que lhe diz: "de fato o senhor está detido, mas não como um ladrão, então é ruim, mas este tipo de detenção... A mim me parece algo de sábio, desculpe-me se estou dizendo uma tolice, a impressão que eu tenho é de algo sábio, que não entendo, mas que também não é preciso entender" (KAFKA, 1997, p. 31). A senhoria diz tudo sobre o processo... parece algo tolo e ao mesmo tempo sábio.

Mas o próprio fato de não se ser culpado parece também se con-

ISSN: 2317-9570

tradizer. Por exemplo: K. sabe da vida "suja" da vizinha. Mas em uma conversa com a senhoria sobre essa vizinha, diz: "se quer conservar limpa a pensão, precisa primeiro me despejar" (KAFKA, 1997, p. 34). Aqui poderíamos dizer que estamos diante de uma autocalúnia. Aliás, K. chegou até em pensar em castigar a senhoria Grubach por ter dito algo sobre sua vizinha – há aqui um jogo de culpa e de quem pode falar em nome da lei. Nos diz também, de algum modo, como K. começa a direcionar a culpa aos outros: a enunciar uma culpa que é contra alguma lei que ele nem sabe qual é.

Começa-se, então, um primeiro momento de reconhecimento do processo:

> (...) só é um processo se eu o reconhecer como tal. Mas neste momento eu o reconheço [senhor juiz no primeiro inquérito], de certa forma por piedade. Não se pode ter outra coisa senão piedade, se se deseja leva-lo em consideração. Não digo que seja um processo desleixado, mas gostaria de lhe oferecer essa definição como forma de autoconhecimento (KAFKA, 1997, p. 56).

Como se percebe, é somente K. que fala sobre esse processo nesse momento – daí a ideia de autocalúnia. Mas seria mesmo uma autocalúnia ou uma forma de fazer K. acreditar no que estão o acusando – um reconhecimento de uma culpa que vem do social?

Nesta altura da obra, em que K. vai ao tribunal para se defender, ele só acusa. Aliás, acusa a todos: diz se tratar de uma acusação que vem de alguma organização estranha que ele não reconhece etc. e se depara com esta questão: "estava agora cara a cara com a multidão. Tinha julgado certo as pessoas? Tinha confiado demais no efeito do seu discurso?" (KAFKA, 1997, pp. 62-63). Afinal, acusa a todos por estarem lhe testando – só ele fala; e afirma sua inocência diante de todos. Não há retruca nem dos ouvintes, nem do juiz, nem de ningém presente no tribunal! Há somente olhares de reprovação (por não confessar sua culpa ou por estar anunciando de forma negada uma culpa?). Ou seja, nada é anunciado "por outro", mas somente "por olhares". "Todos" estamos "de olho" em você? – como um panóptico, talvez?

Depois de um "espetáculo" de K. (no sentido de dizer muito e acusar a todos e receber palmas e reprovações) diante daqueles que estão presentes na "tribuna" – que se tratava do absurdo do processo

–, diz o juiz na tentativa de retirada de K.: "um momento", barrando a passagem de K. para sair. E complementa: "Só queria chamar a sua atenção – disse o juiz – para o fato de que o senhor hoje – isso ainda não deve ter chegado à sua consciência – se privou da vantagem que um inquérito, de qualquer modo, representa para o detido" (KAFKA, 1997, p. 64). A impressão que se dá é: ao culpar todos e não deixar o juiz falar senão ao barrar e anunciar algo na saída, de algum modo, reafirma que a lei atrai a culpa. Ou melhor: a lei ficaria muda até certo ponto, por isso K. ficou na espera de uma ligação: ele sabia que o processo não poderia terminar por ali! Deve ter algo mais importante que o barra – um juiz que não diz nada senão quando o acusado tenta sair de "juízo".

Sem ser convocado, K. volta no outro domingo ao tribunal, mesmo não sabendo se tem ou não uma audiência. Aqui se mostra um compromisso com a culpa ou um "passar" a acreditar que existe uma culpa ou que não há culpa alguma e por isso se deve estar pronto para se defender? A mulher que o recebe mostra o lugar vazio da lei: *onde se senta o juiz*. Neste lugar vazio, K. não quer perder a oportunidade de ver os livros que estão em cima da mesa. Vê ali códigos que condenam só quem é "inocente" – mas os códigos apareciam como temas obscenos sem qualquer relação com o que se espera da lei. E avalia: "são estes os códigos de lei estudados aqui – disse K. –, é por homens assim que devo ser julgado"? (KAFKA, 1997, pp. 68-69) – homens que colocam em cima da mesa essa pornografia que consideram a lei?

Talvez a grande virada do livro é quando K. começa a acreditar que ele deve ser processado por algo, mesmo que seja por uma lei promíscua que nem anuncia a sua acusação; diz, afinal, mesmo que ironicamente, "o meu processo" (KAFKA, 1997, p. 70) – passa agora do "deve haver um processo" a "há um processo"; "não reconheço o processo" a "enfim, há um processo".

A loucura do processo começa a se tornar real, porque todos sabem deste processo e melhor que o próprio K.! A mulher que o recebe no tribunal num domingo sem audiência diz a ele, por exemplo: "com tudo isso eu queria apenas dizer que o juiz de instrução escreve efetivamente muitos relatórios, principalmente sobre o senhor, pois o seu inquérito foi na certa um dos principais objetos da audiência domingo" (KAFKA, 1997, p. 72). Afinal, a loucura parece estabelecida: tem-se muito a falar de algo que nem foi pronun-

ISSN: 2317-9570

ciado pela voz de um juiz – e quem pode falar em nome da lei? O que se tem é uma escrita infindável em livros estranhos, com códigos pornográficos. Livros que contêm acusações de algo a alguém sem anunciar a própria acusação.

Eis um exemplo completamente desconcertante: em uma conversa de K. com um oficial de justiça, este lhe pede um favor estranho e imprevisto e se justifica dizendo que K. é acusado e não pode alegar nada (*cf.* KAFKA, 1997, p. 80) – uma imposição de uma possibilidade de culpa e da assunção da acusação?

Mas K. não está entendendo o processo e nem o porquê deste acontecimento – ele simplesmente está em um processo. Vê seus companheiros, outros que estão sendo acusados, trabalhando domingo no tribunal e agindo de forma humilhante. Ao perguntar o que estão esperando ali, respondem que estão esperando o resultado de sua causa. Ou seja, há outros como K. e mais engajados em seus processos – não consigo imaginar uma descrição mais clara de uma patologia social. O que isso significa?

Podemos pensar a patologia social fazendo uma certa analogia com a patologia individual. Não temos dúvida quando dizemos que um indivíduo sofre de uma dada patologia. Mas e quando dizemos que há uma patologia social? Por analogia, podemos dizer que o social como um todo sofre de uma patologia. Em *O processo*, todos agem *como se fosse natural* que haja um processo sem que seja explicitada a acusação, levando cada indivíduo acusado a acreditar nessa acusação. É o social, como um todo, que parece estar enlouquecendo e não um indivíduo particular. É o social que leva K. a crer em seu processo, anulando sua capacidade de duvidar da legitimidade ou não do processo, como veremos. Daí porque podemos dizer que há uma patologia social (e não de K.).

Assim, tudo parece indicar uma lei com a qual K. está apreendendo a lidar. Por exemplo, há um presságio da lei declarado a K.: "a punição é não só justa como inevitável" (KAFKA, 1997, p. 107). Impossível não pensar em *destino* aqui, pois a culpa é "natural". Mas o estranho é que se aceite o processo, como se se aceitasse uma "loucura social". Afinal, K. aceita a lógica desse processo?

Parece que sim. Por exemplo, quando é "levado" a presenciar o espancamento de dois guardas denunciados na primeira inquisição. Neste momento, K. faz esta pergunta maluca: "não existe nenhuma possibilidade de poupar os dois do espancamento?" (KAFKA, 1997, p. 107). Ou seja, ele não disse que o acontecimento é absurdo, mas se era ou não necessário aquilo... Eis a instauração da loucura social sobre o sujeito. Assim, "atormentava-o não ter conseguido impedir o espancamento, mas não era culpa sua o fato de não tê-lo conseguido" (KAFKA, 1997, p. 111).

O processo continua. Ao conversar com seu tio, a única adjetivação que até agora encontramos do processo aparece: "é um processo criminal" (KAFKA, 1997, p. 117). E o tio perguntou de que processo se tratava? A resposta é que *não se trata de um processo de um tribunal comum*. Só isto satisfez o tio que já começou a especular uma resolução do caso que parece ter ficado claro com a resposta de K. – *parece normal* ser processado por uma lei que não diz seu nome.

Doravante o processo tem poder sobre K. O tio o aconselha a acompanhá-lo ao campo. Mas K. responde: "Eles poderiam me proibir de viajar" (KAFKA, 1997, p. 120) – o processo já tem uma força de lei que o tio reforça:

> (...) [você, K.,] sempre teve uma capacidade de compreensão tão correta, e logo agora ela o abandona? Quer perder o processo? Sabe o que isso significa? Significa que vai ser simplesmente riscado do mapa. E que todos os parentes também serão arrastados, ou pelo menos humilhados até o chão. Josef, concentre-se. Sua indiferença me tira do sério. Quando se olha para você, quase que se acredita no ditado: 'Ter um processo desses já significa tê-lo perdido' (KAFKA, 1997, p. 121).

Quer dizer, K. tem agora várias referências do processo em sua "realidade" – vindo inclusive de seu tio, pois hora nenhuma alguém colocou em dúvida esse processo, somente K.! E por isto mesmo, por só ele colocar tal processo em dúvida, ele parece ser *o sujeito mais sóbrio nesse meio social*! Aquele que pode negar os outros personagens ("kein"). Ele é o único que não se entregou completamente a tal processo, porque não o entende, mas que passa, ao mesmo tempo, por um "processo de entrega" em uma crença: um processo de "domesticação" de si que ele tenta negar. Por exemplo, com o convite do tio a ir ao campo. Nesse momento, K. responde algo assim: apesar de vantajosa a proposta, não se poderia aceitar o convite, pois isso significa fuga e consciência de culpa.

