

UNIVERSIDADE FEDERAL DA PARAÍBA

LEONARDO WEBER CASTOR DE LIMA

Alfred Tarski e o Conceito de Verdade em Linguagens Formalizadas

JOÃO PESSOA – PARAÍBA

2004

UNIVERSIDADE FEDERAL DA PARAÍBA

LEONARDO WEBER CASTOR DE LIMA

Alfred Tarski e o Conceito de Verdade em Linguagens Formalizadas

Dissertação apresentada como requisito para a obtenção do grau de Mestre em Filosofia, no Programa de Pós-graduação em Filosofia, do Centro de Ciências Humanas, Letras e Artes, da Universidade Federal da Paraíba.
ORIENTADOR: Prof. Dr. Garibaldi Monteiro Sarmiento.

JOÃO PESSOA – PARAÍBA

2004

TERMO DE APROVAÇÃO

LEONARDO WEBER CASTOR DE LIMA

Alfred Tarski e o Conceito de Verdade em Linguagens Formalizadas

Dissertação aprovada como requisito para a obtenção do grau de Mestre em Filosofia, no Programa de Pós-graduação em Filosofia, do Centro de Ciências Humanas, Letras e Artes, da Universidade Federal da Paraíba, pela seguinte banca examinadora:

Prof. Dr. Garibaldi Monteiro Sarmiento – Orientador
Programa de Pós-graduação em Filosofia (PPGF / UFPB)

Prof. Dr. Marco A. Ruffino
Universidade Federal do Rio de Janeiro - UFRJ

Prof. Dr. Matias Francisco Dias
Programa de Pós-graduação em Filosofia (PPGF / UFPB)

JOÃO PESSOA – PARAÍBA

JANEIRO DE 2004

AGRADECIMENTOS

Ao CNPq, pela bolsa que financiou minhas pesquisas durante o mestrado.

Ao Prof. Dr. Marco A. Ruffino, pela gentileza de aceitar o convite para compor a banca examinadora desta dissertação.

Ao meu orientador, Prof. Dr. Garibaldi Monteiro Sarmento, pelas aulas sempre ricas, pela amizade e paciência, pela educação e respeito, pelo apoio e estímulo em momentos difíceis, e por ter refinado significativamente meus estudos em lógica.

À Prof.^a Dr.^a Ana Leda de Araújo e ao Prof. Dr. Matias Francisco Dias, pessoas inigualáveis em humanidade e solicitude, e aos quais devo uma parcela considerável de minha formação intelectual.

Ao Prof. Dr. André Leclerc, com quem iniciei minhas atividades de pesquisa científica ainda na Graduação, e com quem, ainda me lembro, assisti à primeira aula sobre Alfred Tarski.

Aos amigos do peito Carlos Eduardo Pimentel, André Joffily Abath e José Belizário Neto, pelo intercâmbio sempre saudável de idéias e sentimentos, bem como pela estabilidade longa de nossas relações.

A Emília de Moraes e Lourenço Pereira, estimados amigos, pelo cuidado afetuoso e amizade sincera.

A minha linda Manuela, pelo carinho inestimável durante todos esses anos, e por me fazer muito feliz ao seu lado.

Aos meus sete irmãos, Jorge, Beca, Lindo, Luciano, Fernando, Thiago e Erich, e à minha irmã, Coeli, pelas boas risadas que sempre damos juntos, e pelo amor e bem-estar que sempre me transmitem.

A minha mãe, Ivanize, pela vida, pelo amor, por tudo.

Ao meu pai – Polycarpo Macedo Lima (in memoriam)

Sumário

Resumo

Introdução

Capítulo 1

Prelúdio Histórico

1.1 A fase pré-Tarskiana da teoria de modelos e o caráter informal das noções semânticas.....	1
1.2 Os seminários de Varsóvia, o método de eliminação dos quantificadores e as investigações semânticas de Tarski.....	6
1.3 O problema das antinomias semânticas.....	9
1.4 O diagnóstico de Tarski das antinomias semânticas.....	12
1.5 Uma nuance conceitual: as noções de verdade-numa-linguagem <i>versus</i> verdade-numa-estrutura.....	15

Capítulo 2

O programa Tarskiano de formalização da semântica

2.1 Formalização – definições e axiomatizações.....	17
2.2 Correção formal de uma definição qualquer.....	22
2.3 Correção formal de uma definição de verdade (1).....	25
2.4 Correção formal de uma definição de verdade (2).....	27
2.5 Adequação material de uma definição qualquer.....	29
2.6 Adequação material de uma definição de verdade.....	33

Capítulo 3

O conceito de verdade formalizado

3.1 O problema de uma definição recursiva direta do conceito de verdade.....	36
3.2 Satisfação por uma seqüência denumerável de objetos.....	39
3.3 Tipos semânticos.....	41

3.4 A definição formal do conceito de verdade por uma seqüência denumerável, e do conceito de verdade.....	43
--	----

Capítulo 4

Para ler *The Concept of Truth*

4.1 Preliminares: fundamentos da metamatemática.....	46
4.2 A teoria das categorias semânticas de Husserl-Lesniewski.....	47
4.3 Tipologia Tarskiana das linguagens formalizadas.....	50
4.4 A ambigüidade semântica do conceito de satisfação.....	52
4.5 Linguagens de ordem infinita e a versão absoluta do teorema da indefinibilidade da verdade.....	56
4.6 Teoria axiomática da verdade.....	58
4.7 O pós-escrito de <i>The Concept of Truth</i> e a derrogação da teoria das categorias semânticas.....	61

Capítulo 5

Critério de verdade

5.1 Preliminares: definição <i>versus</i> critério.....	63
5.2 Critério de verdade nos artigos informais.....	64
5.3 Critério de verdade em <i>The Concept of Truth</i>	67
5.4 Critério de verdade numa classe de estruturas.....	72
5.5 Mais algumas indicações histórico-conceituais.....	77

Considerações finais

Retrospecto.....	79
Prospecto.....	82

Apêndices

<i>Apêndice 1</i> – O paradoxo de Eubúlides-Epimênides.....	i
<i>Apêndice 2</i> – Duas demonstrações do teorema da indefinibilidade da verdade, e sua relação com o primeiro teorema de incompletude de Gödel.....	vi

Índice de abreviações

Referências

RESUMO

Nesta dissertação, oferecemos uma apresentação geral das investigações semânticas de Alfred Tarski, focalizando, em particular, sua contribuição à formalização matemática do conceito de verdade. A apresentação inclui, especialmente, um estudo sobre a monografia *The Concept of Truth in Formalized Languages*, e um estudo sobre a noção Tarskiana de critério de verdade – uma noção relativamente negligenciada na literatura lógico-filosófica.

ABSTRACT

In this work we offer a general presentation of Alfred Tarski's semantical investigations. Particularly we focus on his contribution to the mathematical formalization of the concept of truth. Our presentation includes especially a study of the monograph The Concept of Truth in Formalized Languages, as well as a study of the Tarskian notion of criterion of truth – a relatively neglected notion in logico-philosophical literature.

Introdução

O objetivo desta dissertação é desenvolver uma *apresentação geral das investigações de Tarski sobre a semântica de linguagens formalizadas*, e, em especial, *sobre o conceito de verdade*. Ao lado de Gödel, Tarski figura entre os maiores lógicos do século XX. Aliás, existe uma forte afinidade entre Gödel e Tarski. Assim como os teoremas de incompletude do primeiro representam uma das pérolas da lógica matemática, de profundas conseqüências técnicas e filosóficas (relacionadas ao conceito de *demonstrabilidade*), as investigações semânticas do segundo, em especial, a definição do conceito de verdade em linguagens formalizadas, também representam uma das grandes realizações da lógica matemática, e igualmente tiveram profunda repercussão no meio lógico e filosófico (desta vez, relacionadas ao conceito de *verdade*). Qualquer estudioso da lógica matemática ou da filosofia no século XX, precisa ficar a par das principais contribuições de Gödel e Tarski às ciências formais.

No entanto, um simples manual técnico de lógica matemática, mesmo um daqueles sofisticados, não proporciona uma visão detalhada do contexto em que as principais realizações desses dois lógicos vieram à luz. De fato, isso até o extraviaria de seus propósitos técnicos. Apenas um trabalho complementar, de natureza mais historiográfica ou filosófica, é capaz de proporcionar uma tal visão, com os pormenores relevantes. Nosso propósito aqui é justamente fornecer as informações suplementares que julgamos relevantes para esclarecer o contexto em que se desenvolveram as investigações semânticas de Tarski.

Façamos duas observações preliminares – uma observação geral (sobre a forma) e uma observação específica (sobre o conteúdo) – a respeito do que vamos encontrar nas próximas páginas.

(1) Em primeiro lugar, esta dissertação se inscreve no âmbito da *história* da lógica matemática. Não deve, por conseguinte, ser confundida com uma dissertação de natureza *técnica*, nem com uma dissertação de natureza *filosófica*. Não é nossa intenção desenvolver técnica e formalmente algum tópico ainda não explorado sobre o conceito de verdade. Tampouco é nossa intenção elaborar argumentos filosóficos novos ou contrargumentar com posições filosóficas correntes sobre o conceito de verdade. Agora, evidentemente, dada a própria natureza do tema que escolhemos para esta dissertação, torna-se inevitável a introdução de alguns formalismos (embora não de maneira sistemática) e de algumas considerações filosóficas suaves.

(2) Em segundo lugar, não temos a menor pretensão de exaurir todos os tópicos relacionados com as investigações semânticas de Tarski. Portanto, talvez seja útil dizer quais assuntos não vamos abordar nesta dissertação. John Etchemendy¹ chama a atenção para uma bifurcação que surge a partir do trabalho original de Tarski, e que é interessante assinalar aqui. De um lado, desenvolve-se o ramo da chamada ‘semântica formal’, cuja preocupação maior é investigar a semântica das linguagens naturais (é nesse ramo que encontramos, por exemplo, os trabalhos de Davidson e Montague). De outro lado, desenvolve-se o ramo da chamada ‘análise dos paradoxos semânticos’, cuja preocupação maior é revisar e propor novas soluções a paradoxos como o de Eubúlides-Epimênides, o de Berry, o de Richard, o de Grelling-Nelson, entre outros (é nesse ramo que encontramos, por exemplo, os trabalhos de Kripke, Gupta, Belnap, o próprio Etchemendy ao lado de Barwise etc.). Ambos os ramos, extremamente importantes na atual filosofia da lógica e da linguagem, herdam muito de Tarski. Mas disciplinaremos o ânimo dispersivo: estaremos restritos ao nó da bifurcação, isto é, ao próprio Tarski, o que já é bastante laborioso.

Todo o trabalho está distribuído em cinco capítulos. Os quatro primeiros capítulos estão bastante interligados. O tema central que os unifica é o problema da definição do conceito de verdade, tal como Tarski o abordou. O Capítulo 1 expõe o ambiente lógico-matemático que motivou o desenvolvimento de suas primeiras investigações sobre a semântica de linguagens formalizadas (em especial, sobre o conceito de verdade). O Capítulo 2 delinea o contexto programático geral de formalização da semântica, no qual o problema da definição do conceito de verdade está inserido. O Capítulo 3 apresenta o método desenvolvido por Tarski para resolver esse problema. O Capítulo 4 tenta guiar o leitor por dentro do texto da principal monografia de Tarski sobre o conceito de verdade – *The Concept of Truth in Formalized Languages*. O Capítulo 5 é relativamente independente dos capítulos anteriores. Nele, exploramos um tema um pouco apagado na literatura (seja técnica, seja filosófica, seja historiográfica), a saber, a noção de critério de verdade em Tarski. Estaremos satisfeitos se o resultado final servir como um guia razoável de leitura dos trabalhos de Tarski sobre a semântica de linguagens formalizadas e o conceito de verdade. De fato, essa foi uma das principais motivações pessoais na elaboração deste trabalho.

¹ ETCHMENDY, J. ‘Tarski on truth and logical consequence’, JSL, v. 53, n.º 1, 1988, pp. 51-79

Capítulo 1— *Prelúdio histórico*

SUMÁRIO. A fase ‘pré-Tarskiana’ da teoria de modelos e o caráter informal das noções semânticas. Os seminários de Varsóvia, o método de eliminação dos quantificadores e as investigações semânticas de Tarski. O problema das antinomias semânticas. O diagnóstico de Tarski das antinomias semânticas. Uma nuance conceitual: as noções de verdade-numa-linguagem e verdade-numa-estrutura.

A fase ‘pré-Tarskiana’ da teoria de modelos e o caráter informal das noções semânticas

Antes de Tarski publicar sua famosa monografia *The Concept of Truth in Formalized Languages*², no ano de 1933, a investigação de certos tópicos pertencentes ao que hoje chamamos de ‘teoria de modelos’, já tinha sido encetada havia alguns anos. Para termos uma idéia do que se produzira até então, façamos um rápido escrutínio dos principais resultados alcançados nesse período inicial.

- Em um artigo de 1915, Löwenheim publicou a primeira versão de um importante teorema, parte integrante de uma família de resultados conhecidos como os teoremas de Löwenheim-Skolem (referidos às vezes como os teoremas de Löwenheim-Skolem-Tarski)³. O que Löwenheim demonstrou foi o seguinte:

(1) *Seja s uma sentença elementar (i.e., de primeira ordem). Se s tem um modelo, então s tem um modelo contável.*

- Ainda no mesmo artigo, Löwenheim demonstrou outro resultado de natureza modeloteórica, embora menos relevante que o anterior:

(2) *Seja s uma sentença elementar cujos símbolos de predicado que nela ocorrem são todos monádicos. Se s é verdadeira em todos os modelos finitos, então s é verdadeira em todos os modelos.*

² Doravante, indicaremos essa monografia pelas iniciais ‘CTFL’.

³ ‘Über Möglichkeit im Relativkalkül’ (1915).