K. entra num processo que o insulta, o indigna, o faz entrar em um jogo que não entende as re-

ISSN: 2317-9570

gras. Que se veja a perplexidade de K. ao consultar com seu tio um advogado que vai representá-lo nesse processo: o advogado já sabia da causa, do processo, mesmo estando doente e de cama. Cada detalhe vai fazendo com que o processo ganhe mais relevância. Afinal, todos estão envolvidos *no* processo. É isto o enlouquecedor! – como se o social soubesse mais sobre o seu processo do que você mesmo. Todos sabem mais do que ele sobre seu caso, mas sabem algo que não é contável, ou inacessível. Mesmo a enfermeira de seu advogado tem algum conselho para dar:

> por favor, não pergunte nomes, mas corrija os seus erros, não seja mais tão inflexível, contra esse tribunal não é possível se defender, é preciso fazer uma confissão. Na próxima oportunidade, faça essa confissão. Só aí existe a possibilidade de escapar – só aí. No entanto, mesmo isso não é possível sem ajuda externa, mas não precisa se angustiar por causa dessa ajuda, eu mesma vou providenciá-la (KAFKA, 1997, p. 135).

Todos parecem saber mais do que K. sobre seu processo e o incentivam a aceitá-lo – não é ele que está tentando convencer os outros que haja um processo ou que haja alguma culpa; K. somente está entrando aos trancos e barrancos num jogo cuja regra desconhece. Se há um processo, algo inevitável, que todos aceitam, "deve haver" uma culpa. A partir do capítulo sétimo (*O advogado. O industrial. O pintor*) parece que K. entra completamente nesse jogo. Ele agora começa a julgar a si de desleixo com o processo e, "numa manhã de inverno (...) não conseguia mais deixar de pensar no processo" (KAFKA, 1997, p. 140). Que se veja:

> não conseguia mais deixar de pensar no processo. Já tinha refletido com frequência se não seria bom redigir um documento de defesa e apresenta-lo ao tribunal. Queria expor nele um breve relato de vida e, a propósito de cada acontecimento relevante, explicar os motivos pelos quais tinha agido daquela forma, se esse comportamento devia ser censurado ou aprovado segundo o seu juízo atual, e que razões podia invocar em relação a este ou aquele (KAFKA, 1997, p. 140).

Ou seja, passado um tempo, em uma manhã de inverno, K. está tão crente de seu processo que começa a se censurar.

Começar a se censurar... seria uma autocalúnia como sugere Agamben?

A passagem para esse capítulo sétimo é tão brusca que, das dúvidas e insatisfações que K. sentia, isto é, depois da conversa com o tio, da visita ao advogado e da sugestão da enfermeira, parece que tudo muda: *não há mais dúvida*. K. se diz culpado. A questão agora é saber se se trata de um pecado pessoal ou social. Tratar-se-ia de uma autocalúnia ou da instauração de uma lei que seja partilhada e que não faz senão dizer que você é culpado? Nesta última possibilidade, K. seria um homem que poderia se autocaluniar devido à imposição de uma lei que procura uma culpa.

Segundo Freud, em *A questão da análise leiga*, o Super-eu pune o sujeito de uma forma que teríamos que supor que há uma necessidade de sentimento de culpa; mais especificamente: como se o sujeito *necessitasse de uma doença de punição* (*cf.* KAFKA, 1997, p. 180). Assim, a lei encontra esses que necessitam dessa doença, uma vez que o sujeito teria um ganho na doença, com esse sentimento de culpa inconsciente. O que Kafka parece descrever é esse ganho na doença, mas numa doença que é imposta socialmente, que faz com que K. faça uma passagem da dúvida à certeza sobre o processo... Com o tempo, convence-se que se deve prestar mais atenção em sua acusação, porque todos dizem que é assim e que deve ser assim... Ou ainda outra possibilidade: é como se a lei buscasse culpar um sujeito que não se sentia culpado. Seria essa a violência do processo? – fazer com que o homem se culpe sem ter culpa "pessoal" de ter feito algo? Se esse for o caso, então esse é um ato da lei e não algo espontâneo do sujeito – como se a lei procurasse uma culpa que fosse "natural" ao sujeito.

Ao procurar uma culpa sem qualquer acusação, a lei tem força de lei porque "todos" (um anonimato em geral) sabem de sua força, menos o próprio acusado. Eis porque se trata de uma força a meu ver: porque mesmo "os funcionários são em muitos aspectos como crianças" diante da lei (KAFKA, 1997, p. 149) – apenas seguem o que se prescreve sem saber o porquê. A questão está na ideia de que "eles" sabem que se é culpado de algo que não se sabe o que é. Neste sentido, K. se passa como o único herói da história – porque é o único que não entrou no jogo de "saber" o que não se é sabível; o único que não recaia em uma lei insana que nada anuncia, simplesmente exigindo uma culpa e não uma confissão – porque não há o que confessar. Ele é um herói por duvidar. K. é esse que pode formar a negação do artigo indeterminado e que pode negar um nome qualquer. Mas um herói que decai, porque acaba acreditando na lei num certo momento.

Uma outra genialidade da obra é fazer do leitor um cumplice do processo. Em certo momento, nos flagramos acreditando nesse processo, sem saber que tipo de petições estão sendo colocadas a algum juiz que nem sabemos quem é. Começam a surgir petições que são enviadas do advogado ao tribunal e que jamais são lidas. Surge um réu que se defende por alguma representação de alguém que não diz seu nome, mas que está presente em forma de lei: "os documentos do tribunal, sobretudo o auto de acusação, permaneciam inacessíveis ao acusado e à sua defesa, por isso geralmente não se sabia, ou pelo menos não se sabia com precisão, contra o que a primeira petição precisava se dirigir, de forma que só por acaso ela continha, em verdade, algo relevante para a causa" (KAFKA, 1997, p. 142). O que a lei determina tem força de lei, mas sua palavra não é pública – nem mesmo o acusado deve saber sobre o que se trata. Um dos acusados que K. conheceu diz que o conteúdo de uma petição é de tal forma escrita: "cheguei até a ler pessoalmente uma delas [petições] por deferência de um funcionário do tribunal. Ela era de fato erudita, mas o seu conteúdo na realidade nulo" (KAFKA, 1997, p. 216).

Assim, o advogado é somente admitido ("tolerado") – não mais do que isso. Uma espécie de mistério que recai sobre o acusado: todos "sabem" da acusação, mas nada sobre o que *diz* a lei. E, por isso, a sugestão estranha de Agamben: trata-se de uma autocalúnia, uma vez que, ao não ter uma acusação clara, se espera que o acusado se auto acuse!

Há uma longa descrição de uma falta de lógica do trâmite de qualquer processo, assim como as vantagens que se pode ganhar ao usar essa misteriosa representação da força de lei que não se coaduna com o sentido da lógica tradicional. Afinal, nesta lógica estranha, mesmo os funcionários da "lei" sofrem, porque não sabem exatamente o que estão fazendo. Há uma lógica, mas uma lógica que não sabemos exatamente qual é, a não ser por metáforas.

No fim, o que interessa é que o processo está em andamento... pouco importa a acusação. Mas, curiosamente, é importante ter boas relações com os acusadores: com os funcionários em geral, para tentar se mostrar solícito (disponível e adepto à lei). Ao aceitar o processo, K. se coloca de forma favorável para que o processo ande.

A loucura da força de lei parece agora não ter mais fim... Em uma conversa com os advogados, K. entende que "sempre havia progressos, mas nunca se podia informar de que tipo eram" (KAFKA, 1997, p. 152). Uma loucura instaurada: o importante é que haja "progresso", mesmo sem saber qual. Aliás, nenhum processo surge se não tiver outros que acreditam no processo: "o desprezo que antes manifestava pelo processo já não era válido. Se estivesse sozinho no mundo, poderia com facilidade não levá-lo em conta, embora fosse certo, nesse caso, que o processo simplesmente não teria surgido" (KAFKA, 1997, p. 154). Tudo muda quando é o "todos" que está em jogo: "– em suma, ele [K.] praticamente não tinha mais a escolha de aceitar ou rejeitar o processo, estava no meio dele e precisava se defender. Se estava cansado, isso era ruim" (KAFKA, 1997, p. 154). Isso mostra que houve uma mudança durante os meses de processo: K. não é o mesmo! Ele agora está, de fato, envolvido *no* processo.

Daí porque discordo da leitura de Agamben, pois mesmo aceitando a culpa, lemos:

> se quisesse conseguir alguma coisa era necessário, acima de tudo, repelir previamente qualquer ideia de uma possível culpa. Não havia culpa. O processo não era nada senão um grande negócio, como os que ele já havia fechado com vantagem para o banco; um negócio no interior do qual, conforme a regra, espreitavam diversos perigos que tinham de ser conjurados. Para esse objetivo, entretanto, não se podia jogar com os pensamentos a respeito de alguma culpa, mas sim com o pensamento de se ater o mais possível ao próprio interesse (KAFKA, 1997, pp. 154-155).

K. não pode confessar, mas não tem o que confessar. K. não pode se sentir culpado, porque não há culpa. K. se calunia? Ou ele é imposto a se autocaluniar já que todos "sabem"? A questão vira outra: como faço para ter acesso à lei? E que lei é essa? O que está em jogo aqui? Ora, "hoje K. não conhecia mais essa vergonha ["de não saber"]: a petição tinha de ser feita" (KAFKA, 1997, p. 156). É como se ele tivesse entrado no jogo que ele desconhecia, porque já não se pode igno-

ISSN: 2317-9570

rar o processo que está em andamento, fazendo-o ter um sentimento vago: "permaneceu sentado assim longamente, sem saber o que de fato o preocupava" (KAFKA, 1997, p. 161). Não penso que seja uma autocalúnia, mas uma questão em relação à crença, uma *aceitação* de algo que ele desconhecia. Neste caso, um vazio de determinação; mas uma força de lei, que faz K. assumir uma defesa. Daí sua preocupação estranha em seu trabalho: pedir férias? Como? Já que "(...) tratava-se, na verdade, de todo um processo, cuja duração era imprevisível" (KAFKA, 1997, p. 162).

A loucura do processo vai tomando forma – algo que te acusa que não tem uma culpa:

> totalmente desconhecido o processo não era, embora ainda não estivesse muito claro quem sabia dele, e quanto. (...) O que tinha K. então a esperar? Talvez enfraquecesse com essas reflexões sua capacidade de resistência, mas era preciso também não se enganar e ver tudo tão claro quanto possível naquele momento (KAFKA, 1997, p. 162).