- Numa série de artigos (datados de 1920, 1922, 1928 e 1929)⁴, Skolem estendeu o resultado (1) de Löwenheim, dando-lhe o seguinte conteúdo:

(3) *Seja S um conjunto de sentenças elementares. Se S tem um modelo, então S tem um modelo contável.*

- Em 1928, durante os seminários que ministrou na Universidade de Varsóvia, dos quais falaremos mais adiante, Tarski demonstrou pela primeira vez uma forma da denominada versão ‘ascendente’ (*upward*) do teorema de Löwenheim-Skolem, embora não tenha publicado sua prova:

(4) *Seja S um conjunto de sentenças elementares. Se S tem um modelo infinito, então S tem um modelo não-denumerável.*

- Mais ainda, nesse mesmo ano, Tarski demonstrou uma versão mais forte de (4) – versão que, um pouco mais tarde, em 1936, seria provada independentemente por Mal’cev:

(5) *Seja S um conjunto de sentenças elementares. Se S tem um modelo infinito, então S tem um modelo não-denumerável, em cada cardinal transfinito.*⁵

- O grupo de Hilbert em Göttingen também produziu uma série de resultados nos quais estavam envolvidas noções de caráter modelo-teorético. Por exemplo, em um artigo⁶ de 1928, Bernays e Schönfinkel demonstraram, entre outros teoremas similares, a proposição abaixo:

(6) *Se s é uma sentença elementar que pode ser reduzida à forma normal prenex de tipo ‘ $\forall x_1, \dots, \forall x_m M(x_1, \dots, x_m)$ ’, e s tem um modelo com m indivíduos, então s é uma verdade lógica.*

- Na primeira edição dos *Grundzüge der theoretischen Logik*, de 1928, Hilbert e Ackermann lançaram o seguinte problema – naquele momento, em aberto:

⁴ ‘Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen’ (1920); ‘Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre’ (1922); ‘Über die mathematische Logik’ (1928); ‘Über einige Grundlagenfragen der Mathematik’ (1929).

⁵ Cf. VAUGHT, R. ‘Model theory before 1945’, *Proceedings of the Tarskian Symposium*, p. 160, e ‘Alfred Tarski’s work in model theory’, *JSL*, v. 51, n.º 4, 1986, p. 870. Em CTFL, Tarski enuncia *en passant* esse resultado: *[E]very sentence which is correct in one infinite domain is also correct in every other infinite domain (without reference to its cardinal number)* (p. 205).

⁶ ‘Zum Entscheidungsproblem der mathematischen Logik’ (1928).

Seja P um sistema axiomático para o cálculo de predicados de primeira ordem (ou cálculo funcional restrito, ‘engere Funktionenkalkül’, na terminologia dos autores), e seja V o conjunto das sentenças (da linguagem de P) verdadeiras em todos os domínios de indivíduos, i.e., o conjunto das verdades lógicas elementares. Problema: o conjunto dos teoremas de P e o conjunto V são co-extensivos? Dito de outro modo: podemos deduzir em P todas as sentenças de V ?

Esse é o conhecido problema da completude (semântica) da lógica de primeira ordem, que, como sabemos, foi respondido na afirmativa por Gödel em sua tese de doutorado⁷, em 1929. Na tese, Gödel também mostrou que o teorema de completude é equivalente à seguinte proposição:

(7) Seja S um conjunto de sentenças elementares. Se S é consistente, então S tem um modelo.

Bem, os teoremas (1)-(7) de nosso inventário acima (à exceção, talvez, de (2) e (6)) são os exemplos mais significativos dos primeiros resultados modelo-teóricos. Podemos chamar a fase em que eles foram demonstrados, de fase ‘pré-Tarskiana’ da teoria de modelos. Observe que o conceito central em todos os teoremas enunciados, é o conceito de *verdade de uma sentença em um domínio de indivíduos*. É justamente pelo fato de todos esses teoremas partilharem dessa noção, que os incluímos no âmbito da teoria de modelos. Por que resolvemos denominar essa primeira fase de ‘pré-Tarskiana’? Por uma razão bastante simples, envolvendo um fato bem conhecido: antes da publicação de CTFL, nenhuma tentativa de formalizar (no interior de um sistema fundacional matemático) o conceito de sentença verdadeira, havia sido empreendida. Apesar da enorme importância e da grande difusão dos teoremas da fase inicial (principalmente, dos teoremas de Löwenheim-Skolem e, depois, do teorema de completude de Gödel), *o uso que se fazia da noção de verdade era simplesmente intuitivo e informal*. O conceito de verdade era sempre assumido, implicitamente, como um termo semântico primitivo, com um significado bastante óbvio, e com várias propriedades intuitivamente claras (como o princípio de não-contradição, o princípio do terceiro excluído, entre outros). Entretanto, parecia inadmissível para Tarski (cuja conduta intelectual pautava-se nos cânones do estrito rigor formal, típico dos lógicos e filósofos poloneses) que o conceito fundamental envolvido em resultados tão relevantes permanecesse não formalizado. Há uma pequena história relacionada ao teorema de completude de Gödel, lembrada por Kreisel, que ilustra bem esse estado de coisas:

⁷ ‘Über die Vollständigkeit des Logikkalküls’ (1929), reescrita como artigo e publicada nos *Monatshefte für Mathematik und Physik* sob o título ‘Die Vollständigkeit der Axiome des logischen Funktionenkalküls’ (1930).

*According to Mostowski, in a conversation in Tarski's presence, the latter and his students had no confidence in Gödel's paper when they saw the relevant issue of the 'Monatshefte für Mathematik und Physik' in Warsaw. Why? Gödel had not formally defined validity!*⁸

Em CTFL, há uma passagem significativa em que Tarski assinala sua insatisfação com o uso informal do conceito de verdade nos resultados da primeira fase da teoria de modelos:

It is worthy of note that – in spite of the great importance of these terms⁹ for metamathematical investigations – they have hitherto been used in a purely intuitive sense without any attempt to define their meaning more closely¹⁰.

É curioso observar também que essa insatisfação não era sintomática apenas em Tarski. Já Herbrand, à mesma época, invectivava contra o uso informal de noções semânticas nos resultados da fase pré-Tarskiana:

Similar results have already been stated by Löwenheim (1915), but his proofs, it seems to us, are totally insufficient for our purposes. [...] he gives an intuitive meaning to the notion 'true in an infinite domain', hence his proof [...] does not attain the rigor that we deem desirable¹¹.

Um pouco antes da passagem acima, ele anotara:

It is absolutely necessary to adopt such definitions if we want to give a precise sense to the words 'true in an infinite domain', words that have frequently been used without sufficient explanation, and also if we want to justify a

⁸ Kreisel, G. 'Gödel's excursions into intuitionistic logic', *apud* Wolenski, J. 'From intentionality to formal semantics (from Twardowski to Tarski)', *Erkenntnis*, 56, 2002, p.19. O negrito é nosso.

⁹ Com 'these terms', Tarski está-se referindo a termos como 'sentença verdadeira em um domínio de indivíduos', 'sentença verdadeira em um domínio com k indivíduos' e 'sentença verdadeira em todos os domínios de indivíduos', termos que, como ele mesmo chama a atenção, ocupavam um lugar especial nas investigações metamatemáticas então em curso, em particular naquelas desenvolvidas pela escola Hilbertiana de Göttingen; v. CTFL, pp. 199-200.

¹⁰ CTFL, p. 200; uma outra passagem em tom similar encontra-se no final do §4, p. 241.

¹¹ Herbrand, J. 'Investigations in proof theory' (1930), in Heijenoort, J. *From Frege to Gödel*, pp. 558-559. Imediatamente depois, Herbrand critica também o uso informal da noção de 'satisfatibilidade' que Ackermann faz num artigo de 1928.

*proposition proved by Löwenheim, a proposition to which many refer without clearly seeing that Löwenheim's proof is totally inadequate for our purposes and that, indeed, the proposition has no precise sense until such a definition has been given*¹².

Aliás, é importante observar que Tarski, em nota à passagem de CTFL que citamos há pouco, menciona Herbrand como uma exceção à sua observação, pelo menos no que concerne ao conceito de verdade em um domínio *finito* de indivíduos¹³. Mas seja sublinhado um detalhe: nas duas passagens citadas, Herbrand julga ter atribuído um significado formal preciso ao conceito de verdade em um domínio *infinito* de indivíduos, e, no entanto, Tarski, como acabamos de assinalar, reconhece a definição de Herbrand apenas para o caso finito. A razão para tal discrepância, contudo, merece um estudo à parte, e não nos interessa persegui-la aqui.

Nossa discussão inicial sobre a fase pré-Tarskiana da teoria de modelos ficaria certamente incompleta, se deixássemos de fazer menção a dois pontos. Em primeiro lugar, nosso escrutínio de resultados partiu de 1915, ano em que Löwenheim publicou a primeira versão de seu famoso teorema. Entretanto, se formos historicamente mais liberais, podemos considerar algumas investigações anteriores como já relacionadas com a teoria de modelos. Em um artigo de 1904, Veblen lançou mão da noção de categoricidade para caracterizar um sistema de axiomas para a geometria¹⁴. E, como lembram William e Martha Kneale, o conceito de categoricidade já era conhecido por Dedekind em seu *Was sind und was sollen die Zahlen?*, de 1887¹⁵. Por outro lado, não podemos esquecer o artigo pioneiro de Padoa na teoria da definibilidade, de 1900, em que o problema da independência (ou ‘irredutibilidade’, em sua terminologia original) de um sistema de termos primitivos e de um sistema de axiomas, é tratado com o auxílio da noção de modelo¹⁶. Se nos fosse concedido um pouco mais de liberalidade, lembraríamos sem dúvida a introdução do conceito de Universo de Discurso por De Morgan e Boole, e

¹² *Idem*, pp. 552-553.

¹³ CTFL, p. 200, nota.

¹⁴ O artigo de Veblen é ‘A system of axioms for geometry’. Por diversas vezes, Tarski faz remissão à noção de categoricidade introduzida por Veblen: em CTFL, p. 174; em ‘Some observations on the concepts of ω -consistence and ω -completeness’, p. 282, entre outros. Não devemos esquecer que Veblen integrava o grupo dos chamados “american postulate theorists”, do qual Huntington fazia parte. Uma boa apresentação dos trabalhos desse grupo, que já lançava mão de noções semânticas, pode ser lida no artigo de Michael Scanlan ‘Who were the american postulate theorists?’, JSL, v. 56, n. 3, 1991.

¹⁵ Kneale W. & Kneale M. *O Desenvolvimento da Lógica*, pp. 393-394. Trata-se do § 134, em que Dedekind observa que todos os ‘sistemas’ que satisfazem os axiomas de sua teoria dos ‘sistemas simplesmente infinitos’ (entre os quais se encontra o ‘sistema dos números naturais’) são ‘similares’ entre si. Cf. a tradução inglesa *The Nature and Meaning of Numbers*, pp. 95-96.

¹⁶ Padoa, A. ‘Logical introduction to any deductive theory’, in Heijenoort, *op.cit.*, pp. 118-123. Três décadas mais tarde, Tarski iria aprimorar algumas idéias de Padoa em ‘Some methodological investigations on the definability of concepts’, de 1934.

também a análise de Bolzano dos conceitos de analiticidade e consequência lógica (que antecipou, em pontos importantes, a análise que Tarski faria mais tarde).

Nosso segundo ponto diz respeito aos famosos seminários que Tarski ministrou na Universidade de Varsóvia, no final da década de vinte. Em tais seminários, alguns resultados modelo-teóricos importantes foram obtidos por Tarski. Esses resultados, contudo, ainda estão inseridos no que vimos chamando de fase ‘pré-Tarskiana’ da teoria de modelos, e podem ser vistos como a motivação técnica imediata para a formalização da semântica, que se inicia com a publicação do artigo, não tão conhecido, ‘On Definable Sets of Real Numbers’ (DSRN, 1931) e, dois anos mais tarde, do clássico CTFL. Reservamos inteiramente o próximo tópico para apresentar mais pormenorizadamente o conteúdo desses seminários.

Os seminários de Varsóvia, o método de eliminação dos quantificadores e as investigações semânticas de Tarski

Pouca ou, talvez seja mais apropriado dizer, nenhuma atenção especial foi dirigida, até agora (até onde sabemos), à relação do método de eliminação dos quantificadores com o surgimento dos trabalhos de Tarski sobre o conceito de verdade em linguagens formalizadas. Essa relação não é, de fato, direta e evidente, e, até certo ponto, é mesmo acidental. Apenas alguma pesquisa histórica sobre o período que antecede imediatamente a primeira publicação de CTFL ou, para sermos cronologicamente mais precisos, a primeira comunicação pública dos resultados contidos nessa monografia (no ano de 1930, diante da Seção de Lógica da Sociedade Filosófica, em Varsóvia), é capaz de torná-la explícita.

Alguma informação conceitual preliminar sobre esse método nos será útil. Seja Σ um conjunto de sentenças, e sejam duas fórmulas ϕ e γ . Dizemos que ϕ e γ são Σ -equivalentes, se $\Sigma \vdash \phi \leftrightarrow \gamma$. Considere agora uma teoria axiomática T , cujos axiomas denotaremos por T_a , e vejamos esquematicamente como se desenvolve a aplicação do método de eliminação dos quantificadores a T . Todo o processo pode ser descrito em termos da realização de duas tarefas básicas: em primeiro lugar, selecionamos convenientemente um determinado conjunto de fórmulas da linguagem de T , chamado de ‘conjunto básico’; então, em segundo lugar, mostramos que toda fórmula da linguagem de T é T_a -equivalente a uma combinação booleana de fórmulas do conjunto básico (entendendo-se por ‘combinação booleana’ uma fórmula em que ocorrem apenas conectivos proposicionais, i.e., uma fórmula sem quantificadores). A demonstração deste segundo item fornece os materiais para uma prova de completude de T e, colateralmente,

um procedimento efetivo de decisão para T - por conseguinte, realizar a segunda tarefa é o objetivo principal¹⁷.

Segundo Vaught¹⁸, o método de eliminação dos quantificadores remonta a Löwenheim, que o utilizou (embora de maneira não completamente desenvolvida) na solução positiva do problema da decisão para a lógica monádica de primeira ordem, em 1915. Skolem, contudo, foi quem primeiro o aplicou de maneira cabal, na demonstração da completude e da decidibilidade da teoria das álgebras de conjuntos, num artigo de 1919. Depois, Langford o empregou em suas investigações sobre a teoria elementar das ordens densas e sobre a teoria elementar das ordens discretas com primeiro (e sem último) elemento, em trabalhos de 1926 e 1927¹⁹.

Nos anos de 1927 a 1929, Tarski ministrou uma série de seminários na Universidade de Varsóvia, centrados, em sua grande maioria, no método de eliminação dos quantificadores²⁰. Vários resultados famosos foram obtidos nesse período: Tarski estendeu o resultado de Langford à teoria elementar de todas as ordens discretas, demonstrou a completude e a decidibilidade da álgebra elementar dos números reais, da geometria elementar e da teoria das álgebras booleanas com átomos²¹; além disso, há o conhecido resultado de Presburger²², que, sob a orientação de Tarski, mostrou que a aritmética aditiva é completa e decidível.

Além desses resultados, Tarski descobriu novos fatos relacionados com o método de eliminação dos quantificadores. Diferentemente dos resultados anteriores, todos de natureza sintática, alguns desses fatos diziam respeito agora a determinadas propriedades semânticas das teorias investigadas. Podemos sumariá-los, mais ou menos sistematicamente, da seguinte maneira. Seja T uma teoria de primeira ordem que admite eliminação dos quantificadores (por conseguinte, uma teoria completa). Por meio da aplicação do método de eliminação dos quantificadores a T , Tarski mostrou como é possível obter uma classificação de todas as extensões completas de T . Essa foi uma das novidades. Agora, seja α um modelo de T . A outra novidade importante descoberta por Tarski, foi esta: o conjunto das sentenças da linguagem de T verdadeiras em α , $T(\alpha)$, é uma extensão completa de T . Considere dois modelos α e β quaisquer de T . Tarski

¹⁷ Para detalhes técnicos sobre o método de eliminação dos quantificadores, v. CHANG C. C. & KEISLER H. J. *Model theory*, pp. 49-58.

¹⁸ VAUGHT, R. 'Alfred Tarski's work in model theory', p. 870; 'Model theory before 1945', pp. 159-160.

¹⁹ Löwenheim, *op. cit.* (1915); Skolem, 'Untersuchungen über die Axiome des Klassenkalküls und über Produktions- und Summationsprobleme, welche gewisse Klassen von Aussagen betreffen' (1919); Langford, 'Some theorems on deducibility' (1926) e 'Theorems on deducibility' (1927).

²⁰ Cf. Vaught, referências na nota 17 acima.

²¹ A publicação madura de parte desses resultados está contida em outra monografia famosa de Tarski: *A Decision Method for Elementary Algebra and Geometry* (1948), preparada por McKinsey.

²² Presburger, 'Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt' (1930).

introduz, então, no contexto dessas descobertas, o conceito de equivalência elementar: α e β são *elementarmente equivalentes*, $\alpha \equiv \beta$, se, e somente se, $T(\alpha) = T(\beta)$. Ora, é fácil ver que a relação de equivalência elementar é uma relação de equivalência. Segue, então, das duas novidades acima, que podemos particionar a classe de todos os modelos de T em tantas classes de modelos elementarmente equivalentes, quantas são as extensões completas de T. Ou seja, **com base na classificação das extensões completas de T, obtemos paralelamente uma classificação de todos os modelos de T que são indistinguíveis modulo equivalência elementar**²³.

Considere, p.e., a teoria elementar das ordens lineares densas (que denotaremos por Td). Recuperando os resultados alcançados por Langford, que, como dissemos, demonstrou, usando o método de eliminação dos quantificadores, a decidibilidade e a completude de Td, Tarski mostrou que existem exatamente quatro extensões completas de Td (com primeiro e último elementos, com primeiro e sem último elementos, sem primeiro e com último elementos, e sem primeiro e último elementos). Usando a noção de equivalência elementar e o fato de que, para qualquer modelo α de Td, $T(\alpha)$ é uma extensão completa de Td, Tarski deduziu que a classe de todos os modelos de Td está particionada em quatro classes de modelos elementarmente equivalentes²⁴.

Note que tais resultados envolvem noções inequivocamente semânticas: a noção de equivalência elementar entre modelos, que repousa na noção de modelo de um conjunto de sentenças, que, por sua vez, repousa na noção de verdade de uma sentença em um domínio de indivíduos. Além disso, envolvem fatos que dependem essencialmente de uma formalização do conceito de verdade, como o fato de que, para todo modelo α , $T(\alpha)$ é um sistema dedutivo *completo* (com respeito à negação). Pareceu, então, a Tarski, uma tarefa preliminar indispensável definir formal e precisamente tais noções, a fim de demonstrar rigorosamente os resultados a que ele havia chegado. Um bom indício de que esse foi realmente o caso, é o fato de que foi preciso aguardar mais de meia década para publicar tais resultados. Com efeito, é apenas em ‘Foundations of the Calculus of Systems’ (FCS, 1935-1936) que eles aparecem publicados pela primeira vez, depois de Tarski já ter à disposição uma formalização satisfatória das principais noções semânticas neles envolvidas. Tarski reforça esse indício na nota historiográfica a FCS, em que, referindo-se à classificação dos modelos elementarmente equivalentes de uma teoria completa, diz o seguinte:

[...] I was able to state these facts in a correct and precise form only with the help of the methods which were developed

²³ Cf. DONER, J. & HODGES, W. ‘Alfred Tarski and decidable theories’, JSL, v. 53, n.º 1, 1988, pp. 32-33.

²⁴ Tarski, A. ‘Foundations of the calculus of systems’ (1935-6), pp. 372-375.

*in [On definable sets of real numbers] and [The concept of truth in formalized languages], during the years 1929-30*²⁵.

No entanto, já em CTFL, Tarski antecipa alguma coisa: ele demonstra alguns teoremas relativos ao cálculo de classes²⁶ (ou ‘álgebra atomística da lógica’, como ele também o denomina), uma das teorias investigadas durante os seminários de Varsóvia, e cuja linguagem é utilizada para ilustrar a formalização do conceito de verdade. Um resultado que Tarski menciona, mas que não demonstra, diz respeito à classificação de todas as extensões completas do cálculo de classes (mais precisamente, de uma extensão do sistema apresentado em CTFL)²⁷. Não há alusão, contudo, à classificação dos modelos elementarmente equivalentes do cálculo de classes.

Para encerrar este tópico, não deixemos de lembrar, *en passant*, que data também da época dos seminários de Varsóvia uma demonstração de Tarski (não publicada) de duas formas da versão ascendente do teorema de Löwenheim-Skolem, a que já nos referimos no tópico anterior (exemplos (4) e (5)).

O problema das antinomias semânticas

Até esta altura, nosso propósito foi chamar enfaticamente a atenção para o seguinte fato: Tarski deparou-se com uma constelação de noções semânticas (entre elas, a noção de verdade) informalmente utilizadas em importantes resultados metamatemáticos (alguns dos quais obtidos por ele mesmo e seu grupo em Varsóvia). A necessidade de formalizá-las, depurando-as de certa obscuridade e vagueza de conteúdo, foi uma das motivações imediatas para que ele buscasse algum método de defini-las. Mas essa não foi a motivação mais relevante. Mais do que o uso informal das noções semânticas, a existência de paradoxos em que elas estavam essencialmente envolvidas, foi o motivo principal que levou Tarski a estabelecer os fundamentos teóricos da semântica de linguagens formalizadas. E isso é bastante compreensível. Enquanto não se mostrasse que as noções semânticas podem, sob condições controladas, ser usadas sem o risco de antinomias, não havia garantia de que os resultados modelo-teóricos obtidos até então (dos teoremas de Löwenheim-Skolem ao teorema de completude de Gödel) estavam isentos de contradição.

Mas antes de examinar a profilaxia desenvolvida por Tarski contra as antinomias semânticas, convém tecer algumas considerações gerais sobre paradoxos. Um paradoxo

²⁵ FCS, p. 383.

²⁶ Trata-se da teoria das álgebras booleanas com átomos, a que nos referimos mais acima. Para os axiomas apresentados por Tarski em CTFL, cf. Capítulo 4 desta dissertação.

²⁷ Cf. CTFL, p. 205, nota 2.

pode ser caracterizado da seguinte maneira: temos um argumento $p_1, \dots, p_n \therefore q$, tal que (i) a aceitação das premissas $p_1, \dots, p_n$ é aparentemente inócua; (ii) a dedução de q a partir de $p_1, \dots, p_n$ é logicamente válida, e (iii) a conclusão q é uma sentença contraditória²⁸. Costuma-se aplicar a denominação ‘paradoxo’ à conclusão q , mas não há empecilho em aplicar essa denominação ao próprio argumento descrito acima – de fato, esse segundo uso parece-nos mais natural. Observe-se ainda que, normalmente, entre as premissas $p_1, \dots, p_n$, figura uma sentença p que, de algum modo, está no cerne da construção do paradoxo, e, por vezes, também se aplica a denominação ‘paradoxo’ a essa sentença p .

Segundo uma classificação conhecida, originada com Ramsey²⁹, os paradoxos dividem-se em duas famílias, a saber, a dos paradoxos *lógicos* e a dos paradoxos *semânticos*. Como nosso interesse recai sobre os paradoxos semânticos, não nos estenderemos muito sobre a primeira família. Os paradoxos lógicos situam-se todos no interior da teoria intuitiva dos conjuntos. As antinomias que mais se destacam nessa família são o paradoxo de Cantor (do cardinal do conjunto de todos os conjuntos), o paradoxo de Burali-Forti (do conjunto de todos os ordinais) e o paradoxo de Russell. Desses três paradoxos, o de Russell é o mais saliente, não só por sua simplicidade em relação aos outros, que envolvem algum grau de elaboração técnica, como também por ter solapado o projeto logicista Fregeano, ao revelar a inconsistência do sistema das *Grundgesetze der Arithmetik*.

Vejamos, então, apenas o paradoxo de Russell. Intuitivamente, é concebível que muitos conjuntos não pertençam a si mesmos: o conjunto das vogais, por exemplo, não é elemento de si mesmo. É também concebível que alguns conjuntos pertençam a si mesmos: o conjunto dos entes abstratos, sendo ele próprio um ente abstrato, é um elemento de si mesmo. Agora, dado que qualquer predicado define um conjunto (assumindo-se a versão irrestrita do axioma de compreensão), podemos construir um conjunto R cujos elementos são todos os conjuntos que não pertencem a si mesmos, i.e.,

$$R =_{df.} \{x: x \notin x\}$$

Ora, sendo R um conjunto, convém perguntar se ele pertence ou não a si mesmo. Ou R pertence a si mesmo, ou R não pertence a si mesmo. Por um lado, se R pertence a R , então R não satisfaz o predicado que o define; logo, R não pertence a R . Por outro lado,

²⁸ Podemos elasticar essa caracterização, substituindo (ii) por (ii)': a *inferência* (dedutiva/indutiva) de q a partir de $p_1, \dots, p_n$ é *correta*; e substituindo (iii) por (iii)': a conclusão q é uma sentença intuitivamente inaceitável (não sendo necessário que q seja contraditória).

²⁹ ‘The foundations of mathematics’ (1925). A classificação de Ramsey, contudo, não é unanimemente aceita, a começar por Russell, que atribuía a origem de todos os paradoxos lógicos e semânticos à violação do chamado ‘Princípio do Círculo Vicioso’.

se R não pertence a R, então R satisfaz o predicado que o define; logo, R pertence a R. Portanto, R pertence a R, se, e somente se, R não pertence a R – o que é absurdo.

A história subsequente à descoberta dos paradoxos lógicos é bem conhecida: em 1908, Russell³⁰ propôs um antídoto bastante radical contra essas antinomias, a teoria ramificada dos tipos (que, alguns anos depois, foi substituída por uma versão mais moderada, mas ainda radical – a teoria simples dos tipos de Ramsey e Chwistek); no mesmo ano, Zermelo³¹ propôs uma via que mais tarde se tornaria canônica, a via da axiomatização da teoria intuitiva dos conjuntos (várias axiomáticas famosas foram elaboradas, a começar pela teoria do próprio Zermelo, aperfeiçoada depois por Fraenkel; há também a versão de Bernays, Gödel e von Neumann, a teoria de Kelley e Morse, e os sistemas de Quine em *Mathematical Logic* e *New Foundations*). Mas encerremos aqui o tópico dos paradoxos lógicos, e passemos ao que mais nos interessa, que é a segunda família de paradoxos.

Os paradoxos semânticos caracterizam-se pelo fato de envolverem noções semânticas, como os conceitos de verdade e definibilidade. Dos paradoxos semânticos, o mais antigo e famoso é o Mentiroso, que envolve o conceito de verdade, e que é atribuído ao megárico Eubúlides e ao cretense Epimênides³². Uma versão do paradoxo de Eubúlides-Epimênides é desenvolvida da seguinte maneira. Considere a sentença abaixo:

(EE) Esta sentença não é verdadeira.

Convém perguntar se (EE) é verdadeira ou não. Ou a sentença (EE) é verdadeira, ou a sentença (EE) não é verdadeira. Por um lado, se a sentença (EE) é verdadeira, então o que ela diz é verdadeiro. Ela diz que ela mesma, (EE), não é verdadeira. Logo, a sentença (EE) não é verdadeira. Por outro lado, se a sentença (EE) não é verdadeira, então o que ela diz não é verdadeiro. Ela diz que ela mesma, (EE), não é verdadeira. Logo, ela é verdadeira. Portanto, a sentença (EE) é verdadeira, se, e somente se, a sentença (EE) não é verdadeira – o que é absurdo.

O paradoxo de Eubúlides-Epimênides é o principal obstáculo a uma formalização consistente do conceito de verdade. No Apêndice 1, apresentamos uma versão mais elaborada desse paradoxo, devida a Lukasiewicz, e que serviu de modelo para a análise de Tarski do problema geral das antinomias semânticas. Dois outros paradoxos importantes são os paradoxos de Berry e Richard, que envolvem o conceito de definibilidade semântica. Esse conceito é tão importante, do ponto de vista

³⁰ ‘Mathematical logic as based on the theory of types’ (1908).

³¹ ‘Untersuchungen über die Grundlagen der Mengenlehre I’ (1908).

³² A atribuição do Mentiroso a Epimênides baseia-se numa passagem bíblica: ‘Um deles [dos cretenses], seu próprio profeta, disse: os cretenses são sempre mentirosos’, Epístola a Tito, 1:12.

metamatemático, quanto o conceito de verdade, e as investigações de Tarski sobre a noção de definibilidade são tão relevantes, quanto suas investigações sobre a noção de verdade. Vejamos uma versão do paradoxo de Berry. Considere a seguinte frase declarativa:

(B) x é o menor número natural não definível por uma frase declarativa em português com menos de noventa letras.

Somos compelidos a admitir a existência de um número b que satisfaz essa frase, pelas razões aduzidas a seguir: uma vez que o alfabeto da língua portuguesa consiste em um conjunto finito de letras, o número de concatenações com menos de noventa letras é finito. Assim, o conjunto D dos números naturais definíveis por frases declarativas da língua portuguesa com menos de noventa letras, é finito. Ora, sendo finito, seu complementar em N é não-vazio. Como N é bem ordenado pela ordem usual $\leq$, o complementar de D em N possui menor elemento. Esse elemento é precisamente b , o menor número natural não definível por uma frase declarativa em português com menos de noventa letras, que, no entanto, foi definido por uma frase declarativa em português com oitenta e nove letras – o que é absurdo.