O que parece ser descrito é que *a pessoa vai se enlouquecendo com a passagem da dúvida à crença*. Fica tão estranho a certeza que há um processo que qualquer sugestão que possa ajudar K. "deve" ser levada muitíssimo a sério – parar tudo que estava fazendo e direcionar-se ao caso. Eis a loucura estabelecida: "(...) [K.] deixou o banco quase feliz com o fato de poder se dedicar, por algum tempo, mais completamente à sua causa" (KAFKA, 1997, p. 170). Que se veja a precisão: quase completamente – é um caminho de crença numa causa que não existe *senão na fantasia social*. Ou seja, o que está em jogo aqui é: ter êxito na fantasia.

K., num certo momento, entra em contato com um pintor que tem acesso aos que estão no tribunal. O pintor representa um dos senhores que estão no tribunal em um quadro. E K. não perde a oportunidade de perguntar: "o senhor pintou a figura como ela realmente fica no trono?" (KAFKA, 1997, p. 177). Mais ou menos isso: é esse o trono da lei? O mais incrível é a resposta do pintor: "não – disse o pintor. – Não vi a figura nem o trono, tudo é invenção, mas me indicaram o que eu tenho de pintar" (KAFKA, 1997, p. 177). É como se se representasse uma figura que não exista senão na imaginação social, porque não há como ter uma verificabilidade entre o retrato e o "mundo", mas somente uma crença.

Em certo momento, K. se diz inocente, mas sem dizer do quê, porque a pergunta dirigida a ele era somente se se é ou não inocente! Mas sua suposta inocência não o livra do caso. Diz K.: "minha inocência não simplifica o caso (...). Depende de muitas coisas sutis, nas quais o tribunal se perde. Mas no final emerge, de alguma parte onde originariamente não existia nada, uma grande culpa" (KAFKA, 1997, p. 181).

K. vai se convencendo de sua culpa. Lemos: "mas todas [pessoas] concordam em que não são levantadas acusações levianas e que o tribunal, quando acusa, está firmemente convencido da culpa do acusado e só com dificuldade é dissuadido dessa convicção" (KAFKA, 1997, pp. 181-182). A obra vai se desdobrando de uma forma em que só é possível acompanhá-la se se entrar no jogo. Por exemplo, o pintor diz a K.: "essas meninas [crianças] também fazem parte do tribunal. (...) Tudo pertence ao tribunal" (KAFKA, 1997, p. 183). Se este tipo de passagem *se passar como natural*, se até as crianças participam "daquele" tribunal, a loucura está estabelecida..., pois todos estão envolvidos em algo em que há regras e critérios tão secretos que ninguém pode ter acesso a eles – quando "tudo e todos" condenam K.

Quem sabe sobre a lei? O pintor é um bom entendedor dela e um conselheiro, – e diz:

> na lei – de qualquer modo não a li – consta, naturalmente, por um lado, que o inocente é absolvido, mas por outro ali não consta que os juízes podem ser influenciados. Ora, a minha experiência é justamente o contrário. Não sei de nenhuma absolvição real, mas sem dúvida de muitas formas de influência. Claro que é possível que em todos os casos que eu conheci não existisse inocência. Mas não é uma coisa improvável? Em tantos casos, nenhuma inocência sequer? (...) As decisões finais do tribunal não são publicadas, não são acessíveis nem mesmo aos juízes, daí que só se conservaram lendas sobre velhos casos judiciais (KAFKA, 1997, p. 187).

O grande segredo desse tribunal parece ser que não há um juiz visível e plausível, senão a fantasia social que jamais dará a oportunidade de uma absolvição completa do acusado, pois "(...) esse direito só o tem o tribunal

ISSN: 2317-9570

supremo, inteiramente inacessível ao senhor, a mim e a todos nós" (KAFKA, 1997, p. 192).

Mas a certeza ainda não está completamente estabelecida. Para sair do atelier, K. fica assustado e o pintor diz: "são cartórios do tribunal. Não sabia que aqui há cartórios? Eles estão em quase todos os sótãos, por que deveriam faltar logo aqui?" (KAFKA, 1997, p. 199). Em todos lugares alguém te julga... e K. parece não mais se assustar senão por ser "pego" de surpresa:

> K. não se assustou tanto por ter encontrado ali cartórios do tribunal; assustou-se principalmente consigo mesmo, com o seu desconhecimento das coisas do tribunal; parecia-lhe ser uma regra básica do comportamento de um acusado estar sempre preparado, não se deixar nunca colher de surpresa, não olhar desprevenidamente para a direita quando o juiz estava à esquerda, ao seu lado – e era justamente essa a regra fundamental que ele sempre violava (KAFKA, 1997, p. 200).

O caso vai se tornando importante – sua vida pessoal, seu trabalho etc. se tornam desimportantes. Trata-se de um caso importante demais, pois como diz um comerciante que K. conheceu: "quando alguém quer fazer algo pelo seu processo, só pode se ocupar pouco de outras coisas" (KAFKA, 1997, p. 212).

A ideia é que K. vai perdendo aos poucos todo o interesse por tudo o que não seja o processo: é algo que vai cada vez tomando seu tempo, sua energia, seus esforços. Seria isto a autocalúnia que nos diz Agamben?

Mesmo no final do livro, não consigo acompanhar Agamben (que diz que a autocalúnia aparece desde a primeira frase da obra):

> não – disse o sacerdote. Mas temo que vá terminar mal. Consideram-no culpado. Talvez o seu processo não ultrapasse nem mesmo um tribunal de nível inferior. No momento, pelo menos, consideram provada a sua culpa. – Mas eu não sou culpado – disse K. – É um equívoco. Como é que um ser humano pode ser culpado? Aqui somos todos seres humanos, tanto uns como outros. – É verdade – disse o sacerdote. – Mas é assim que os culpados costumam falar (KAFKA, 1997, p. 258).

Trata-se de uma espécie de construção em que o sujeito vai se engajando na crença de um processo, na crença social de que o processo é real, e que vai transformando K. que nem sabe a que tipo de tribunal está servindo. Eis como, na cena do sacerdote com K., o processo é explicado: "a sentença não vem de uma vez, é o processo que se converte aos poucos em veredicto" (KAFKA, 1997, p. 258).

Nessa conversa com o sacerdote, em que encontramos a famosa passagem do sujeito diante da porta da lei, o mais estranho parece ser a interpretação do sacerdote da função do porteiro desta porta: ele não responde nada; venera os superiores; não é subornável; não sabe o que vigia; aceita presentes para dizer que não é indiferente; é simplesmente um cumpridor de deveres.

Diante desse porteiro, é-se culpado? Afinal, por que o acusado "se senta no banquinho do lado da porta e ali permanece durante toda a vida, isso ocorre voluntariamente, a história não fala de coação alguma"? (KAFKA, 1997, p. 267). Se se senta voluntariamente ao lado da porta da lei, eis a possibilidade de uma aceitação de uma lei que passa a imperar no sujeito, mesmo que não seja nada verdadeira a acusação, senão uma necessidade da lei (*cf.* KAFKA, 1997, p. 269). Uma ambiguidade sem fim, porque a lei convoca e nada disso acontece sem que você acredite na acusação: "pertenço pois ao tribunal – disse o sacerdote. – Por que deveria querer alguma coisa de você? O tribunal não quer nada de você. Ele o acolhe quando você vem e o deixa quando você vai" (KAFKA, 1997, p. 271). K. acredita nesse processo e paga sua vida por ele:

> devo então demonstrar que nem sequer o processo de um ano me serviu de lição? Devo acabar como um homem obtuso? Será que podem dizer de mim que no início do processo eu quis terminá-lo e agora, no seu fim, quero reiniciá-lo? Não quero que digam isso. Sou grato por terem me dado como acompanhantes estes senhores semimudos, que não entendem nada, e pelo fato de terem deixado para mim a incumbência de dizer a mim mesmo o que é necessário (KAFKA, 1997, p. 275).

A obra termina com a execução de K. que é arrastado por dois homens gordos que o levam para fora da cidade e o matam com uma faca fina fincada em seu coração – tal como se mata um porco.

Mesmo durante esse arrastar, K. ainda reluta. Mas num certo momento "mantém sua dignidade" e aceita a decisão final desse tribunal. Morre como um cão por homens estranhos que não têm ideia do porquê desse ato. Uma autocalúnia? Não. Uma loucura social que leva o homem a crer em uma culpa que não existe senão na fantasia social. Eis o que denominaria, de fato, uma patologia social.

Afinal, trata-se de algo trágico ou cômico? A tragédia, segundo Agamben, é a realização de um destino inexorável, que não se pode mudar. O cômico seria, por sua vez, a realização de uma espécie de redenção do destino trágico. No caso, pela "purificação" da lei. Difícil a questão, porque até mesmo o que consideramos sem qualquer lógica deveria ser lido sob a lente cômica – basta ver como se chega um momento em que o sujeito não mais questiona a lei e nem mesmo como aquilo foi possível. Uma redenção pela lei (algo *cômico*) ou o cumprimento de um destino que vem de uma loucura social (sem qualquer redenção possível, pois se paga *tragicamente* com a própria vida)? Poderíamos pensar em algo tragicômico?

Referências

AGAMBEN, Giorgio. *Categorias italianas – Estudos de poética e literatura*. Trad. Carlos Eduardo Schmidt Capela e Vinícius Nicastro Honesko. Florianópolis: Editoda UFSC, 2014.

______. *Nudez*. Trad. Miguel Serras Pereira. Lisboa: Relógio D'Água, 2010.

______. *Profanações*. Trad. Selvino Assmann. São Paulo: Boitempo, 2007.

ARISTÓTELES; HORÁCIO; LONGINO. *A poética clássica*. Trad. Jaime Bruna. São Paulo: Cultrix, 1997.

FREUD, Sigmund. "Conferência 21 – O desenvolvimento da libido e as organizações sexuais (1917)". In: *Obras completas vol. 13– Conferências introdutórias à psicanálise (1916-1917)*. Trad. Paulo César de Souza. São Paulo: Companhia das letras, 2014a.

______. "A negação (1925)". In: *Obras completas vol. 16 – O eu e o id, "Autobiografia" e outros textos (1926-1929)*. Trad. Paulo César de Souza. São Paulo: Companhia das letras, 2011.

______. "A questão da análise leiga: diálogo com um interlocutor imparcial (1926)". In: *Obras completas vol. 17 – Inibição, sintoma e angústia, O futuro de uma ilusão e outros textos (1926-1929)*. Trad. Paulo César de Souza. São Paulo: Companhia das letras, 2014b.