Um outro paradoxo semântico famoso é o de Grelling-Nelson, que envolve a noção de ‘ser verdadeiro de si mesmo’. Dizemos que uma expressão e é ‘autológica’, se, e somente se, e é verdadeira de si mesma; do contrário, dizemos que a expressão e é ‘heterológica’. Por exemplo, o adjetivo ‘mamífero’ é heterológico, pois não é verdadeiro de si mesmo – o adjetivo ‘mamífero’ não é mamífero. Por outro lado, o adjetivo ‘polissilábico’ é autológico, pois é verdadeiro de si mesmo – o adjetivo ‘polissilábico’ é polissilábico. Mas convém perguntar se o adjetivo ‘heterológico’ é ou não heterológico. Ou ‘heterológico’ é heterológico, ou ‘heterológico’ não é heterológico. Por um lado, se ‘heterológico’ é heterológico, então, por definição, ‘heterológico’ não é verdadeiro de si mesmo – por conseguinte, ‘heterológico’ não é heterológico. Por outro lado, se ‘heterológico’ não é heterológico, então ‘heterológico’ é autológico, e, por definição, verdadeiro de si mesmo – por conseguinte, ‘heterológico’ é heterológico. Conclusão: o adjetivo ‘heterológico’ é heterológico, se, e somente se, o adjetivo ‘heterológico’ não é heterológico – o que é absurdo.

Como solucionar um paradoxo? Voltemos ao esquema $p_1, \dots, p_n \therefore q$, com o qual iniciamos este tópico. O fato de a sentença q , deduzida de forma correta de premissas aparentemente inocentes, ser contraditória, é sintoma de que algo está errado. Mas o problema não aparece na superfície do argumento paradoxal: a dedução de q é logicamente válida (portanto, não estamos diante de uma falácia), e não há, propriamente

falando, premissas falsas entre $p_1, \dots, p_n$ (e é por isso que sua aceitação parece inócua). O problema está num nível mais profundo, e, para detectá-lo e resolvê-lo, temos que: (a) explicitar todas as pressuposições que foram admitidas de modo a tornar possível a construção do argumento antinômico, e (b) identificar as pressuposições que devem ser eliminadas, isto é, aquelas que são a causa efetiva do paradoxo.

O diagnóstico de Tarski das antinomias semânticas

Tarski dedicou dois artigos específicos ao problema das antinomias semânticas, cada um dos quais voltado para uma noção semântica particular. O primeiro artigo publicado foi ‘On Definable Sets of Real Numbers’ (DSRN, 1931), em que ele mostra como usar formal e controladamente o conceito de definibilidade semântica³³, que os paradoxos de Richard (e Berry) haviam posto em xeque, e cuja legitimidade, por conseguinte, era vista como duvidosa pelos matemáticos e lógicos mais exigentes. O segundo foi CTFL, em que ele mostra como formalizar consistentemente o conceito de verdade, um conceito até certo ponto espúrio, na medida em que está envolvido no paradoxo de Eubúlides-Epimênides. Além desses dois artigos, de caráter formal, Tarski publicou três outros, de caráter informal e de divulgação, que contêm considerações sobre os paradoxos semânticos. No pequeno artigo ‘The Establishment of Scientific Semantics’ (ESS, 1936), o problema das antinomias semânticas é tratado sob uma perspectiva mais geral, e está inserido no amplo contexto do programa de formalização da semântica. Em ‘The Semantic Conception of Truth and the Foundations of Semantics’ (SCTFS, 1944) e ‘Truth and Proof’ (TP, 1969), Tarski reapresenta a análise do paradoxo de Eubúlides-Epimênides, já efetuada em CTFL. A apresentação desse paradoxo em SCTFS (e mesmo em TP) é, em certos aspectos, mais cuidadosa do que a apresentação em CTFL.

O exame mais pormenorizado que Tarski fez de uma antinomia foi o exame do paradoxo de Eubúlides-Epimênides. Em DSRN, ele se limitou apenas a expor a profilaxia contra paradoxos como o de Richard (e Berry), deixando de lado uma análise de tais antinomias – uma análise que justificasse essa mesma profilaxia. De qualquer modo, a análise particular de cada uma das antinomias semânticas, depois da publicação de CTFL, pareceu desnecessária a Tarski, já que o exame do paradoxo de Eubúlides-Epimênides poderia ser tomado como modelo para um diagnóstico geral dos paradoxos semânticos, e para as medidas a serem tomadas para evitá-los. Deixamos para o Apêndice 1 uma

³³ Para fins de ilustração, Tarski considera o caso particular do conceito de definibilidade em $\mathfrak{R} = \langle \mathbb{R}; \leq, +, 1 \rangle$, sendo $\mathbb{R}$ o conjunto dos números reais (i.e., o grupo aditivo dos reais, com a ordem usual). A linguagem escolhida por Tarski baseia-se na versão simples da teoria dos tipos, e os conjuntos definíveis focalizados são os conjuntos ‘de primeira ordem’, i.e., conjuntos de números reais (e não conjuntos de conjuntos de números reais etc.).

apresentação mais detalhada da análise de Tarski desse paradoxo. Cumpre, no entanto, fazer aqui algumas considerações gerais.

Como já observamos, a solução de um paradoxo exige, em primeiro lugar, o exame das assunções implícitas que são relevantes para a sua produção, e, em segundo lugar, a eliminação das assunções que realmente conduzem a ele. Vejamos os resultados a que Tarski chega, após examinar o Paradoxo de Eubúlides-Epimênides, e vejamos qual método ele sugere para que o evitemos.

Dois pressupostos fundamentais foram assumidos. Em primeiro lugar, assumiu-se a validade das leis ordinárias da lógica clássica (os princípios de não-contradição e terceiro excluído, a lei de Leibniz etc.). Por conta das consequências drásticas de uma rejeição da lógica clássica, Tarski julga insensato recusar esse pressuposto. Ele não considera alternativas não-clássicas. Resta, então, a exclusão do segundo pressuposto identificado.

O fulcro do paradoxo de Eubúlides-Epimênides – e, de um modo geral, de qualquer antinomia semântica – está no que Tarski chama de ‘fechamento semântico’ da linguagem em que o paradoxo é construído. Esse é o segundo pressuposto. Intuitivamente, linguagens semanticamente fechadas são linguagens que permitem qualquer forma de auto-referência, ou, mais precisamente, permitem que sentenças a elas pertencentes atribuam a si mesmas quaisquer propriedades. Entre estas últimas, figuram, em especial, as propriedades semânticas, concernentes, por exemplo, ao conceito de verdade (como no paradoxo de Eubúlides-Epimênides) ou ao conceito de definibilidade (como nos paradoxos de Richard e Berry). Para que a auto-referência semântica seja possível, é necessário que a linguagem contenha pelo menos um nome para cada uma de suas expressões. Além disso, é necessário que a linguagem contenha termos que expressem propriedades semânticas, como ‘é verdadeiro’, ‘denota’ etc. Nem toda auto-referência é daninha, e mesmo algumas sentenças semanticamente auto-referentes, como ‘esta sentença é verdadeira’, são completamente inócuas. Mas, de uma maneira geral, o fenômeno da auto-referência semântica, típico de linguagens semanticamente fechadas, engendra antinomias. Por essa razão, Tarski não hesita em concluir que, se quisermos afastar o risco de antinomias semânticas, temos que evitar a todo custo linguagens semanticamente fechadas. De acordo com ele, qualquer linguagem natural (como o português, o inglês, o polonês etc.) é semanticamente fechada, permitindo o aparecimento de paradoxos semânticos como o de Eubúlides-Epimênides, o que as torna essencialmente inconsistentes. Em virtude dessa característica, as linguagens naturais revelam-se inidôneas para servir de suporte a uma definição satisfatória da noção de verdade.

Pois bem. Uma vez detectada a causa do paradoxo de Eubúlides-Epimênides (e de outros da mesma família), que tipo de medidas devem ser tomadas? A resposta está no estabelecimento de níveis lingüísticos. A linguagem no âmbito da qual o conceito de verdade (e outros conceitos de natureza semântica) é definido, não pode ser a linguagem para cujas sentenças esse conceito é definido. Claro, uma vez afastado o fechamento semântico desta última linguagem, o nome de suas expressões, bem como os termos semânticos a ela concernentes, são legados a uma outra linguagem. Por conseguinte, são estabelecidos pelo menos dois níveis lingüísticos: de um lado, o nível (inferior) da linguagem cuja semântica está sendo investigada, i.e., o nível da linguagem que é objeto de investigação – a ‘linguagem-objeto’; de outro, o nível (superior) da linguagem da teoria em cujo âmbito se desenvolve o estudo da semântica da linguagem-objeto – a ‘metalinguagem’. A distinção não é absoluta: a própria metalinguagem pode ser a linguagem-objeto de uma linguagem de nível superior. Se quisermos investigar a semântica da metalinguagem, temos, pelas mesmas razões anteriores, de galgar mais um nível lingüístico e ascender à ‘meta-metalinguagem’, e assim por diante. No próximo capítulo, voltaremos a falar sobre níveis lingüísticos, e acrescentaremos oportunamente mais detalhes sobre a estrutura da metalinguagem. É importante observar que a distinção entre níveis lingüísticos já está presente em artigos não voltados ao problema das antinomias semânticas, publicados antes de DSRN e CTFL. Implicitamente, ela está presente em ‘On Some Fundamental Concepts of Metamathematics’ (1930) e em ‘Fundamental Concepts of the Methodology of the Deductive Sciences’ (1930). Explicitamente, ela está presente em ‘Investigations into the Sentential Calculus’ (1930)³⁴.

Uma nuance conceitual: as noções de verdade-numa-linguagem e verdade-numa-estrutura

Até agora, descuramos de chamar a atenção para um fato importante. Estivemos o tempo todo falando, pura e simplesmente, na maior parte das vezes, em ‘o conceito de verdade’, mas essa forma de falar é, pelo menos para Tarski, pouco precisa (embora conste do título de sua famosa monografia), e pode dar a impressão de que ele está tentando formalizar algo pomposo e cheio de majestade como *a* Verdade. De fato, as coisas são bem mais modestas (e por isso, para muitos filósofos, decepcionantes). O que Tarski pretende definir é, antes, o conceito de ‘sentença verdadeira’, e uma vez que toda sentença é sempre um item pertencente a uma determinada linguagem (em contraste com

³⁴ Cf. pp. 39-40.

itens lingüísticamente apátridas, como proposições), o que Tarski pretende definir é, falando mais rigorosamente, o conceito de ‘sentença verdadeira numa determinada linguagem’³⁵. Os lógicos habituados com a moderna terminologia da teoria de modelos acharão estranho esse conceito, pois sua expectativa seria ler ‘sentença (de uma determinada linguagem) verdadeira numa determinada estrutura’. Em outras palavras, os lógicos atuais não falariam em ‘verdade-numa-linguagem’, mas em ‘verdade-numa-estrutura’. Mas será esse um descuido meramente terminológico de Tarski?

Wilfrid Hodges³⁶ chamou a atenção para um detalhe formalmente importante nos trabalhos de Tarski sobre o conceito de verdade em linguagens formalizadas. Esse detalhe diz respeito à ausência em CTFL do conceito de verdade-numa-estrutura. A omissão desse conceito se repete nos artigos informais de divulgação ESS, SCTFS e TP. Costuma-se dizer que o conceito sobre o qual Tarski se debruça nesses textos é o conceito de verdade *simpliciter*, ou verdade-numa-linguagem. A primeira ocorrência de uma definição formal do conceito de verdade-numa-estrutura, dá-se em ‘Arithmetical Extensions of Relational Systems’ (AERS, 1957)³⁷, escrito em parceria com Robert Vaught (o que Tarski chama de ‘sistema relacional’ é o que se entende, na terminologia corrente dos dias de hoje, por ‘estrutura’). Antes de tal artigo, contudo, a noção aparece informalmente esboçada no primeiro capítulo de *Undecidable Theories* (UT, 1953)³⁸, livro escrito em parceria com Robinson e Mostowski. Poucos anos antes, no entanto, Leon Henkin e John Kemeny já se haviam adiantado a essa tarefa, o primeiro em ‘The Completeness of First-Order Functional Calculus’ (1949), e o segundo em ‘Models of Logical Systems’ (1948). Hodges sugere que isso se deve à natureza das linguagens formais utilizadas por Tarski nesse primeiro período. Em contraste com as linguagens usadas no âmbito da teoria de modelos (linguagens de primeira ordem, em especial), Tarski se vale extensivamente de linguagens denominadas por Hodges de ‘linguagens Frege-Peano’. Em tais linguagens, o significado dos símbolos (incluindo os quantificadores) é fixado implicitamente como parte da própria linguagem. Numa linguagem Frege-Peano, uma sentença é incondicionalmente verdadeira ou falsa; numa linguagem modelo-teorética, a verdade ou falsidade de uma sentença depende, é claro, da estrutura que lhe serve de suporte de referência. Embora seja um problema delicado identificar as razões por que Tarski não partiu desde o princípio da definição de verdade

³⁵ Detalhe importante: o ‘L’ de ‘sentença verdadeira em L’, no âmbito de CTFL, denota sempre a linguagem de uma teoria especificamente considerada (como a linguagem do cálculo de classes, da teoria elementar das ordens lineares densas etc.), não sendo nunca uma linguagem qualquer, não especificada. Mais precisamente falando: o ‘L’ de ‘sentença verdadeira em L’ é sempre uma constante, e nunca uma variável que percorre uma classe de linguagens. Apenas no final de CTFL (pp. 263-4) é que Tarski se reporta brevemente a esse conceito “translingüístico” de verdade, e às dificuldades a ele associadas.

³⁶ HODGES, W. ‘Truth in a structure’, *Proceedings of the Aristotelian Society*, NS 86, 1986.

³⁷ AERS, pp. 84-85.

³⁸ UT, pp. 8-9.

em uma estrutura, e por que só veio a fornecê-la duas décadas depois da definição original, não há maiores dificuldades técnicas no trânsito do contexto *simpliciter* para o contexto modelo-teorético. A única coisa que temos a fazer é simplesmente explicitar a interpretação da linguagem em uma dada estrutura, e proceder segundo o método original de Tarski³⁹. Com isso, todas as noções semânticas definidas originalmente no contexto *simpliciter*, migram automaticamente para o contexto modelo-teorético usual.

Feitas essas considerações, advertimos que continuaremos a utilizar a expressão ‘o conceito de verdade’, por razões de economia estilística, no lugar de ‘o conceito de sentença verdadeira numa determinada linguagem’. •

³⁹ É o que faz, p.e., Mendelson, em *Introduction to Mathematical Logic*, pp. 50-2. Cf. também observação de Hodges, *op.cit.*, p. 145.

Capítulo 2 – *O programa Tarskiano de formalização da semântica*

SUMÁRIO. Formalização – definições e axiomatizações. Correção formal de uma definição qualquer. Correção formal de uma definição de verdade (1). Correção formal de uma definição de verdade (2). Adequação material de uma definição qualquer. Adequação material de uma definição de verdade.