______. "A repressão (1915)". In: *Obras completas vol. 12 – Introdução ao narcisismo, Ensaios de metapsicologia e outros textos (1914-1916)*. Trad. Paulo César de Souza. São Paulo: Companhia das letras, 2010.

FOUCAULT, Michel. *Histoire de la sexualité I – la volonté de savoir*. Paris: Gallimard, 2005.

KAFKA, Franz. *O processo.* Trad. Modesto Carone. São Paulo: Companhia das Letras, 1997.

LACAN, Jacques. *Le séminaire VIII – Le transfert (1960-1961)*. Paris: Seuil, 1991.

Schelling e a Questão dos Postulados Práticos em "Cartas Filosóficas sobre o Dogmatismo e Criticismo"

[Schelling and the Question on the Practical Postulates in "Philosophical Letters on Dogmatism and Criticism]

Marília Cota Pacheco*

Resumo: Em "Cartas Filosóficas Sobre o Dogmatismo e o Criticismo" Schelling demonstra que tanto o dogmatismo quanto o criticismo têm um ponto comum: a incapacidade de demonstrar a dedução do processo de individuação de onde o múltiplo surge a partir de uma unidade absoluta. De um lado, Schelling ressalta que Espinosa podia de certo formular a questão "como o Absoluto sai de si mesmo", entretanto, lhe era inconcebível como o Absoluto pode sair de si mesmo. Por outro lado, Schelling considera que o criticismo pode provar a necessidade de proposições sintéticas *para o domínio da experiência*. "Mas ... por que há, em geral, um domínio da experiência?" A solução para as duas questões supracitadas se torna necessariamente um postulado prático. Pelo viés dos postulados práticos apresentaremos uma das raízes, por assim dizer, da dialética da imaginação no interior do pensamento de Schelling, a saber, a noção kantiana de incondicionado.
Palavras-chave: postulados práticos, criticismo, dogmatismo, incondicionado.

Abstract: According to the text "Philosophical Letters on Dogmatism and Criticism" Schelling demonstrates that both dogmatism and criticism have one common point: the inability to demonstrate the deduction of the individuation process from which the multiple arises from an absolute unity. On the one hand, Schelling points out that Spinoza could indeed formulate the question "how the Absolute comes out of himself", however, it was inconceivable to him how the Absolute can come out of itself. On the other hand, Schelling considers that criticism can prove the need for synthetic propositions for the domain of experience. "But ... why is there, in general, a domain of experience?" The two aforementioned issue's solution becomes necessarily a practical postulate. This paper aims to present one of the roots, so to speak, of the dialectic of imagination within Schelling's thought, namely: Kant's notion of unconditioned.
Keywords: practical postulates, criticism, dogmatism, unconditioned.

*Professora substituta no Departamento de Filosofia da Universidade de Brasília (UnB). Doutora em Filosofia pela USP. E-mail: mariliacota@gmail.com.

DOI: https://doi.org/10.26512/rfmc.v6i2.19682

I - Introdução

O texto "Cartas Filosóficas Sobre o Dogmatismo e o Criticismo" foi publicado entre 1795 e 1796 no *Philosophisches Journal einer Gesellschaft Teutscher Gelehrten* de *Jena*. Friedrich Wilhelm Joseph Schelling tinha 21 anos. Neste texto, Schelling demonstra que tanto o dogmatismo quanto o criticismo têm um problema comum: a incapacidade de demonstrar a dedução do processo de individuação de onde o múltiplo surge a partir de uma unidade absoluta. De um lado, Schelling ressalta que Espinosa podia de certo formular a questão "como o Absoluto sai de si mesmo", entretanto, lhe era inconcebível como o Absoluto pode sair de si mesmo. Por outro lado, Schelling considera que o criticismo pode provar a necessidade de proposições sintéticas *para o domínio da experiência*. "Mas ... por que há, em geral, um domínio da experiência?" A solução para as duas questões supracitadas se torna necessariamente um postulado prático.

A originalidade de Schelling constitui no fato de que o autor especula a possibilidade de solucionar tal problema, considerando que a unidade absoluta da razão só pode aparecer, por assim dizer, pela necessidade de a razão em geral tornar-se prática. Contudo, é uma necessidade que não vale "para uma razão determinada, aprisionada nos grilhões de um sistema isolado; (...) a razão "deveria, onde o saber cessa, criar ela mesma um novo domínio, isto é, teria de tornar-se, de razão meramente cognoscente, uma razão criadora."[1]

Antes de tratarmos do texto "Cartas Filosóficas Sobre o Dogmatismo e o Criticismo", apresentaremos uma das raízes, por assim dizer, da dialética da imaginação no interior do pensamento de Schelling, pelo viés da noção kantiana de incondicionado nos primeiros escritos de Schelling. Para tanto nos apoiaremos no ensaio de Eric Watkins "*The early Schelling on the unconditioned*"[2] onde a noção de incondicionado em Schelling é apresentada através de dois de seus primeiros escritos, a saber: *Über die Möglichkeit einer Form der Philosophie überhaupt* –1794 (*Sobre a possibilidade de uma forma absoluta de filosofia*) e *Vom Ich als Prinzip der Philosophie oder über das Unbedingte im menschlichen Wissen* – 1795 (Do eu como o princí-

[1] Schelling, F.W.J. "Cartas Filosóficas sobre o Dogmatismo e o Criticismo", in: Textos escolhidos. São Paulo, Ed. Nova Cultural, 1989; tradução de Rubens Rodrigues Torres Filho, p.

[2] Watkins, Eric. "The early Schelling on the unconditioned", in: *Interpreting Schelling: Critical Essays*. Edited By Lara Ostaric, Cambridge University Press, 2014, pp. 10-31.

ISSN: 2317-9570

pio da filosofia, ou Sobre o Incondicionado no Conhecimento Humano).

II – Eric Watkins: Kant e o incondicionado

Em seu ensaio, Watkins ressalta o projeto global de Kant na Crítica da Razão Pura:

> Os resultados a que Kant chega são: podemos ter cognição sintética a priori somente de objetos que nos são dados em sensibilidade (aparências) e que objetos que não nos são dados na sensibilidade (coisas em si) podem ser pensados mas não percebidos. Os objetos da metafísica tradicional, como Deus, a alma e nossa liberdade, não são, Kant pensa, objetos que podem ser dados a nós em sensibilidade, sua análise revela limites fundamentais sobre o que podemos conhecer, limites que destroem a possibilidade da metafísica tradicional. (...) Kant acrescenta que devemos ainda usar nossas ideias dos objetos da metafísica tradicional, mas apenas como princípios reguladores, de modo a organizar as aparências de maneiras sistemáticas, (...).[3]

Dentre as muitas descrições da razão que Kant fornece, Watkins considera que a mais interessante e inovadora é que "a razão é uma faculdade ativa que busca a totalidade das condições para o que é condicionado e pode assim encontrar um lugar de descanso apropriado apenas identificando o incondicionado."[4] Kant introduz essa caracterização da razão mostrando como ela pode explicar os silogismos na lógica por meio do uso lógico da razão. A ideia básica é que um silogismo é uma relação racional entre duas premissas e uma conclusão onde as premissas de um silogismo contêm as condições de sua conclusão, as quais podem, portanto, ser condicionadas pelo que está contido em suas premissas.[5]

Segue-se que, tipos diferentes de silogismos dependem dos diferentes tipos de relações de condicionamento lógico (por exemplo, categórico, hipotético e disjuntivo). Além disso, uma vez que a razão tenha identificado as premissas a partir das quais uma

[3] Idem, pp. 13-14.
[4] Idem, p. 14.
[5] C.f. Watkins, op. cit. p. 14-15.

determinada cognição segue, ela continuará a procurar por outras premissas a partir das quais as premissas originais seguirão, por sua vez, como conclusões. O resultado claro é que a razão procura identificar uma série de silogismos interconectados, que vão desde cognições mais específicas, até as cognições mais gerais mais elevadas. Kant estende o escopo da razão para além dos silogismos, introduzindo o que ele chama de uso real da razão. Enquanto o uso lógico da razão diz respeito às diferentes relações de condicionamento entre os conceitos de silogismos, o uso real da razão diz respeito às relações de condicionamento que se estabelecem entre os objetos.[6]

Assim como com o uso lógico da razão, também no seu uso real, a razão busca pela completude das condições para os objetos condicionados e, portanto, para o incondicionado. Ao estender as relações condicionadoras reais ao incondicionado, a razão exige que os objetos sejam incondicionados em relação a (1) sua magnitude e, assim, constituem o mundo como uma totalidade; (2) sua composição, de tal modo que esse mundo consiste, em última análise, em uma série simples ou infinitamente divisível de partes; (3) sua causalidade, exigindo causas não causadas, ou liberdade absoluta, ou uma regressão infinita de causas; (4) sua atualidade, exigindo um ser necessário; (5) seu status como estados mentais (para representações), exigindo um sujeito absoluto, ou alma; e (5) sua possibilidade, requerendo um *ens realissimum* como base de sua possibilidade.[7]

Ao tentar explicar os objetos condicionados que experimentamos, a razão está comprometida com a existência de algo incondicionado em cada um desses aspectos, porque, pelo menos para nossa razão, a existência de algo condicionado implica a existência de algo não condicionado. Ao mesmo tempo, Kant também tenta desmascarar as falácias que são supostamente cometidas em todos os argumentos dogmáticos que pretendem transmitir a cognição do mundo como uma totalidade, a alma e Deus. Na Crítica da Razão Pura,[8] Kant faz uma distinção entre condições cognitivas, as exigências da razão, que podem ser lógicas, epistêmicas ou metafísicas. Por conseguinte, é perfeitamente apropriado sustentar que a

[6]Idem, p. 15.
[7]Ibidem.
[8]Ibidem.

ISSN: 2317-9570

razão exige que *postulemos* a existência de objetos incondicionados, mesmo que não seja possível para nós conhecê-los. Isto é, mesmo que nos seja negada a plena cognição de tais objetos, ainda é possível ter um certo tipo de fé ou crença racional [Glaube] em tais objetos.