Formalização – definições e axiomatizações

Considerando o que foi exposto no capítulo anterior, não é difícil perceber que a tarefa de formalizar o conceito de verdade está inserida num programa mais geral que Tarski assumiu: *formalizar, de maneira consistente, toda a semântica de linguagens formalizadas*. Agora, implicitamente, estivemos o tempo todo associando a idéia de “formalização” com a idéia de “definição” – não por acaso, mas porque o programa de Tarski é conhecido principalmente pelos métodos elaborados para a definição de noções semânticas. E, de fato, nossa ênfase, tanto neste capítulo, como no próximo, recairá sobre definições. Entretanto, é preciso fazer algumas considerações cuidadosas aqui, porque *a contribuição de Tarski à formalização da semântica se estende para além de seus métodos definicionais*. Veremos logo mais que, além da via definicional, uma outra via se abre para o empreendimento de formalização, a saber, a via da axiomatização.

(i) A formalização geral da semântica

Três artigos técnicos revelam o engajamento de Tarski num programa de formalização geral da semântica de linguagens formalizadas. A série inicia-se com ‘On Definable Sets of Real Numbers’. O §1 desse artigo abre-se com as seguintes palavras:

The problem set in this article belongs in principle to the type of problems which frequently occur in the course of mathematical investigations. Our interest is directed towards a term of which we can give an account that is more or less precise in its intuitive content, but the

meaning of which has not at present been rigorously established, at least in mathematics.[...] ⁴⁰

Nesse trecho, as palavras de Tarski refletem sua insatisfação com o uso informal de conceitos em investigações de uma determinada disciplina (matemática ou lógica), como enfatizamos no início do Capítulo 1. Ele continua:

[...] We then seek to construct a definition of this term which, while satisfying the requirements of methodological rigour, will also render adequately and precisely the actual meaning of the term.[...] ⁴¹

Quando Tarski escreve “satisfying the requirements of methodological rigour”, ele está a se referir ao que, na maioria das vezes, chama de condição de *correção formal* de uma definição, e quando ele escreve “render adequately and precisely the actual meaning of the term”, ele está a se referir ao que, também na maioria das vezes, chama de condição de *adequação material* de uma definição. Ele continua, exemplificando:

[...] It was just such problems that the geometers solved when they established the meaning of the terms ‘movement’, ‘line’, ‘surface’, or ‘dimension’ for the first time.[...] ⁴²

E então, finalmente, ele apresenta o objetivo de seu artigo:

[...] Here I present an analogous problem concerning the term ‘definable set of real numbers’.⁴³

Não haveria qualquer problema se, no lugar da expressão “definable set of real numbers”, inseríssemos qualquer outro termo referente a uma noção semântica informalmente utilizada (por exemplo, a expressão “true sentence”). O objetivo de Tarski, em última instância, é bem mais geral, como já assinalamos: formalizar consistentemente toda a semântica de linguagens formalizadas. O conceito de definibilidade semântica de um conjunto de números reais, é apenas um caso particular selecionado para ilustrar o seu método geral. E as palavras acima citadas bem que calhariam na abertura de CTFL, o

⁴⁰ DSRN, pp. 111-112.

⁴¹ *Idem*, p. 112.

⁴² *Ibidem*, p. 112.

⁴³ *Ibidem*, p. 112.

segundo artigo da série. Essa última monografia, em que a formalização da semântica é exemplificada com a importante definição do conceito de verdade, inicia-se com as seguintes palavras:

*The present article is almost wholly devoted to a single problem – **the definition of truth**. Its task is to construct – with reference to a given language – **a materially adequate and formally correct definition of the term ‘true sentence’**.*⁴⁴

Um outro artigo que revela o compromisso de Tarski com a formalização da semântica é ‘On the Concept of Logical Consequence’ (CLC, 1936), o terceiro da série. O problema geral é o mesmo: formalizar uma determinada noção semântica intuitiva. Dessa vez, trata-se de definir o conceito de consequência lógica (em sua acepção semântica). Há apenas três pequenas diferenças entre CLC, de um lado, e DSRN e CTFL, do outro. Em primeiro lugar, nestes últimos, o problema consiste não só em formalizar os conceitos intuitivos de definibilidade e verdade, mas também, como vimos, em neutralizar seu potencial antinômico. Em CLC, o problema é mais restrito, e consiste simplesmente em formalizar a noção intuitiva de consequência lógica, já que não há paradoxos envolvendo esse conceito. Além do mais, diga-se de passagem, enquanto as análises conceituais de DSRN e CTFL mostraram-se, em certo sentido, bem sucedidas, a definição de consequência lógica oferecida por Tarski em CLC não foi completamente feliz, e ele mesmo alimentava dúvidas a seu respeito⁴⁵. Mas esse tema não nos interessa aqui.

Entre esses três artigos, podemos interpolar ainda dois outros trabalhos de Tarski, os quais evidenciam também seu engajamento com os fundamentos formais da semântica. Trata-se de ‘The Establishment of Scientific Semantics’ (ESS, 1936) e ‘The Semantic Conception of Truth and the Foundations of Semantics’ (SCTFS, 1944). Ambos são artigos informais, com caráter de divulgação de suas pesquisas. O primeiro apresenta um delineamento geral do programa de formalização da ‘semântica científica’ (ou semântica da linguagem das ciências – especialmente, as formais), bastante sintonizado com as idéias centrais do positivismo lógico (e de Carnap, em particular, que seria extremamente

⁴⁴ CTFL, p. 152.

⁴⁵ A propósito, cf. Etchemendy, J. ‘Tarski on truth and logical consequence’, JSL, v. 53, n.º 1, 1988, pp. 51-79. Tarski registra suas dúvidas na seguinte passagem: *I am not at all of the opinion that in the result of the above discussion the problem of a materially adequate definition of the concept of consequence has been completely solved* [CLC, p. 418]. Tarski também contende sobre a adequação material das definições correntes do conceito de consequência, em ‘Some Observations on the Concepts of ω -Completeness and ω -Consistency’, p. 294. Em CLC, ainda, ele põe igualmente em dúvida a adequação material das definições que Carnap elabora em *Logische Syntax der Sprache* para a noção de consequência lógica (p. 416, nota 1).

influenciado por Tarski), e o segundo é uma espécie de “versão didática” de CTFL para filósofos. Ao longo desta dissertação, faremos diversas remissões a esses dois trabalhos.

(ii) Formalização definicional *versus* formalização axiomática

Há alguns pontos nos excertos mencionados no tópico anterior que pretendemos examinar separadamente no restante deste capítulo. Trata-se das condições de correção formal e adequação material de uma definição. Antes, contudo, parece-nos importante partir de uma perspectiva mais geral, no que concerne à formalização lógico-matemática de noções intuitivas. As diversas passagens que transcrevemos acima sugerem uma única via de formalização, a via definicional. Mas a formalização não é um processo que se restringe ao estabelecimento de definições formais de noções que se encontram em nível intuitivo. A formalização de noções intuitivas também pode ser feita por uma outra via, além da definicional, a saber, pela via axiomática. Já em CTFL, Tarski assinala esses dois enfoques possíveis de formalização do conceito de verdade: num caso, a formalização é efetuada por meio de uma definição do conceito de verdade no interior de uma dada teoria; no outro caso, a formalização é efetuada por meio do estabelecimento de uma teoria axiomática do conceito de verdade, em que este último é assumido como primitivo semântico⁴⁶. Em ESS, Tarski considera os dois procedimentos gerais de formalização da semântica:

At this point two procedures come into consideration. In the first the semantical concepts (or at least some of them) are introduced into the metalanguage as new primitive concepts and their basic properties are established by means of axioms. Among these axioms are included all statements which secure the materially adequate use of the concepts in question.[...] ⁴⁷

[...] In the second procedure, [...] the semantical concepts are defined in terms of the usual concepts of the metalanguage and are thus reduced to purely logical concepts, the concepts of the language being

⁴⁶ Num sentido mais amplo da palavra ‘definição’, o enfoque axiomático também poderia ser chamado de ‘definicional’, se tivermos em mente que é costume, desde Gergonne, considerar um sistema de axiomas como uma *definição implícita* dos termos primitivos que nele ocorrem. No entanto, marcamos a distinção entre os dois enfoques, porque uma definição *sensu stricto* é, ao contrário de uma definição implícita, uma *regra específica de eliminação de termos*.

⁴⁷ ESS, p. 405.

*investigated and the specific concepts of the morphology of the language.*⁴⁸

Podemos generalizar as considerações acima: a formalização de uma noção intuitiva c se dá ou por meio de uma definição no interior de alguma teoria, ou por meio de uma axiomatização, em que c é assumida como constante primitiva, e os axiomas fixam suas propriedades fundamentais⁴⁹. Esses dois enfoques, o definicional e o axiomático, parecem esgotar as possibilidades de formalização (pelo menos para Tarski). O conceito de função algorítmica é um exemplo famoso de questões de formalização. Há um episódio envolvendo Church e Gödel, digno de nota. Antes de enunciar sua famosa tese, Church sugeriu a Gödel, em 1934, a identificação dos conceitos de função efetivamente calculável e de função recursiva geral (ou, então, λ -definível). Gödel julgou insatisfatória a sugestão. Não havia razões cogentes, em sua opinião, para identificar os dois conceitos por meio de uma definição – i.e., não havia razões cogentes para que uma tal definição fosse considerada materialmente adequada. A idéia que Gödel julgava mais satisfatória era estabelecer axiomáticamente uma teoria das funções efetivamente calculáveis. Alguma suspeita o demovia da empresa de tentar definir formalmente a noção de função efetivamente calculável. Essa desconfiança só foi desfeita quando Gödel leu o artigo em que Turing apresentou sua versão, após analisar minuciosamente a noção intuitiva de calculabilidade, em 1936⁵⁰.

Se estivermos seguindo o outro enfoque possível de caracterização formal de uma dada noção intuitiva (ou de um determinado conjunto de noções intuitivas), isto é, a via axiomática, devemos estar atentos a uma exigência natural mínima: que o sistema de axiomas seja categórico. A existência de modelos não-isomórficos seria uma prova de que a noção intuitiva não foi exatamente capturada, ou pelo menos de que não conseguimos fixar de maneira unívoca o seu significado. Tarski estava perfeitamente consciente dessa exigência. Em ‘Some Methodological Investigations on the Definability of Concepts’ (SMIDC, 1934), ele faz considerações explícitas a esse respeito (ao discutir o problema da completude de um sistema de termos primitivos):

A non-categorical set of sentences (especially if it is used as an axiom system of a deductive theory) does not give the impression of a closed and organic unity and

⁴⁸ *Idem*, p. 406.

⁴⁹ No caso de formalizações axiomáticas, o quadro pode ser um pouco mais complexo: em vez de procurar estabelecer axiomáticamente as propriedades de uma única noção intuitiva, podemos estar ocupados em estabelecer axiomáticamente as propriedades de – e as relações mútuas entre – várias noções intuitivas mutuamente independentes.

⁵⁰ A propósito, cf. o estudo bastante informativo de Biraben, R.E. *Tese de Church: algumas questões histórico-conceituais*, pp. 31-32 e 46-51.

*does not seem to determine precisely the meaning of the concepts contained in it.*⁵¹

E, como não poderia ser diferente, em CTFL, essa é a exigência principal a ser observada por qualquer teoria axiomática da verdade:

*[...] It seems natural to require that the axioms of the theory of truth, together with the original axioms of the metatheory, should constitute a categorical system. It can be shown that this postulate coincides in the present case with another postulate, according to which the axiom system of the theory of truth should unambiguously determine the extension of the symbol 'Tr' which occurs in it [...].*⁵²

Podemos dizer que a exigência de categoricidade cumpre, no âmbito do enfoque axiomático, a mesma função que a condição de adequação material cumpre no âmbito do enfoque definicional de formalização.

(iii) Duas digressões

Uma pequena digressão. A escolha da via de formalização de uma noção intuitiva levanta alguns problemas interessantes. O que nos leva, às vezes, a escolher o enfoque definicional e não o enfoque axiomático? E o que nos leva, às vezes, a proceder da maneira inversa, isto é, a escolher o enfoque axiomático e não o definicional? Haverá casos em que apenas o enfoque axiomático é possível, ou, em outros termos, haverá conceitos absolutamente indefiníveis? Ou haverá apenas casos de indefinibilidade relativa? Certamente, questões como essas estavam envolvidas quando Gödel rejeitou a sugestão de Church de identificar, definicionalmente, a noção intuitiva de calculabilidade efetiva com os conceitos formais de recursividade ou λ -definibilidade. No Capítulo 4, veremos detalhadamente por que razão Tarski propõe a via axiomática para formalizar o conceito de verdade em certos casos. Podemos, contudo, antecipar o cerne da resposta. Em certos tipos de linguagens formais, o conceito de verdade é *absolutamente indefinível*. De uma maneira geral, isso parece sugerir que, sempre que um teorema de indefinibilidade absoluta de uma determinada noção é demonstrado, temos que abandonar

⁵¹ SMIDC, p. 311. O negrito é nosso. Entrementes, Tarski parece ter feito algumas reservas a essas considerações, como se depreende da última nota a SMIDC, acrescida em 1956, p. 319.

⁵² CTFL, p. 258.

sem esperanças o enfoque definicional. Nesse caso, não resta outra alternativa, senão caracterizar formalmente essa noção por meio de um conjunto de axiomas que fixem suas propriedades fundamentais.

Mais digressão. Considere, por exemplo, a ‘estrutura intuitiva’ dos números naturais (com as operações de adição e multiplicação, e respectivos elementos neutros). Formalmente, podemos caracterizá-la por meio de um conjunto determinado de axiomas – os axiomas de Peano em primeira ordem, digamos. Ou podemos caracterizá-la formalmente seguindo o enfoque definicional, construindo-a no interior de uma teoria suficientemente abrangente – ZF ou BGN, por exemplo. Agora, onde esse processo de formalização vai parar? Parece que a via definicional não pode ser a última instância desse caminho, porque uma definição é sempre assentada numa teoria subjacente. No final, parece que teríamos de ter uma teoria fundacional última, cujos conceitos primitivos seriam absolutamente indefiníveis. O símbolo ‘ $\in$ ’ da teoria dos conjuntos parece ter um significado intuitivo mais ou menos claro (embora, evidentemente, impreciso em muitos aspectos), mas até agora o único enfoque utilizado para formalizá-lo foi o axiomático. Seria o conceito de pertinência absolutamente indefinível, e, por conseguinte, um genuíno conceito matemático primitivo? Seria uma teoria axiomática de conjuntos um sistema fundacional último (ao lado, é claro, de outras teorias concorrentes)? Trata-se apenas de especulações⁵³. É melhor voltar à apresentação do programa de Tarski. Daqui por diante, limitar-nos-emos a considerar o enfoque definicional.