Watkins ressalta:

> Epistemologicamente, a razão exige que procuremos um princípio incondicionado a partir do qual todas as outras cognições sejam derivadas (via silogismos). (...) Metafisicamente, a razão exige que usemos nossas ideias dos objetos tradicionais da metafísica como princípios reguladores, mas (...) também exige que postulemos a existência real de algo não condicionado que condiciona os diferentes objetos que experimentamos como condicionados em várias maneiras. Na prática, a razão exige crença na realidade dos objetos incondicionados de Deus, da liberdade e da imortalidade da alma.[9]

Com esse esboço da noção kantiana de incondicionado, passemos ao trabalho filosófico do jovem Schelling.

III – o incondicionado no jovem Schelling

No período entre 1794-1800 a questão fundamental é a relação de Schelling com Fichte. Entretanto, na medida em que há uma diferença essencial quanto à origem filosófica de ambos, seria errado caracterizar Schelling como um simples discípulo de Fichte. Segundo Schulz, Fichte foi completamente determinado pela filosofia de Kant. Fichte já vê na apercepção transcendental (o *eu penso* que necessariamente deve acompanhar todas as minhas representações) a possibilidade de colocar o Eu como princípio essencial da filosofia, na medida em que o sujeito também se torna objeto através do esquematismo da imaginação. Enquanto fundamento da possibilidade do saber da consciência finita, esse Eu deve ser pressuposto, contudo sempre permanece velado.

O jovem Schelling, por sua vez, não parte de uma tal crítica sobre a possibilidade do saber. Ele se esforça para alcançar uma unidade última onde a totalidade esteja em concordância. Esse inte-

[9] Watkins, op. cit. p. 15-16.

resse por uma unidade originária foi proporcionado essencialmente pela filosofia de Espinosa, através do texto de Jacobi *Über die Lehre des Spinoza in Briefen an Moses Mendelssohn*. Schelling parece ter encontrado em Espinosa o autor que se decide por um incondicionado conhecido (*bekannte*) que não existe por si enquanto pessoa colocada acima do mundo. A expressão formal dessa unidade é *hen kai pan*. Nos primeiros escritos de Schelling encontramos várias vezes tal expressão formal da unidade, que ele procura integrar à filosofia de Fichte.

Segue-se que as reflexões de Fichte sobre o Eu como princípio conduzem Schelling ao seguinte discernimento: o incondicionado não pode ser buscado no âmbito da essência material disponível (*dinglich-vorhandenen*) na medida em que, necessariamente, uma coisa surge a partir do Eu, ou seja, ao conhecer todos os objetos, o Eu torna os objetos cognoscíveis. Em concordância com Fichte, Schelling esclarece que o incondicionado nunca pode encontrar-se naquilo que pode ser determinado ou conhecido; o incondicionado só pode se situar naquilo que é determinante (*bedingend*). Enquanto tal, o incondicionado só pode ser o Eu absoluto enquanto princípio de todo saber. Essa passagem de Espinosa a Fichte carrega consigo um sério problema. Para Schelling, o Eu absoluto enquanto princípio incondicionado não se manifesta como princípios fundamentais reflexivos que se tornam conhecidos da consciência finita. Ele entenderá o Eu absoluto como um último princípio ainda acima de toda separação que se encontra fora da unidade originária.

Nesse sentido, permanece espinosista, mas, por outro lado, numa carta a Hegel[10], esclarece o que o diferencia de Espinosa: o incondicionado só pode ser o Eu como liberdade absoluta. Ao colocar o Eu absoluto em oposição à esfera da determinação finita, Schelling se defronta obrigatoriamente com duas questões: 1) como o incondicionado em si mesmo pode ser alcançado por nós? 2) como, a partir desse incondicionado enquanto puro ser, podemos conhecer o mundo determinado pela diversidade? Os textos de Schelling entre 1794 e 1800 mostram as diversas tentativas do autor para solucionar o problema em questão e, por conseguinte, mostram também as diferentes maneiras como foi formulado.

No texto *Vom Ich...* Schelling modela o princípio de todo saber

[10]C.f. in: *System des transzendentalen Idealismus*, Hamburg, Meiner, 2000; "Einleitung", p. XII.

ISSN: 2317-9570

de modo genuinamente idealista. A argumentação é: todo saber é saber de algo, caso contrário não haveria nenhum saber real; entretanto, deve haver no saber um fundamento originário de toda a realidade. Esse fundamento originário só pode ser incondicionado, ou aquilo que só é concebível através do seu ser, isto é, o fundamento originário só pode ser aquilo *pelo que* o ser e o pensar coincidem. Esse *pelo que* é a absolutez do *eu sou*, de tal modo que o meu eu contém um ser que antecede todo pensar e todo representar; esse ser *é* antes e depois de ser pensado simplesmente porque ele *é* enquanto pensa a si mesmo, ou ainda: ele é porque ele apenas pensa a si mesmo e ele pensa a si mesmo porque ele é[11] e, justamente por isso, a forma originária do Eu absoluto está de acordo com a sua pura identidade e sua absoluta liberdade. Schelling demonstra a determinação do Eu absoluto em suas formas subordinadas de acordo com as categorias kantianas: o Eu absoluto da quantidade é segundo a unidade absoluta, o da qualidade segundo a realidade absoluta, o da relação segundo a substancialidade e causalidade absoluta e o da modalidade segundo o puro e absoluto ser[12].

Como bem ressalta Watkins, o prefácio do texto “Do eu como o princípio (...)” deixa margem para acreditarmos que os primeiros esboços do pensamento de Schelling dizem respeito ao início de seus estudos sobre Kant, sobretudo a noção kantiana de incondicionado. Por quê? Ora, se se aceita o argumento de Kant de que, através das relações condicionantes, o conceito de condicionado leva ao de incondicionado, então o que Schelling explicitamente afirma no “Do eu como o princípio (...)” é que “o único critério que alguém pode usar para estabelecer o primeiro princípio da filosofia é aquele de ser absolutamente incondicionado.”[13] Kant não explicita nenhum primeiro princípio de sua filosofia, contudo, Schelling como que toma os conceitos de condicionado e incondicionado que Kant enfatizou e estende o uso deles no contexto dos princípios, ainda no espírito de Kant.

Ao que tudo indica, em contrapartida à dedução reflexiva de Fichte, Schelling apresenta uma descrição ontológica do Eu absoluto sob o prisma das determinações finitas para integrá-las à ab-

[11] In: *Vom Ich als Princip der Philosophie oder über das Unbedingte im menschlichen Wissen* (1795); Surkamp Verlag, Band I, p. I/1, 167.

[12] Idem, 160.

[13] Watkins, op. cit. pp. 18-19.

solutez do Eu e elevá-las acima da esfera finita. Então, como a integração é mantida se o Eu absoluto está acima de seu próprio eu consciente? Schelling esclarece que a autoconsciência pressupõe o risco de se perder o Eu[14], pois o eu que aparece na consciência não é mais o Eu absoluto puro, para o qual, de um modo geral, não há nenhum objeto e que tampouco pode vir a ser objeto. O Eu absoluto em si mesmo não pode vir à consciência de modo mediato, ele só pode ser incluído na consciência de modo imediato. A única inclusão imediata é a intuição. Mas a intuição do Absoluto não pode ser de modo algum uma intuição sensível, ela só pode ser intelectual, pois a intuição sensível só é produzida a partir de seu objeto concreto, enquanto a intuição intelectual encerra em si mesma uma unidade imediata de um *intuinte* e de um *intuído*: "Onde há objeto, há intuição sensível e vice-versa. Logo, onde não há *nenhum* objeto, isto é, no Eu absoluto, não há nenhuma intuição sensível, portanto, ou não há absolutamente intuição intelectual, ou há absolutamente intuição intelectual. *O Eu absoluto é, para si mesmo, determinado como puro Eu na intuição intelectual*".[15] Aqui, a intuição intelectual, entendida como uma espécie de entrega do eu consciente ao puro Eu onde não há oposição, assegura e mantém a integração das determinações finitas no Eu absoluto que está acima de seu eu consciente.

Contudo, ainda permanece a questão: como é possível, a partir do infinito, pormos o Absoluto como uma unidade indistinta em contraposição à esfera finita sem perdermos a relação de pura identidade do Eu absoluto e, ao mesmo tempo, produzirmos juízos sintéticos? Ela é, em suma, a questão sobre a síntese da unidade infinita e da diversidade finita e é justamente a questão principal do texto de 1795 "Cartas Filosóficas sobre o Dogmatismo e o Criticismo" (*Philosophischen Briefen über Dogmatismus und Kriticismus)*.

Nesse texto, na medida em que o Absoluto está posto como identidade acima de toda oposição e, de um modo geral, uma síntese só pode surgir através da oposição, o nosso próprio questionamento sobre a síntese do infinito e do finito constitui a condição de tal

[14] *Vom Ich* Idem, 180.

[15] "*Wo Objekt ist, da ist sinnliche Anschauung, und umgekehrt. Wo also **kein** Objekt ist, d.h. im absoluten Ich, da ist keine sinnliche Anschauung, also entweder gar keine oder **intellektuale** Anschauung. **Das Ich also ist für sich selbst als blosses Ich in intellektualer Anschauung bestimmt**.*" In: *Vom Ich als Princip der Philosophie oder über das Unbedingte im menschlichen Wissen* (1795); Surkamp Verlag, Band I, p. I/1, 181. Tradução nossa.

ISSN: 2317-9570

síntese. Assim, a condição da síntese está dada, porém, o problema não pode ser totalmente resolvido no âmbito teórico, pois a intuição intelectual, na sua unificação imediata, joga fora, por assim dizer, toda oposição. Nessa imediatez da intuição intelectual, a oposição está numa espécie de estado de morte (*Zustand des Todes*).

A reflexão sobre esse estado da oposição se apresenta necessariamente como um regresso em que nos encontramos privados de nós mesmos e, portanto, não fornece nenhuma resposta. Segue-se que, justamente porque nessa reflexão a oposição do intuinte e do intuído está suspensa, essa reflexão será explicada ou de maneira espinosista (suspensão de toda egoidade num objeto absoluto), ou de maneira idealista (suspensão de toda objetividade num Eu absoluto).