Correção formal de uma definição qualquer

Quando falamos em definições formais, sempre estamos pressupondo uma teoria subjacente na qual tais definições são estabelecidas. Para que a definição de uma constante c (que pode ser uma constante individual, relacional ou funcional, ou mesmo um conectivo proposicional ou um quantificador) seja considerada *formalmente correta*, é preciso que (i) seja especificado o conjunto C de constantes em termos das quais c será definida, e (ii) sejam especificadas as regras formais de definição. Os requisitos (i) e (ii) são satisfeitos a partir do momento em que especificamos a estrutura formal da teoria subjacente e correspondente linguagem em que a definição será construída. O conjunto C de constantes contém as constantes primitivas da teoria, mas também pode conter

⁵³ Essas suaves digressões filosóficas estão relacionadas à teoria da definibilidade sintática.

constantes previamente introduzidas por meio de definições⁵⁴. Para ilustrar o item (ii), vejamos um exemplo de uma regra de definição:

Definição de constantes predicativas monádicas. Seja T uma determinada teoria. Uma equivalência Df introduzindo uma nova constante predicativa monádica P é uma definição em T , se, e somente se, Df é da forma $P(x) \leftrightarrow \varphi$, e as seguintes condições são satisfeitas: (i) x é a única variável livre que ocorre em φ , e (ii) as únicas constantes não-lógicas que ocorrem em φ são as constantes primitivas da linguagem de T ou outras constantes previamente definidas em T .

Lesniewski foi o primeiro a apontar explicitamente os critérios básicos para que as regras de definição de uma dada teoria sejam consideradas satisfatórias⁵⁵. São dois: o critério de *eliminabilidade* e o critério de *não-criatividade*. Pelo primeiro, as regras devem ser tais que, dada uma constante definida c , elas possibilitem a eliminação de c em qualquer contexto em que ocorra. Pelo segundo, as definições introduzidas devem ser não-criativas, no sentido de não incrementar o poder dedutivo da teoria. Em outras palavras, definições não devem expandir o conjunto original de teoremas.

Algumas definições violam abertamente o critério de eliminabilidade. É o caso das definições condicionais e das definições recursivas. Mas há técnicas conhecidas que contornam o problema. Não vamos entrar em detalhes sobre tais técnicas. Limitemo-nos a uma explicação superficial de uma técnica conhecida, pela qual podemos transformar uma definição recursiva numa definição explícita. Trata-se do método de Dedekind-Frege. Esse método é citado e utilizado várias vezes por Tarski em CTFL, sendo de importância para o seu programa de formalização da semântica, como perceberemos mais adiante. Bem, vejamos mais de perto o problema metodológico das definições recursivas. Considere, por exemplo, uma determinada linguagem para o cálculo proposicional clássico, cujo alfabeto é o conjunto $\{p_i\}_{i \in \mathbb{N}}$ de parâmetros proposicionais, além de um conjunto adequado de conectivos, digamos $\{\neg, \vee\}$. Considere agora a seguinte definição recursiva do conceito de fórmula bem formada nessa linguagem:

- (P1) $Fbf(x) \Leftrightarrow_{df}$ (1) $x \in \{p_i\}_{i \in \mathbb{N}}$, ou
 (2) $\exists y(Fbf(y) \wedge x = \neg y)$, ou
 (3) $\exists y \exists z(Fbf(y) \wedge Fbf(z) \wedge x = y \vee z)$, e
 (4) *Fecho*.

⁵⁴ Esses requisitos são uma generalização do que Tarski diz em SCTFS, pp. 48-49.

⁵⁵ Para uma apresentação bastante didática das regras de definição e dos critérios de Lesniewski, cf. Suppes, P. *Introduction to logic*, cap. 8, e tb. *Axiomatic set theory*, cap. 2.

A definição acima parece envolver algum tipo de circularidade, na medida em que o *definiendum*, a expressão definida (à esquerda do bicondicional), ocorre também no *definiens*, a expressão definidora (à direita do bicondicional). Por conta dessa circularidade (que, de fato, não é viciosa), não podemos eliminar o predicado ‘Fbf’ de todos os contextos em que ele ocorre⁵⁶. É por infringir o critério de eliminabilidade de Lesniewski, que uma definição recursiva como (P1) é metodologicamente problemática. Mas, feita uma certa assunção de natureza conjuntista, podemos aplicar o método de Dedekind-Frege a uma definição recursiva, e transformá-la numa definição explícita. A aplicação desse método a (P1) consiste simplesmente em substituir o *definiens* original por uma fórmula que, intuitivamente, diz ‘x pertence à menor classe de expressões que contém as letras proposicionais, e é fechada sob as operações de negação e disjunção’. Formalmente:

$$(P2) \text{ Fbf}(x) \Leftrightarrow_{df.} \forall X((\forall y(y \in \{p_i\}_{i \in \mathbb{N}} \rightarrow y \in X) \vee \forall y(y \in X \rightarrow \neg y \in X) \vee \forall y \forall z(y, z \in X \rightarrow y \vee z \in X)) \rightarrow x \in X)$$

A teoria em que (P2) é estabelecida, deve conter axiomas que garantam a existência de um conjunto não-vazio que satisfaça o seu *definiens*. Essa é a assunção de natureza conjuntista a que nos referimos há pouco.

Bem, então, de um modo geral, o esquema da definição explícita que resulta, pelo método de Dedekind-Frege, de uma definição recursiva de um predicado ‘P’, é o seguinte:

$$(1) P(x) \Leftrightarrow_{df.} \forall X(\varphi(X) \rightarrow x \in X),$$

sendo ‘ $\varphi(X)$ ’ uma transcrição das cláusulas usuais de fechamento: ‘a classe X contém os itens básicos tais e tais, e é fechada sob as operações tais e tais’. Uma vez que temos a seguinte equivalência:

$$\forall X(\varphi(X) \rightarrow x \in X) \Leftrightarrow x \in \cap \{X: \varphi(X)\},$$

podemos substituir a fórmula (1) pela seguinte:

$$(2) P(x) \Leftrightarrow_{df.} x \in \cap \{X: \varphi(X)\}$$

⁵⁶ O problema é mais facilmente constatável, por exemplo, na definição recursiva da operação de adição no conjunto dos números naturais: o operador ‘+’ é eliminável no contexto ‘2 + 3’, mas não o é em ‘x + x’.

Então, ‘ $P(x)$ ’ significa, intuitivamente falando, ‘ x pertence à menor classe que contém os itens básicos tais e tais, e que é fechada sob as operações tais e tais’.

Em CTFL, o esquema de definição explícita (‘normal’, na terminologia de Tarski) que substitui uma definição recursiva, é:

$P(x)$ se, e somente se, a fórmula $x \in X$ vale para toda classe X que satisfaz a condição $\varphi(X)$,

que, na verdade, é apenas uma versão coloquial de (1).

Correção formal de uma definição do conceito de verdade (1)

Considerando o que foi dito no início do tópico anterior, para que uma definição do conceito de verdade satisfaça a condição de correção formal, temos que (i) especificar o conjunto de constantes por meio das quais o conceito de verdade será definido, e (ii) especificar as regras de definição de acordo com as quais o conceito de verdade será definido. Os itens (i) e (ii) correspondem, de fato, a uma descrição da estrutura formal da linguagem e da teoria em que o conceito de verdade será definido.

No que diz respeito ao item (i), Tarski assume, em CTFL, a seguinte “máxima reducionista”:

*What terms are we to use in constructing the definition of truth? [...] In this construction **I shall not make use of any semantical concept if I am not able previously to reduce it to others concepts.***⁵⁷

Para compreendermos as razões da máxima reducionista de Tarski, é interessante considerar a seguinte questão: *Qual a finalidade básica de uma definição?* Há pelo menos duas respostas a essa pergunta, não necessariamente excludentes entre si. A primeira é que uma definição serve a fins pragmáticos de *economia* simbólica. À medida que vamos usando uma teoria (o que significa dizer, demonstrando teoremas em seu interior), é útil condensar expressões extensas em expressões mais compactas, a fim de facilitar o manuseio formal dos símbolos. A segunda resposta é que uma definição serve a fins semânticos de *explicação* do significado de um termo – explicação que se dá por meio de termos cujo significado já foi estabelecido anteriormente (e que, por conseguinte, já conhecemos). Nesse caso, uma definição tem a importante finalidade de esclarecimento conceitual – finalidade que às vezes não é atingida com sucesso. Quando uma definição é

⁵⁷ CTFL, p. 153. Os negritos são nossos. Onde se lê ‘outros conceitos’, subentenda-se ‘conceitos não-semânticos’.

feita com um fim explicativo, mas contém em seu *definiens* (i.e., a expressão definidora) termos cujo significado ainda não foi esclarecido, ou cujo significado é menos claro do que o significado do *definiendum* (i.e., a expressão definida), dizemos que a definição incorre no vício de definir *obscurum per obscuriora*.

Ao longo da história da teoria das definições, encontramos esses dois pontos de vista. Por exemplo, Blaise Pascal (que muito influenciou os lógicos de Port-Royal, Arnauld e Nicole) oferece simultaneamente ambas as respostas:

*Leur [des definitions] utilité et leur [idem] usage est d'éclaircir et d'abrégé le discours, en exprimant, par le seul nom qu'on impose, ce qui ne pourrait se dire qu'en plusieurs termes [...].*⁵⁸

Por outro lado, se consultarmos o que Padoa diz a respeito do assunto, veremos que ele parece inclinar-se mais à função pragmática das definições:

*In any deductive theory a 'symbolic definition' is nothing but the convention of replacing a sequence **a** of already considered symbols [...] by a new symbol **x** [...], namely, the convention of taking **x** = **a** from now on. Therefore, although symbolic definitions are very useful for the conciseness of the propositions that follow them, they are not necessary. In fact, by a simple 'symbolic translation' (that is, replacing each defined symbol by the sequence of symbols that defines it) we can reduce the system of particular symbols of the theory to the system of undefined symbols.*⁵⁹

Em CTFL e SCTFS, a posição de Tarski é inequívoca: a função precípua de uma definição é a função explicativa. De fato, em SCTFS, ao postular que a teoria em que pretendemos definir o conceito de verdade não deve conter primitivos semânticos, ele escreve:

*[I]f this postulate is satisfied, the definition of truth, or of any other semantic concept, will fulfill what we intuitively expect from every definition; that is, **it will explain the meaning of the term being defined in terms***

⁵⁸ PASCAL, B. *De l'esprit géométrique*, p. 69. Os negritos são nossos.

⁵⁹ PADOA, A. *op.cit.*, p.120.

*whose meaning appears to be completely clear and unequivocal.*⁶⁰

Parafraseando o excerto acima: o que ‘intuitivamente esperamos’ de uma definição é que ela seja explicativa. Complementa a passagem que citamos, a seguinte observação em CTFL:

*The question how a certain concept is to be defined is correctly formulated only if a list is given of the terms by means of which the required definition is to be constructed. If the definition is to fulfill its proper task, the sense of the terms in this list must admit of no doubt.*⁶¹

Subentende-se, então, que a ‘tarefa própria’ de uma definição é fornecer uma explicação do significado do *definiendum*.

Uma passagem ilustrativa de seu famoso livro *Introduction to Logic and to the Methodology of Deductive Sciences* (1941), reforça a concepção explicativa de Tarski a respeito das definições:

*Imagine, for instance, that in arithmetic the symbol “ $\leq$ ” has not as yet been employed but that one wants to introduce it now into the considerations. For this purpose it is necessary to define this symbol first, that is, to explain exactly its meaning in terms of expressions which are already known and whose meaning are beyond doubt*⁶².

Tarski mantém essa posição mesmo muitos anos depois de seu trabalho original. Em ‘Truth and Proof’ (1969), há passagens em que Tarski diz explicitamente que o propósito de uma definição do conceito de verdade é ‘explicar o seu significado corrente’. Apesar disso, ele adota uma posição mais flexível do que a que defende em CTFL e SCTFS, e chama a atenção para outros fins eventuais de uma definição⁶³.

Mas convém, a esta altura, perguntar: que termos, então, serão utilizados na definição do conceito de verdade? Que termos serão utilizados para explicar o conceito de

⁶⁰ SCTFS, pp. 54-5. Os negritos são nossos.

⁶¹ CTFL, pp. 152-3. Os negritos são nossos.

⁶² *Introduction to Logic and to the Methodology of Deductive Sciences*. Dover, 1995, pp. 33-4. O negrito é nosso.

⁶³ Cf. TP, p. 63.

verdade? Basicamente, termos de natureza sintática e termos pertencentes a um sistema fundacional matemático – p.e., termos pertencentes ao vocabulário de alguma teoria axiomática de conjuntos. E por quê tais termos? Simplesmente porque, para Tarski, o significado de termos sintático-conjuntistas é mais claro (e, talvez, menos perigoso) do que o significado de termos semânticos. Assim, a máxima reducionista de Tarski tem o seguinte objetivo: *mostrar que termos semânticos, epistemicamente obscuros e logicamente envolvidos com antinomias, são redutíveis a termos não-semânticos, de natureza sintática e conjuntista, epistemicamente mais claros e logicamente menos suspeitos*. Ao mesmo tempo que explicamos em termos claros e precisos as noções semânticas, removemo-lhes a pecha de intrinsecamente inconsistentes. Em última instância, então, é isso o que se ganha com a formalização do conceito de verdade, pela via definicional.

Correção formal de uma definição do conceito de verdade (2)

Ao encerrarmos a penúltima seção do Capítulo 1 com o diagnóstico de Tarski das antinomias semânticas, vimos que a precaução básica a ser tomada é a estratificação da linguagem, ou, mais precisamente, o estabelecimento de uma hierarquia de níveis lingüísticos: pelo menos dois níveis devem ser distinguidos, a saber, o nível da linguagem-objeto, que é a linguagem cuja sintaxe e semântica são os objetos de investigação, e o nível da metalinguagem, que é a linguagem da teoria em cujo âmbito se desenvolve a investigação da sintaxe e da semântica da linguagem-objeto. Ao lado dessa hierarquia de linguagens, temos que levar igualmente em consideração uma correspondente hierarquia de teorias. De fato, normalmente, não estamos interessados apenas em investigar uma linguagem *per se*, mas em investigar uma linguagem enquanto suporte lingüístico de alguma teoria (seja essa teoria axiomática ou não). E é assim que Tarski procede em CTFL. É sempre a linguagem de alguma teoria determinada que é objeto de investigação. Por outro lado, não há como se investigar a sintaxe ou a semântica de uma linguagem (e de uma teoria) unicamente por meio de uma metalinguagem. Rigorosamente falando, isso nem faz sentido. A investigação só pode ser desenvolvida no âmbito de uma *teoria* sobre essa linguagem (e sua correspondente teoria), isto é, no âmbito de uma meta-teoria. Em suma, ao lado da distinção entre linguagem-objeto e metalinguagem, temos que levar em consideração uma correspondente distinção entre ‘teoria-objeto’ (embora Tarski não use esse termo) e meta-teoria.

Usualmente, escolhe-se como meta-teoria algum sistema fundacional matemático, como, por exemplo, alguma versão da teoria dos tipos, ou alguma teoria axiomática dos conjuntos, digamos, ZF ou BGN em primeira ordem. Em DSRN, por exemplo, a meta-

teoria escolhida é uma versão da teoria simples dos tipos. De qualquer modo, a meta-teoria deve satisfazer uma condição que Tarski chama de *riqueza essencial* em relação à linguagem-objeto e sua correspondente teoria. No caso de a linguagem-objeto possuir uma gramática baseada na teoria dos tipos, a condição de riqueza essencial pode ser expressa da seguinte forma:

(RE1) A linguagem da meta-teoria (i.e., a metalinguagem) deve conter variáveis de ordem superior à maior ordem das variáveis da linguagem-objeto.

No caso de a gramática da linguagem-objeto não ser baseada na teoria dos tipos, a condição de riqueza essencial pode ser expressa assim:

(RE2) A meta-teoria deve estar munida de axiomas que garantam a existência de pelo menos um modelo para a teoria-objeto⁶⁴.

A condição de riqueza essencial é necessária e suficiente para uma definição satisfatória do conceito de verdade. De fato, é um requisito tão importante que, se não for satisfeito, podemos, em certo sentido, colapsar a distinção entre metalinguagem e linguagem-objeto (e meta-teoria e teoria-objeto), e transformar esta última numa linguagem semanticamente fechada. Se a condição de riqueza essencial não for satisfeita, podemos traduzir a meta-teoria e a metalinguagem na teoria-objeto e na linguagem-objeto, e reconstruir no âmbito destas últimas o paradoxo de Eubúlides-Epimênides.

Em CTFL, Tarski faz questão de descrever a estrutura formal da linguagem e da teoria em que o conceito de verdade é definido, como exige a condição de correção formal⁶⁵. O quadro abaixo resume essa estrutura, pressupondo a satisfação da condição de riqueza essencial:

	Metalinguagem	Meta-teoria
1	Expressões de carácter lógico e matemático em geral	Axiomas de carácter lógico e matemático em geral
2	Expressões com o mesmo significado das expressões da linguagem-objeto	Axiomas com o mesmo significado dos axiomas da teoria-objeto ou suficientes para a dedução de seus teoremas
3	Expressões de carácter descritivo-estrutural	Axiomas de carácter descritivo-estrutural

⁶⁴ A condição (RE2), que não ocorre em CTFL, é enunciada pela primeira vez em ‘On undecidable statements in enlarged systems of logic and the concept of truth’ (USES, 1939), p. 110. Veja-se também a nota 16 de SCTFS.

⁶⁵ CTFL, pp. 169-173 e 210-212.

Os dois primeiros grupos de expressões e axiomas dispensam comentários. O terceiro grupo de expressões contém aquelas que são usadas para denotar sinais isolados e expressões da linguagem-objeto (p.e., os símbolos do alfabeto), classes de tais expressões (p.e., o símbolo da classe das fórmulas bem formadas), seqüências de tais expressões, relações existentes entre expressões e entre classes de expressões etc. O terceiro grupo de axiomas compreende o que hoje chamamos de teoria da concatenação, cujo primitivo é o operador ‘^’ de concatenação de expressões. Tarski, aliás, foi o primeiro a axiomatizar essa teoria, em CTFL⁶⁶.

No que diz respeito às regras metalingüísticas de definição, não há muito a acrescentar ao que já dissemos na seção anterior: são adotadas as usuais, desde que satisfaçam os critérios de Lesniewski de eliminabilidade e não-criatividade. É importante que tais critérios sejam satisfeitos, especialmente o primeiro, porque, se não pudéssemos eliminar de qualquer contexto os termos semânticos definidos, ficaríamos com um problema sério: as intenções reducionistas de Tarski não seriam completamente realizadas e, por conseguinte, os objetivos explicativos do enfoque definicional não seriam satisfatoriamente alcançados.

Adequação material de uma definição qualquer

Um problema que sói acontecer em matemática é o seguinte. Percebemos que as demonstrações de um dado setor de investigação usam uma certa noção intuitiva. Muitas vezes, é o caráter evidente do conceito e de suas propriedades básicas que relaxa a necessidade de formalizá-lo. Mas, por razões de rigor formal, resolvemos formalizar a noção. Amiúde, a formalização se dá por meio de uma definição no interior de uma teoria que serve de base às nossas investigações. Então, além de respeitar as regras formais de definição, temos que cuidar de formalizar adequadamente a noção, e isto significa que, uma vez formalizado, o conceito deve preservar seu significado intuitivo original e suas propriedades básicas conhecidas. A definição serve precisamente ao propósito de formalizar o conceito intuitivamente usado, e, portanto, não faz sentido modificar-lhe o conteúdo. A exigência de conservação do significado e das propriedades informais é o que se pode chamar de condição de adequação material da definição.

Em resumo: temos um conceito *c* que satisfaz um dado conjunto de propriedades intuitivas conhecidas. Uma formalização do conceito *c* deve ser tal que todas essas propriedades intuitivas sejam preservadas. Se, de alguma maneira, é possível selecionar uma propriedade intuitiva básica *P* de *c*, isto é, uma propriedade de *c* de que todas as outras sejam conseqüências, então podemos utilizá-la na formulação exata de um *critério*

⁶⁶ CTFL, pp. 173-174. Veja também SOCωCC, p. 282.

de adequação material da definição de c : uma formalização de c deve ser tal que a propriedade P seja preservada.

Há dois exemplos bastante conhecidos de formalização definicional e correspondente problema de adequação material – o primeiro extraído da análise, e o segundo extraído da teoria da recursão –, que podemos trazer à baila para ilustrar o que foi dito acima. De Newton e Leibniz até a primeira metade do século XIX, o Cálculo baseou-se na noção intuitiva de continuidade de uma função real. Boa parte dos teoremas sobre o contínuo dos números reais baseava-se na intuição espacial – por exemplo, na intuição de uma curva contínua, na intuição de uma reta tangente a um ponto de uma curva etc. Mas quando, no século XIX, Bolzano e Weierstrass mostraram a existência de funções contínuas não diferenciáveis (em linguagem geométrica, curvas contínuas sem reta tangente a qualquer um de seus pontos), um resultado que certamente escapa à nossa intuição espacial, sentiu-se que nem sempre esta última é confiável – pois, se fôssemos basear-nos nela, provavelmente diríamos que o teorema de Bolzano e Weierstrass é falso⁶⁷. Então, se boa parte dos teoremas da fase informal do cálculo baseava-se na intuição espacial, e esta agora se mostrava pouco segura, pareceu razoável reelaborar todos os resultados do cálculo obtidos até então, sobre bases mais seguras. É nesse contexto que surgem os trabalhos de Cauchy e Weierstrass de formalização rigorosa da análise⁶⁸. A noção intuitiva de continuidade de uma função é então explicada em termos formalmente precisos pela conhecida definição por meio das tolerâncias ε e δ . Entretanto, embora essa definição tenha sido introduzida para remover a pouca segurança do uso intuitivo da noção de continuidade, seu valor está certamente adstrito à satisfação do requisito de adequação material: ela seria de pouco valor se não recuperasse, desta vez de maneira formal e precisa, as principais propriedades intuitivamente atribuídas ao conceito de função contínua. E, de fato, nesse sentido, ela pode ser dita materialmente adequada – embora não tenha sido isolada uma propriedade intuitiva básica para servir de critério de adequação material para ela.

Um segundo exemplo é o do conceito de algoritmo ou calculabilidade efetiva⁶⁹. Intuitivamente, um algoritmo é um procedimento mecânico (descrito por um conjunto finito bem definido de instruções elementares bem especificadas) por meio do qual podemos resolver um dado problema (normalmente numérico) num número finito de passos. A partir do século XX, especialmente entre os lógicos, o conceito de algoritmo foi

⁶⁷ A esse respeito, vale a pena conferir uma passagem de Poincaré em *O Valor da Ciência*, Cap. 1, p. 16.

⁶⁸ Sobre esse importante episódio da história da análise, cf. Kneale, W. & Kneale, M. *op.cit.*, pp. 405-406. Nesse mesmo ambiente de fundamentação da análise, é interessante conferir a eloquente introdução de Dedekind ao seu *Continuity and Irrational Numbers*, em que ele discorre sobre a necessidade de uma definição formal do conceito de continuidade do conjunto dos números reais (pp. 1-2).

⁶⁹ Ou, para sermos mais rigorosos, o conceito de função numérica algorítmica.

ganhando cada vez mais importância, e essa importância está particularmente vinculada ao que se chama de *Entscheidungsproblem*, ou problema da decisão, de uma teoria formal. O problema da decisão para um dado sistema formal, é o problema da existência de um procedimento algorítmico para determinar se uma dada sentença (da linguagem desse sistema) é ou não um teorema de tal sistema. Nos primeiros casos solucionados pelos lógicos (p.e., cálculo proposicional clássico, cálculo de predicados monádicos de primeira ordem etc.), não foi necessário definir formalmente o conceito de algoritmo, porque todos eles mostraram-se decidíveis: exibia-se um procedimento de decisão para uma certa teoria, e intuitivamente ficava claro que tal procedimento era efetivo. Por outro lado, até o início da década de 1930, não se tinha ainda uma resposta ao problema da decisão para o cálculo de predicados de primeira ordem e para a aritmética elementar dos números naturais – exceto para alguns de seus fragmentos (já nos referimos, por exemplo, à aritmética aditiva de Presburger). Desconfiava-se que a solução geral era negativa, mas para mostrar isso, era necessário, de antemão, definir formalmente o conceito de algoritmo. Em 1936, Church mostrou que a resposta é negativa, mas precisou antes formalizar definicionalmente o conceito de procedimento efetivo – primeiro, por meio do conceito de λ -definibilidade, depois, por meio de outras noções equivalentes. A Tese de Church pode ser vista como um enunciado que afirma a adequação material da definição formal de Church do conceito de algoritmo (a rigor, de funções numéricas algorítmicas): o conceito de λ -definibilidade (ou outro equivalente) captura exatamente a noção intuitiva de calculabilidade efetiva. Houve quem contestasse a adequação material da definição formal de Church (e outras equivalentes), como Rosza Peter e Laszlo Kálmár, mas a maioria dos lógicos e matemáticos acredita que a Tese de Church é verdadeira. De qualquer maneira, é razoável pensar que, se dispuséssemos de um critério de adequação material, possivelmente já teríamos decidido há muito tempo se a Tese de Church é mesmo, de fato, verdadeira – e teríamos eliminado qualquer rincão de dissidência porventura existente.

Um exemplo de insucesso na adequação material de uma definição é lembrado por Feferman⁷⁰. Aliás, exemplos negativos muitas vezes são mais elucidativos. Intuitivamente, podemos descrever uma curva como a figura que resulta do deslocamento de um ponto no espaço. No século XIX, Camille Jordan propôs a definição de curva como a imagem contínua de um segmento de reta. Entretanto, alguns anos depois, Peano mostrou que é possível preencher a área de um quadrado com uma tal imagem, o que é uma prova de que a definição proposta por Jordan é materialmente inadequada. De fato, dificilmente uma figura como “■ ” seria chamada por nós de ‘uma curva’.

⁷⁰ FEFERMAN, S. ‘Tarski’s conceptual analysis of semantical notions’ (2002), p. 4.

Nos três principais artigos em que Tarski empreende a definição formal de uma noção intuitiva (DSRN, CTFL e CLC), o problema da adequação material é uma questão que está sempre presente. As passagens de DSRN citadas no início deste capítulo são explícitas quanto à necessidade de adequação material de uma definição formal em geral. E, de fato, após definir matematicamente o conceito de definibilidade de um conjunto de números reais, Tarski escreve:

*Now the question arises whether **the definitions just constructed** (the formal rigour of which raises no objection) **are also adequate materially**; in other words, **do they in fact grasp the current meaning of the notion as it is known intuitively?***⁷¹

Em CTFL, a adequação material da definição do conceito de verdade é central em toda a monografia, como veremos na próxima seção. Finalmente, em CLC, como já assinalamos, o problema da adequação material da definição de Tarski do conceito de consequência lógica, é considerado por ele mesmo como um importante problema em aberto:

*I am not at all of the opinion that in the result of the above discussion the problem of a materially adequate definition of the concept of consequence has been completely solved.*⁷²

Aliás, já dissemos, hoje nós sabemos que a definição de Tarski em CLC, de fato, não é materialmente adequada⁷³. Dos três artigos, cumpre notar que apenas o segundo, CTFL, consegue isolar uma propriedade da noção intuitiva para formular um critério de adequação material. Vejamos isso mais de perto agora.

Adequação material de uma definição do conceito de verdade

A formalização de um conceito *c* deve ser tal que as propriedades intuitivas de *c* sejam preservadas. Essa é, em linhas gerais, a condição de adequação material de uma definição formal, como vimos no tópico anterior. Agora, quais são as propriedades intuitivas do conceito de verdade? Bem, há algumas bastante conhecidas. Os princípios

⁷¹ DSRN, pp. 128-129.

⁷² CLC, p. 418. Cf. nota 6.

⁷³ *Idem*.

de não-contradição e do terceiro excluído constituem dois bons exemplos. Uma sentença não pode ser simultaneamente verdadeira e falsa. Uma sentença é ou verdadeira, ou falsa, não havendo outro valor possível. Outros exemplos prosaicos: uma sentença da forma ‘ ϕ e ψ ’ é verdadeira, quando ϕ e ψ são ambas verdadeiras; se uma sentença da forma ‘se ϕ então ψ ’ é verdadeira, e a sentença ϕ é verdadeira, então a sentença ψ também é verdadeira. Poderíamos estender *ad nauseam* a lista com um sem-número de exemplos.