Quando Schelling demonstra que essas duas interpretações são as únicas corretas, ele também está provando que a reflexão, da qual ambas partem, deve ser determinada como Eu absoluto. Consequentemente, no âmbito teórico dessas duas interpretações, a resposta para a questão sobre a passagem do infinito ao finito será sensatamente a mesma: do infinito ao finito não há passagem. A solução desse problema só pode ser apresentada no âmbito prático a partir da nossa tendência natural rumo ao infinito, isto é, pelo nosso próprio empenho para realizar o Absoluto em nós mesmos através da atividade infinita, a ideal. Entretanto, justamente porque a intuição intelectual não proporciona nenhum conhecimento do Absoluto, este não terá mais lugar no homem finito que, a partir do estado imediato da intuição intelectual, precisa voltar para si mesmo na esfera finita e, através da reflexão, banir o esforço rumo ao infinito.

Apenas o uso imanente que fazemos do princípio absoluto, na filosofia prática, para o conhecimento de nossa destinação, nos redireciona para fora do Absoluto. Enfim, nas *Cartas*, para responder à questão sobre a síntese da unidade infinita e da diversidade finita, Schelling mostrou que essa questão só tem resposta no âmbito prático, através do uso imanente que fazemos da ideia do Absoluto.

IV – A questão dos postulados em "Cartas Filosóficas sobre o Dogmatismo e o Criticismo"

Na primeira Carta a questão dos postulados aparece através de uma censura ao kantismo adocicado dos teólogos que lecionavam em Tübingen, sobretudo Johann

Flatt e Gottlieb Storr,[16] isto é, Schelling refuta a chamada prova moral da existência de Deus. Basicamente seu argumento consiste no seguinte: na medida em que os teólogos de Tübingen entendem que a razão teórica é demasiado fraca para conceber um Deus, e a ideia de Deus só é realizável por exigências morais, eles entendem que a necessidade prática é mais premente que a teórica, desconsiderando que a razão teórica não encontra em parte nenhuma espaço para causalidade absoluta, admitem uma causalidade absoluta no âmbito da razão prática. Se assim o é, também precisam admitir que a razão teórica teria de ser ampliada pois uma "admissão prática", segundo Schelling, é um assentimento teórico segundo a forma, mas que, segundo a matéria, segundo o fundamento, é prático.[17] Com isso, Schelling reafirma a máxima da primeira Crítica: "do infinito ao finito não há passagem", isto é, podemos pensar o incondicionado, mas não podemos conhecê-lo.

Na segunda Carta, Schelling entende que a exigência de não se admitir nenhuma objetividade absoluta, é uma exigência própria de nosso ser originário e não apenas uma exigência da faculdade de conhecer. Noutras palavras: para Schelling, tomar a faculdade de conhecer como a própria razão é o grande equívoco dos teólogos de Tübingen:

> Porque a primeira tentativa empreendida contra o dogmatismo só podia partir da faculdade-de-conhecer, vós acreditastes poder imputar audaciosamente "a razão a culpa do malogro de vossa esperança. (...) Segundo vossa crença, o dogmatismo mesmo, que teria um fundamento mais profundo do que a simples faculdade-de-conhecer, zombaria de nossas provas.[18]

Na terceira Carta, Schelling mostra que a *Crítica da Razão Pura* só alcança uma refutação negativa do dogmatismo, isto é, apresenta apenas a indemonstrabilidade teórica do dogmatismo, pois o criticismo só podia partir de um ponto que fosse comum a ele e ao dogmatismo. Por quê? Por-

[16] C.f. Torres Filho, Rubens Rodrigues in: nota de sua tradução de "Cartas Filosóficas sobre o Dogmatismo e o Criticismo", Textos escolhidos. São Paulo, Ed. Nova Cultural, 1989; tradução de Rubens Rodrigues Torres Filho, p. 5.

[17] Schelling, F.W.J. "Cartas Filosóficas sobre o Dogmatismo e o Criticismo", in: Textos escolhidos. São Paulo, Ed. Nova Cultural, 1989; tradução de Rubens Rodrigues Torres Filho, p. 7.

[18] Idem, p. 9.

ISSN: 2317-9570

que só há em geral sistemas diferentes se houver ao mesmo tempo um domínio comum a eles. Esse domínio comum é o que Schelling chama de "exílio do Absoluto; pois sobre o Absoluto seríamos todos unânimes, se nunca deixássemos sua esfera; e, se nunca saíssemos dela, não teríamos nenhum outro domínio para controvérsias."[19] O exílio do absoluto é a finitude, por isso o criticismo só poderia refutar o dogmatismo a partir da faculdade de conhecer, estabelecendo as condições de objetivação da experiência.

E ainda: para Schelling, a Crítica da Razão Pura só começou efetivamente sua luta a partir do problema que diz respeito ao ponto comum de toda filosofia e que Kant apresenta no início de sua obra com a pergunta: *Como chegamos, em geral, a julgar sinteticamente* Para Schelling, expressa de outro modo, a pergunta diz: *Como chego, em geral, a sair do Absoluto e a ir a um oposto?* Recorramos às palavras do autor para esclarecer a analogia estabelecida entre as duas questões.

> Com efeito, uma síntese, em geral, surge pelo conflito da pluralidade com a unidade originária. Pois, sem conflito em geral, nenhuma síntese é necessária; onde não há pluralidade, há unidade pura e simples; mas se a pluralidade fosse o originário, mais uma vez não haveria síntese. (...) embora só possamos conceber a síntese por uma unidade originária contraposta à pluralidade, a Crítica da Razão Pura não podia (...) chegar àquela unidade absoluta, porque, (...) só podia partir precisamente daquele fato de que parte a controvérsia da própria filosofia.[20]

O que Schelling parece explicitar aqui é que o sujeito, assim que julga objetivamente, sai de si mesmo e é obrigado a empreender uma síntese e, justamente por isso, o objeto não pode aparecer como absoluto em nenhuma síntese; se assim o fosse, o objeto como absoluto não deixaria subsistir nenhuma síntese. Até esse ponto o dogmatismo está teoricamente refutado. Contudo, Schelling ressalta[21] que a faculdade de conhecer está longe de ser esgotada com

[19]Idem, p. 10.

[20]Schelling, F.W.J. "Cartas Filosóficas sobre o Dogmatismo e o Criticismo", in: Textos escolhidos. São Paulo, Ed. Nova Cultural, 1989; tradução de Rubens Rodrigues Torres Filho, p. 11.

[21]C.f. idem p. 12.

a ação da síntese. Por quê? Porque na medida em que pressupõe que as ações meramente formais do sujeito não se submetem a nenhuma dúvida, uma crítica da faculdade-de-conhecer busca provar a marcha de toda síntese material pela marcha de toda síntese meramente formal, ou seja, pressupõe que o sujeito é obrigado a ir dos juízos condicionados aos incondicionados (por prosilogismos). Por conseguinte, tem de aceitar que a razão teórica visa necessariamente um incondicionado, "e justamente ao fazer isso, tem de aniquilar aquilo que acaba de edificar. Ou seja: enquanto permanece no domínio da síntese, ela triunfa sobre o dogmatismo; tão logo abandona esse domínio (...) a luta começa de novo."[22]

Schelling considera que a *Crítica da Razão Pura* demonstrou que a controvérsia sobre onde está o princípio daquela unidade que é expressa no juízo sintético não pode ser decidida na filosofia teórica e, ao fazê-lo a partir da faculdade-de-conhecer, a *Crítica da Razão Pura* permite, a partir da essência da razão, a dedução de dois sistemas diretamente opostos entre si, dogmatismo (realismo) e criticismo (idealismo). Para Schelling, a "*Crítica da Razão Pura* ensinou ao dogmatismo como se tornar dogmatismo, isto é, um sistema de realismo objetivo fundado em si mesmo."[23] E se alguém julgar que tal "afirmação é inteiramente contrária ao espírito da Crítica" (...), Schelling contra-argumenta:

> Permita-me, portanto, recordar-lhe, também *de antemão*, apenas uma parte da Crítica, que justamente até agora, foi menos esclarecida de todas: refiro-me à parte que trata da coisas em si. Se se acredita que a Crítica da Razão Pura deve fundar apenas o criticismo, então, justamente nesse ponto, não há como salvá-la, ao que entendo, da acusação de inconsequência.

A noção de *coisa em si* da *Crítica* deixa subsistir o sistema do realismo ou dogmatismo consequente porque Schelling a entende como uma limitação que produz no Eu (enquanto princípio) a tendência para se tornar objeto de modo que nessa sua objetivação ele se compreenderá como sujeito. Esse querer compreender a si mesmo como sujeito sem-

[22] Ibidem.
[23] Schelling, F.W.J. "Cartas Filosóficas sobre o Dogmatismo e o Criticismo", in: Textos escolhidos. São Paulo, Ed. Nova Cultural, 1989; tradução de Rubens Rodrigues Torres Filho, p. 15.

pre impulsiona o Eu para a sua objetivação mais elevada. Entretanto, Schelling demonstra que nesse processo o Eu é parcial, pois nele o Eu consuma apenas o seu agir objetivo e, por conseguinte, ainda não reconhece a si mesmo como sujeito. Esse caminho para a autoconsciência só se torna conhecido em virtude do filósofo que realiza o Eu objetivo fundado na sua própria autoconsciência.[24]

Isso significa que, para Schelling, no particular da sensação do ser limitado, a atividade ideal determina o limite da atividade real como uma esfera de passividade. O Eu se cinde em passividade e atividade apenas para não refletir a sua própria atividade limitante. Não ocorrendo essa reflexão, a passividade aparece necessariamente como uma efetividade estranha ao próprio Eu que a determina, isto é, a passividade aparece como uma efetividade da *coisa em si* que, de fato, é a sombra da atividade ideal que ultrapassou o limite, pois para o Eu intuir a si mesmo como o que sente ele precisa reintegrar ativamente em si o seu oposto, ou seja, reintegrar em si a passividade.[25] E, para Schelling, esta é a possibilidade, em filosofia, de se deduzir dois sistemas diretamente opostos entre si. Por conseguinte, Schelling dirá na V Carta: "a *Crítica da Razão Pura* provou (...) que nenhum sistema – tenha ele o nome que tiver – é, em sua perfeição, objeto do saber, mas apenas objeto de uma ação, praticamente necessária, mas infinita."[26]

Pode-se dizer que, para Schelling, o método dos postulados práticos foi estabelecido tanto para o dogmatismo quanto para o criticismo exprimirem o domínio do Absoluto. Na VI Carta, Schelling considera que Espinosa usou o método dos postulados pela mesma razão do criticismo, isto é: pela impossibilidade de conceber como saímos do absoluto para, pura e simplesmente opor algo a nós. Para Schelling, Espinosa afirma apenas uma causalidade imanente do objeto absoluto meramente porque lhe era inconcebível como o Absoluto pode sair de si mesmo, ou seja, ele podia de certo formular a questão, mas não soluciona-la, daí a função mais ampla do postulado na filosofia de Espinosa.

A ideia de Deus como causa imanente que existe com suas

[24]C.f. *System des transzendentalen Idealismus*, Hamburg, Meiner, 2000, pp. 55-57; Surkamp Verlag, Band I, p. I/3, 373-388.

[25]Ibidem.

[26]Schelling, F.W.J. "Cartas Filosóficas sobre o Dogmatismo e o Criticismo", in: Textos escolhidos. São Paulo, Ed. Nova Cultural, 1989; tradução de Rubens Rodrigues Torres Filho, p.16.

criaturas é atrativa para Schelling porque, assim como Espinosa, ele também parte do pensamento fundamental de que Deus, enquanto Absoluto, não pode ter nenhum objeto fora de si e, por conseguinte, não pode ser entendido, teoricamente, como um objeto, isto é, do lado de "fora". Resta que se, teoricamente, Deus não pode ser entendido como um objeto, então, Schelling tem que torna-*Lo* sujeito do discurso filosófico. Por quê? Vejamos: o axioma fundamental dos dois sistemas (dogmatismo e criticismo) é uma questão insolúvel teoricamente, a solução da questão sobre como chegamos a sair do Absoluto é uma solução que só pode se dar no campo da prática ou da subjetividade. Para o autor, a solução dessa questão se dá por um ato de escolha e, por isso, dizemos que se dá no campo da subjetividade; por conseguinte, dizemos que é por liberdade que encontramos a solução dessa questão.

Se, Schelling *desreifica* Deus, tornando-o sujeito do discurso filosófico, parece razoável considerar que a sua primeira definição para o Absoluto seja sujeito. Contudo, um sujeito que tem a necessidade interna de passar ao objeto, pois, na VI Carta, Schelling concebe uma espécie de faculdade *produtiva realizadora*, não pelo *saber*, mas pelo *agir*, mediante o qual, a razão criaria ela mesma, onde o saber cessa, um novo domínio, sem perder-se em vãs ficções e, assim, teria de tornar-se, de razão meramente *cognoscente*, uma razão criadora – de razão teórica, uma razão prática. "Mas a necessidade de se tornar *prática* vale para a razão *em geral*, não para uma razão determinada, aprisionada nos grilhões de um sistema isolado."[27]

O fato de Schelling considerar Deus o sujeito do discurso filosófico viabiliza uma outra noção de sua filosofia que, por sua vez, fundamenta a razão criadora, a saber: a *identidade imediata*, intrínseca de real e ideal. Essa identidade imediata é a essência do Absoluto e, justamente por isso, não pode ser conhecida mediante definições, mas apenas com a intuição não sensível.

Para finalizar: pelo viés da noção kantiana de incondicionado as "Cartas Filosóficas Sobre o Dogmatismo e o Criticismo" nos fornecem um bom indício do que leva Schelling a conceber posteriormente em *Sistema do Idealismo Transcendental* sua dialética da imaginação como forma sistemática do Eu enquanto princípio de uma consciência do mo-

[27] Idem, p. 19.

ISSN: 2317-9570

mento presente que surge, originariamente, de um passado transcendental. Nos textos de sua filosofia da identidade,[28] essa forma sistemática será ampliada como uma *forma formante* de sujeito e objeto no Absoluto e, nisso, possibilita a fundamentação da essência da identidade alma – Absoluto na intuição intelectual, entendida como princípio e conhecimento que supera infinitamente toda determinação conceitual e, ao mesmo tempo, efetiva o princípio transcendental e a causa imanente da equipossibilidade volitiva e cognitiva do puro sujeito-objeto. Noutras palavras: a filosofia alcança a reconstrução do Absoluto como sujeito-objeto de modo objetivo porque ela contempla algo que se revela numa ação de manifestação de si mesma como sujeito.

Referências

SCHELLING, Friedrich Wilhelm Joseph. *Sämtliche Werke*, ed. K. F. A. Schelling, Stuttgart, 1856/1861 (1° seção, vols. 1-10; 2° seção, vols 1-4).

______. *Ausgewählte Schriften: in 6 Bd.* / Frankfurt am Main: Suhrkamp – Verlag, – 1. Auflage 1985.

______. *A Filosofia da Arte*. Tradução de Márcio Suzuki. São Paulo, Edusp, 2001.

______. *Philosophische Briefe über Dogmatismus und Kriticismus* (1795). In: Friedrich Wilhelm Joseph Schelling, Historisch – Kritische Ausgabe, Reihe I: Werke; Werke 3, herausgegeben von Harmut Buchner, Wilhelm G. Jacobs und Annemarie Pieper. Stuttgart: Frommannn – Holzboog, 1982.

______. *System des transzendentalen Idealismus*, Hamburg, Meiner, 2000.

______. *Textos escolhidos*. São Paulo, Ed. Nova Cultural, 1989; tradução de Rubens Rodrigues Torres Filho.

______. *Vom Ich als Princip der Philosophie oder über das Unbedingte im menschlichen Wissen* (1795); Surkamp Verlag, Band I.

TORRES FILHO, Rubens Rodrigues. "O Simbólico em Schelling", in: Ensaios de Filosofia Ilustrada, São Paulo, Editora Brasiliense, 1987.

[28]Tais como: *A Filosofia da Arte*. Tradução de Márcio Suzuki. São Paulo, Edusp, 2001. *Filosofia e Religione (1804);* A cura di Luigi Pareyson; traduzione di Valério Vera Milano, Mursia ed., 1987. *Propedeutica della filosofia*; traduzione di Fabio Palchetti, Pisa, ETS Editrice, 1991.

WATKINS, Eric. "The early Schelling on the unconditioned", in: *Interpreting Schelling: Critical Essays*. Edited By Lara Ostaric, Cambridge University Press, 2014, pp. 10-31.

ISSN: 2317-9570

RESENHA

Marília Cota Pacheco*

OSTARIC, Lara (Editor) ***Interpreting Schelling: Critical Essays.*** **Cambridge University Press, 2014.**

Interpretando Schelling: Ensaios Críticos é uma coletânea de onze ensaios, organizados por Lara Ostaric[2] . Esses ensaios traçam sistematicamente o desenvolvimento histórico do pensamento de Friedrich Wilhelm Joseph Schelling desde a *Filosofia Transcendental* e *Filosofia da Natureza* (1794-1800), passando pela sua *Filosofia da Identidade* (1801-1809), *Escritos sobre a Liberdade, Idades do Mundo* (1809-1827), chegando até sua *Filosofia Positiva / Negativa e à crítica de Hegel* (1827-1854).

Como bem ressalta a organizadora, o volume oferece uma compreensão mais sutil do idealismo alemão do que a oferecida por uma narrativa super simplificada "de Kant a Hegel,"[3] que retrata esse movimento filosófico como uma progressão teleológica que começa com Kant, é avançada por Fichte e Schelling e culmina no sistema de Hegel que sintetiza todas as visões anteriores.

> É claro que o assim chamado idealismo objetivo de Schelling, com um princípio incondicionado que transcende tanto o sujeito quanto o objeto, marca um afastamento do idealismo subjetivo de Fichte e abre caminho para o sistema de Hegel. No entanto, prestando mais atenção à constelação das ideias que motivaram o pensamento de Schelling, é possível apreciá-lo mais como um pensador original, um pensador cujo impacto ultrapassou o estágio inicial do idealismo alemão e cu-

*Professora substituta do Departamento de Filosofia da Universidade de Brasília.

[2]Lara Ostaric é Professora Assistente de Filosofia na Temple University. Publicou artigos sobre Kant e Schelling e está trabalhando num livro sobre a terceira Crítica de Kant e sua influência na filosofia alemã pós Kant.

[3]*Von Kant bis Hegel* é o título do importante estudo sobre o idealismo alemão de Richard Kroner.

[4]Ostaric, Lara. *Interpreting Schelling: Critical Essays.* Edited By Lara Ostaric, Cambridge University Press, 2014, Introduction, p. 03. A tradução é minha.

jas ideias são importantes para nós hoje.[4]

A coletânea nos mostra muito mais continuidade no pensamento de Schelling do que é geralmente reconhecido; ressalta os diferentes estágios no desenvolvimento de seu sistema filosófico não como um sinal de imaturidade intelectual, nem como o resultado inevitável da influência de muitas e diferentes posturas filosóficas. O conjunto desses ensaios críticos indica que talvez as reformulações de Schelling em seu próprio sistema filosófico sejam "uma indicação de sua modéstia e seu reconhecimento de que, embora rigorosa e sistemática, a reflexão filosófica não seja onipotente diante da complexidade da condição humana," como bem ressalta Ostaric.

Em "The Early Schelling on the Unconditioned," Eric Watkins[5] faz uma análise das passagens centrais de dois ensaios do jovem Schelling: *Über die Möglichkeit einer Form der Philosophie überhaupt* – 1794 (Sobre a possibilidade de uma forma absoluta de filosofia) e *Vom Ich als Prinzip der Philosophie oder über das Unbedingte im menschlichen Wissen* – 1795 (Do eu como o princípio da filosofia, ou Sobre o Incondicionado no Conhecimento Humano). Watkins esclarece como o primeiro Schelling chega a empregar a noção de "incondicionado" no centro de seu projeto filosófico; sem negar a influência de outras figuras no pensamento inicial de Schelling, como Fichte ou Reinhold, Watkins argumenta que são as visões específicas de Kant sobre o incondicionado que desempenham um papel crucial no desenvolvimento de uma série de características fundamentais do pensamento inicial de Schelling.

Michael N. Forster[6], em seu ensaio "Schelling and Skepticism", contesta as acusações de Hegel em sua Introdução à *Fenomenologia do Espírito* sobre a filosofia de Schelling entendida como dogmática e vulnerável ao ceticismo. Segundo Forster, tais acusações não são totalmente justificadas; não se aplicam à carreira de Schelling como um todo, pois as reflexões de Schelling sobre o ceticismo e

[5]Eric Watkins é professor de Filosofia na Universidade da Califórnia em San Diego. Ele é o autor de *Kant and the Metaphysics of Causality* (Cambridge, 2005), o editor de *Kant and the Sciences* (2001), o editor e tradutor de *Kant's Critique of Pure Reason: Background Source Materials* (Cambridge, 2009).

[6]Michael n. Forster é Professor na Alexander von Humboldt, titular da cadeira de Filosofia Teórica e co-diretor do Centro Internacional de Filosofia da Universidade de Bonn. Ele é o autor de *German Philosophy of Language: From Schlegel to Hegel and Beyond* (2011), *After Herder: Philosophy of Language in the German Tradition* (2010), *Kant and Skepticism* (2008), *Wittgenstein on the Arbitrariness of Grammar* (2004), *Hegel's Idea of a Phenomenology of Spirit* (1998), and *Hegel and Skepticism*(1989).

sua relação com a filosofia passaram por três fases diferentes. A primeira é uma posição inspirada, durante o período de 1794-1800; a segunda, uma posição inspirada em Hegel, que ele adotou brevemente em 1802-1803 e a terceira, uma posição inspirada no romantismo que ele adotou por volta de 1821. No final de seu ensaio, Forster considera uma quarta fase da tentativa de Schelling de lidar com o ceticismo: a sua filosofia positiva como uma modificação da sua posição inspirada no Romantismo.

Em "The Concept of Life in Early Schelling," Lara Ostaric mostra como nos estágios iniciais de sua *Naturphilosophie* Schelling é motivado pela questão da correspondência necessária entre o eu e a natureza e, portanto, tenta demonstrar que a natureza não é um objeto inanimado desprovido de autoconsciência, mas algo que é ao mesmo tempo um sujeito e seu próprio objeto. A natureza não deve ser concebida como um mecanismo morto, mas como uma organização viva e como um "análogo da razão"e liberdade, porque ser o próprio sujeito e objeto é ser autodeterminado. É isso que Schelling considera a característica essencial da vida. Nisso, Ostaric mostra que os primeiros escritos de Schelling fazem parte de um desenvolvimento progressivo e contínuo de seu sistema filosófico.

Paul Guyer[7], em seu ensaio "Knowledge and Pleasure in the Aesthetics of Schelling" (Conhecimento e Prazer na Estética de Schelling), analisa passagens centrais do Sistema de Idealismo Transcendental de 1800, e das palestras sobre A Filosofia da Arte 1802-1803 (*Philosophie der Kunst*), para mostrar como Schelling adotou e transformou a concepção estética de Kant. Guyer alega que Kant criou uma síntese da nova estética desenvolvida em meados do século XVIII na Escócia e na Alemanha, com a teoria clássica de que a experiência estética é uma forma distinta de apreensão da verdade. A estética de Schelling favorece uma abordagem puramente cognitiva e a compreensão de que a experiência estética é prazerosa apenas porque nos libera da dor de uma contradição inescapável da condição humana.

Em "Exhibiting the Particular in the Universal': Philosophi-

[7]Paul Guyer é Professor na Brown University. Ele é o autor de nove livros sobre Kant, incluindo *Kant and the Claims of Taste* (Cambridge, 1997, 2nd edn.), *Kant and the Claims of Knowledge* (Cambridge, 1987), *Kant and the Experience of Freedom* (Cambridge, 1993), e *Kant on Freedom, Law, and Happiness* (Cambridge, 2000). O professor Guyer é um dos coeditores gerais da Cambridge Edition de Kant. Seu trabalho de três volumes, *A History of Modern Aesthetics*, foi publicado pela Cambridge em 2014.

cal Construction and Intuition in Schelling's Philosophy of Identity (1801-1804)"(Exibindo o Particular no Universal: Construção Filosófica e Intuição na Filosofia da Identidade de Schelling), Daniel Breazeale[8] discute o método de construção filosófica de Schelling em sua Filosofia da Identidade. Inuenciado pelo texto de Kant *Princípios Metafísicos*... (onde "construir" um conceito é "exibir [*darstellen*] a priori a intuição correspondente a ele") e pelo desenvolvimento posterior que Fichte dá a esse método filosófico, Schelling desenvolve sua própria concepção de construção filosófica. Breazeale concentra-se em oito das características mais importantes do método de construção de Schelling: (1) seu ponto de vista "absoluto", (2) seu princípio (a lei da identidade racional), (3) seu órgão (intuição intelectual), (4) seu método atual (exposição do particular no universal), (5) seus elementos (ideias da razão), (6) seu produto (o Sistema de Identidade), (7) sua verdade e realidade, e (8) a capacidade inata e intocável de intuição intelectual (gênio filosófico). Em sua conclusão, ele oferece um exame e uma crítica à concepção de construção filosófica de Schelling.

Em seu ensaio "Identity of Identity and Non-Identity: Schelling's Path to the Absolute System of Identity," (Identidade de identidade e não-identidade: o caminho de Schelling para o Sistema Absoluto de Identidade), Manfred Frank[9] foca no pensamento central do Sistema Absoluto de Identidade de Schelling, que diz respeito a uma forma de identidade que não é simples, mas concebida de tal modo que duas coisas diferentes pertencem inteiramente a um e mesmo todo. Frank descreve os problemas do início da filosofia moderna, para os quais a noção de identidade de Schelling tenta fornecer uma solução. Discute as figuras da história da filosofia que influenciaram a Filosofia da Identidade madura de Schelling e mostra a relevância da noção de identidade de Schelling para as teorias contemporâneas sobre mente-corpo. Na parte final Frank aborda a diferença entre a noção de identidade de Schelling e He-

[8]Daniel Breazeale é professor de Filosofia na Universidade de Kentucky. Ele é o autor de *Fichte and the Project of Transcendental Philosophy* e numerosos artigos de revistas, capítulos de livros, traduções e edições de / sobre filosofia alemã de Kant a Nietzsche, com um foco de pesquisa sobre a filosofia de J.G. Fichte.

[9]Manfred Frank é Professor Emérito de Filosofia na Eberhard Karls University, Tübingen. Ele é autor de inúmeros artigos, edições e monografias, que foram traduzidos para mais de vinte idiomas. Seus livros incluem *Der unendliche Mangel an Sein. Schellings Hegelkritik und die Anfänge der Marxschen Dialektik* (1975/1992), *Selbstbewutsein und Selbsterkenntnis. Essays zur analytischen Philosophie der Subjektivität* (1991) e *'Unendliche Annäherung.' Die Anfänge der philosophischen Frühromantik* (1997).

gel.

Em "Idealism and Freedom in Schelling's Freiheitsschrift" (Idealismo e Liberdade no Freiheitsschrift de Schelling), Michelle Kosch[10] faz uma distinção entre uma concepção "formal"de liberdade, isto é, uma caracterização do livre-arbítrio que permite uma distinção entre comportamento imputável / não-imputável e uma concepção do livre-arbítrio como fonte de imperativos morais.

No seu ensaio "Beauty Reconsidered: Freedom and Virtue in Schelling's Aesthetics" (Beleza reconsiderada: liberdade e virtude na estética de Schelling), Jennifer Dobe considera que *Freiheitsschrift* (1809) de Schelling, contrariando a visão predominante, oferece recursos para identificar a nova e inovadora abordagem de Schelling à estética. Concentra-se nas principais passagens do discurso de Schelling de 1807 para a *Akademie der Wissenschaften* em Munique (*Über das Verhältnis der bildenden Künste zu der Natur*) e nos fragmentos de Weltalter de 1811-1515. Dobe mostra como Schelling começa a ampliar sua estética com base na nova concepção de liberdade alcançada em *Freiheitsschrift*.

Em "Nature and Freedom in Schelling and Adorno" (Natureza e Liberdade em Schelling e Adorno), Andrew Bowie[11] mostra como a tensão dialética entre existência e seu fundamento, razão autodeterminada e seu outro, no *Freiheitsschrift* de Schelling, abre espaço para uma compreensão não-dogmática da Natureza, isto é, uma compreensão da Natureza como algo que precisa ser legitimado e não algo usado como legitimação, mostrando sua relação com o sujeito e, portanto, com a liberdade.

O ensaio de Günter Zöller, "Church and State: Schelling's Political Philosophy of Religion" (Igreja e Estado: Filosofia Política da Religião de Schelling), enfoca a relação entre Igreja e Estado no Curso de Palestras Privadas de Stuttgart de 1810 e em suas Investigações Filosóficas sobre a Essência da Liberdade Humana de 1809. Num primeiro momento Zöller apresenta o pano de fundo

[10] Michelle Kosch é Professora Associada de Filosofia na Cornell University. Ela é autora de *Freedom and Reason in Kant, Schelling, and Kierkegaard* (2006) e vários artigos sobre Kierkegaard, Fichte e a filosofia continental do século XIX.

[11] Andrew Bowie é professor de Filosofia e Alemão na Royal Holloway, Universidade de Londres. Ele é o autor de *Adorno and the Ends of Philosophy* (2013), *German Philosophy: A Very Short Introduction* (2010), *Music, Philosophy, and Modernity* (Cambridge, 2007), *Introduction to German Philosophy from Kant to Habermas* (2003), *From Romanticism to Critical Theory: The Philosophy of German Literary Theory* (1997), *Schelling and Modern European Philosophy: An Introduction* (1993), e *Aesthetics and Subjectivity from Kant to Nietzsche* (1990).

histórico da filosofia política da religião de Schelling; em seguida, o desenvolvimento de uma concepção liberal e legal do Estado para uma concepção absolutista e ética na filosofia política de Schelling. Por fim Zöller discute a teoria filosófico-teológica de Schelling sobre o estado.

A coletânea finaliza com o ensaio de Fred Rush[12] "Schelling's Critique of Hegel" (A Crítica de Schelling a Hegel), cujo foco central são as preleções de Berlim, apresentadas nas décadas de 1840 e início de 1850, quando Schelling faz uma ampla distinção entre duas abordagens da filosofia: "negativa"e "positiva."O ensaio de Rush levanta a questão de até que ponto a crítica de Schelling a Hegel é válida. Sua principal alegação é que as críticas de Schelling retêm em grande parte sua força, embora algumas delas mostrem que o Schelling tardio está mais próximo de Hegel em alguns pontos do que a polêmica filosófica inicialmente sugeriria.

[12]Fred Rush é Professor Associado de Filosofia na Universidade de Notre Dame. Ele é o autor de *Irony and Idealism* (2014), *On Architecture* (2009) e o editor do *Cambridge Companion to Critical Theory* (Cambridge, 2004).

ISSN: 2317-9570