O problema é que, a rigor, as propriedades do conceito de verdade são em número infinito. Nossa lista acima é acidental e rapsódica. Há propriedades que nem conhecemos. Algumas propriedades, entretanto, parecem ser centrais ao conceito de verdade. Nossa definição tem que dar conta pelo menos delas. O que é crucial é encontrar, se for possível, uma propriedade intuitiva que seja básica, no sentido de todas as outras poderem ser deduzidas com base nela. Tal propriedade dará unidade ao conjunto das propriedades intuitivas do conceito de verdade. Ao mesmo tempo, servirá de critério de adequação material. Em suma, temos que encontrar uma propriedade que faça as vezes da propriedade intuitiva ‘mais forte’ do conceito de verdade. Essa propriedade, Tarski encontra expressa no chamado ‘Esquema T’ – propriedade que, mais tarde, convencionou-se chamar de “descitatória” (*disquotational*)^{74,75}. Tarski acredita que a propriedade descitatória reúne as intuições fundamentais da concepção clássica da verdade, o correspondentismo, especialmente na versão formulada por Aristóteles em sua *Metafísica*:

*Dizer daquilo que é, que não é, e daquilo que não é, que é, é falso, e dizer daquilo que é, que é, e daquilo que não é, que não é, é verdadeiro.*⁷⁶

Tarski cita outras maneiras possíveis de se enunciar a propriedade fundamental do conceito de verdade, quando referida a sentenças, por exemplo:

(C1) Uma sentença é verdadeira, se, e somente se, corresponde aos fatos.

Ou:

⁷⁴ Uma tradução alternativa para ‘disquotational’ seria ‘descitacional’ ou mesmo ‘disquotacional’. No entanto, como não encontramos o registro de ‘citacional’ nos nossos dicionários de referência (*Houaiss*, por exemplo), preferimos usar a forma negativa, embora cacofônica, de ‘citatório’, termo já catalogado.

⁷⁵ É importante não confundir a posição de Tarski com o ponto de vista deflacionista do chamado “disquotacionalismo”. Para os disquotacionalistas, o conceito de verdade é vazio de conteúdo filosófico. Eles defendem que o predicado ‘... é verdadeiro’ é apenas um expediente lógico para: (i) descitar; (ii) expressar, por meios finitos, conjunções e disjunções infinitas; (iii) quantificar objetualmente em contextos substitucionais.

⁷⁶ Livro Γ , 7, 27.

(C2) Uma sentença é verdadeira, se, e somente se, designa um estado de coisas existente.

Contudo, Tarski afirma que tais formulações, embora intuitivamente claras, são, do ponto de vista lógico, vagas e imprecisas, de modo que é preciso apresentar uma formulação menos equívoca, que não contenha termos suscetíveis de serem diversamente interpretados, como ‘correspondência’, ‘fato’, ‘designação’, ‘estado de coisas’ etc.

Que é bastante discutível a crença de Tarski de que a propriedade descitatória captura as intuições básicas do correspondentismo, não nos interessa discutir nesta dissertação. Mas, afinal, em que consiste, do ponto de vista lógico, a propriedade descitatória? Suponhamos uma sentença como ‘a neve é branca’ (é o exemplo que ele usa, e que quase todos os lógicos, em uníssono, gostam de usar quando estão a discorrer sobre o conceito de verdade). O que significa dizer que a sentença ‘a neve é branca’ é verdadeira? Ou melhor, como definiríamos de modo inequívoco e simples a verdade da sentença ‘a neve é branca’? A resposta mais simples e imediata, segundo Tarski, é a seguinte equivalência:

(N1) A sentença ‘a neve é branca’ é verdadeira se, e somente se, a neve é branca.

Note que a expressão ‘a neve é branca’ ocorre duas vezes na equivalência acima: no lado esquerdo do bicondicional, entre aspas (i.e., mencionada, *in suppositio materialis*), e no lado direito do bicondicional, sem aspas (i.e., usada, *in suppositio formalis*). No lado esquerdo do bicondicional, temos o *nome* da sentença. No lado direito, a própria sentença. Observe que encerrar uma expressão entre aspas não é o único expediente disponível para se formar o seu nome. Pode-se recorrer ao que Tarski chama de *nomes descritivo-estruturais*. Um nome descritivo-estrutural de uma expressão é uma descrição exaustiva de todos os sinais que nela ocorrem, bem como da ordem em que estão dispostos. Um nome descritivo-estrutural da sentença *a neve é branca* pode ser a descrição definida abaixo:

A expressão formada pelos seguintes sinais, nesta ordem: letra A, vazio, letra ENE, letra E, letra VÊ, letra E, vazio, letra E (com acento agudo), vazio, letra BÊ, letra ERRE, letra A, letra ENE, letra CÊ, letra A

Dado este nome, a equivalência (N1) pode ser reescrita assim:

(N2) A expressão formada pelos seguintes sinais, nesta ordem: letra A, vazio, letra ENE, letra E, letra VÊ, letra E, vazio, letra E (com acento agudo), vazio, letra BÊ, letra ERRE, letra A, letra ENE, letra CÊ, letra A, é verdadeira, se, e somente se, a neve é branca.

Se abstrairmos das equivalências (N1) e (N2) a sentença ‘a neve é branca’, que é usada no lado direito do bicondicional, e, em seu lugar, indicarmos com a letra esquemática p uma sentença qualquer, e se, além disso, substituirmos o seu nome (seja o formado com aspas, seja o descritivo-estrutural) pela letra s , obtemos o seguinte esquema de equivalência:

(T) A sentença s é verdadeira se, e somente se, p .

Tal é o esquema a que nos referimos alguns parágrafos atrás, e que Tarski batizou de ‘Esquema T’ (‘T’ de *truth*). Tarski afirma que a propriedade expressa no Esquema T captura de modo simples e inequívoco as intenções da concepção clássica da verdade, de modo que tal esquema fornece a propriedade de que necessitamos para estabelecer um critério preciso de adequação material da definição do conceito de verdade. Formulado em termos mais ou menos precisos, o critério de adequação material (ou, como Tarski também o denomina, *Convenção T*) pode ser expresso da seguinte maneira:

(CAM) Seja D uma definição do conceito de verdade numa determinada linguagem L ; D é materialmente adequada, se, e somente se, todas as instâncias do Esquema T para as sentenças de L são dedutíveis de D na meta-teoria.

Em outras palavras, se a definição do predicado de verdade é materialmente adequada, dela se pode deduzir qualquer instância do Esquema T, isto é, equivalências como:

A sentença ‘a neve é branca’ é verdadeira se, e somente se, a neve é branca.

A sentença ‘A Terra gira em torno do Sol’ é verdadeira se, e somente se, a Terra gira em torno do Sol. etc. etc. etc.,

se estivermos, por exemplo, definindo o conceito de verdade para um fragmento do português.

Tarski sugere que cada instância do Esquema T seja vista como uma definição parcial do conceito de verdade, a saber, da verdade da sentença mencionada e usada na instância. A definição geral seria, intuitivamente, a “conjunção” de todas as definições parciais. Daí a exigência de que, da definição geral, sejam dedutíveis todas as instâncias do Esquema T, i.e., de que a propriedade descitatória se aplique a todas as sentenças da linguagem-objeto. •

Capítulo 3 – *O conceito de verdade formalizado*

SUMÁRIO. O problema de uma definição recursiva direta do conceito de verdade e o conceito de satisfação. Satisfação por uma seqüência enumerável de objetos. Tipos semânticos. A definição formal dos conceitos de satisfação e verdade.

Vejamos agora como Tarski procede à formalização do conceito de verdade. Neste capítulo, vislumbraremos apenas o enfoque definicional, que, de fato, reúne suas maiores contribuições aos fundamentos da semântica, e que é o enfoque principal assumido em CTFL. A apresentação do enfoque axiomático será adiada para o próximo capítulo, quando estivermos discutindo um resultado de CTFL intimamente relacionado à teoria axiomática da verdade – a saber, o teorema da indefinibilidade da verdade.

O problema de uma definição recursiva direta do conceito de verdade e o conceito de satisfação

É um fato bastante conhecido que a definição-chave de CTFL não é a do conceito de verdade, mas a do conceito de satisfação. Com efeito, é com base neste último, que Tarski define o conceito de verdade. Mas, afinal, por que motivo ele recorre ao conceito de satisfação para definir o conceito de verdade? Nesta seção, procuraremos responder com alguma clareza a essa pergunta.

Se a linguagem-objeto contivesse apenas um número finito de sentenças, a definição do conceito de verdade seria bastante simples. Ela assumiria a forma de uma definição “por escrutínio” (ou, como diz Tarski, *list-like*). De fato, a única coisa que teríamos a fazer seria construir a conjunção de todas as definições parciais de verdade. Mais precisamente, listaríamos todas as sentenças $s_1, \dots, s_n$ da linguagem-objeto, com seus respectivos nomes descritivo-estruturais $d_1, \dots, d_n$, e formaríamos a seguinte conjunção:

$$(V(d_1) \Leftrightarrow s_1) \wedge \dots \wedge (V(d_n) \Leftrightarrow s_n)$$

A conjunção acima é uma definição do conceito de verdade nessa linguagem. A fim de lhe dar a forma correta de uma definição, transformamo-la na seguinte fórmula equivalente:

$$V(x) \Leftrightarrow_{df.} ((x = d_1 \wedge s_1) \vee \dots \vee (x = d_n \wedge s_n))$$

É evidente que essa definição implica todas as instâncias do Esquema T para as sentenças da linguagem-objeto (por simples eliminação da conjunção). Portanto, além de formalmente correta, ela é materialmente adequada – capturando, assim, a intuição básica subjacente ao conceito clássico de verdade.

O problema é que uma linguagem com um número finito de sentenças é algo bastante exótico. Do ponto de vista prático, não tem a menor serventia. O que nos interessa são as linguagens com base nas quais construímos teorias, e tais linguagens contêm um número infinito (normalmente contável) de sentenças. Mas a definição do conceito de verdade em linguagens com um número infinito de sentenças envolve algumas dificuldades técnicas. A primeira delas é que, no âmbito de uma metalinguagem finitária (i.e., cujas expressões possuem comprimento finito), uma definição por escrutínio é inexeqüível. De fato, uma tal definição teria de assumir a forma de uma conjunção infinita de todas as definições parciais⁷⁷.

Uma saída muito natural para essa dificuldade é utilizar um procedimento recursivo, para dar conta do número infinito de sentenças da linguagem-objeto. Uma primeira idéia seria tentar construir uma definição recursiva direta do conceito de verdade. Definiríamos o conceito de verdade para o caso básico das sentenças mais simples da linguagem-objeto, e, de acordo com as operações por meio das quais construímos sentenças compostas a partir de sentenças mais simples, definiríamos o conceito de verdade de uma sentença complexa em função da verdade de suas sentenças constituintes.

Para as sentenças básicas da linguagem-objeto, considerariamos as instâncias do Esquema T, sendo ‘ d_α ’ o nome descritivo-estrutural de ‘ α ’ :

$$V(d_\alpha) \Leftrightarrow \alpha$$

Definiríamos sem dificuldade a verdade de fórmulas compostas formadas pela aplicação das operações do cálculo proposicional:

⁷⁷ À época de CTFL, a idéia de uma linguagem infinitária era algo bizarra, embora lógicos como Russell e C. I. Lewis, por exemplo, interpretassem fórmulas quantificadas existencial e universalmente como cadeias de disjunções e conjunções, respectivamente, e cujo comprimento dependeria da cardinalidade do domínio de objetos percorridos pelas variáveis; Mal’cev foi o primeiro a considerar seriamente em suas investigações linguagens infinitárias; nas décadas de 50 e 60, Tarski também desenvolveria pesquisas sobre tais linguagens. É interessante notar que Henkin apontou certas dificuldades técnicas na definição do conceito de verdade em linguagens com *strings* infinitas de quantificadores (e também com quantificadores parcialmente ordenados). Ele observou que não é possível aplicar diretamente o método desenvolvido por Tarski a tais linguagens. A constatação de Henkin inspirou o surgimento da chamada ‘game-theoretical semantics’, cujo principal expoente é Hintikka.

$$\begin{aligned}
V(\neg\alpha) &\Leftrightarrow \text{n\~ao ocorre que } V(\alpha) \\
V(\alpha\wedge\beta) &\Leftrightarrow V(\alpha) \text{ e } V(\beta) \\
V(\alpha\vee\beta) &\Leftrightarrow V(\alpha) \text{ ou } V(\beta) \\
V(\alpha\rightarrow\beta) &\Leftrightarrow \text{se } V(\alpha), \text{ ent\~ao } V(\beta) \\
V(\alpha\leftrightarrow\beta) &\Leftrightarrow V(\alpha) \text{ se, e somente se, } V(\beta)
\end{aligned}$$

Por fim, tentaríamos definir a verdade de sentenas compostas formadas por generalizao:

$$\begin{aligned}
V(\forall x\alpha) &\Leftrightarrow \text{para todo } x, V(\alpha) \\
V(\exists x\alpha) &\Leftrightarrow \text{para algum } x, V(\alpha)
\end{aligned}$$

O problema com as duas ltimas clusulas  evidente: normalmente, α no  uma sentena, mas uma frmula aberta⁷⁸; por conseguinte, no faz sentido aplicar-lhe o predicado V , pois frmulas abertas no so verdadeiras, nem falsas. O mesmo problema se apresentaria com as clusulas bsicas, referentes s frmulas atmicas, se a linguagem-objeto no contivesse constantes individuais.

Para solucionar esse problema, somos compelidos a procurar um conceito que satisfaa dois requisitos: (i) o conceito deve ser mais geral do que o conceito de verdade, de tal modo que se aplique tanto a frmulas fechadas como a frmulas abertas da linguagem-objeto; isso nos garante a possibilidade de defini-lo recursivamente, de maneira direta, de modo a dar conta do nmero infinito de frmulas (e, por conseguinte, sentenas) da linguagem-objeto; (ii) quando aplicado a sentenas, o conceito geral deve conduzir imediatamente ao conceito de verdade.

O candidato mais natural que se apresenta  o conceito de satisfao. Intuitivamente, trata-se de uma noo bastante simples e conhecida. Dada uma frmula com n variveis livres e uma n -upla ordenada de objetos, dizemos que a n -upla satisfaz a frmula, se, e somente se, vale a frmula que resulta da substituio de cada varivel livre por cada componente da n -upla. Por exemplo:

$$\begin{aligned}
\langle 2, 3, 5 \rangle &\text{ satisfaz a frmula ' } x + y = z \text{ ' } \Leftrightarrow 2 + 3 = 5 \\
\langle 1, 1 \rangle &\text{ satisfaz a frmula ' } x < y \text{ ' } \Leftrightarrow 1 < 1 \\
\langle 45 \rangle &\text{ satisfaz a frmula ' } x \div 5 = 9 \text{ ' } \Leftrightarrow 45 \div 5 = 9
\end{aligned}$$

⁷⁸ Exceto em casos artificiais irrelevantes, decorrentes da definio das frmulas bem formadas de uma linguagem com quantificadores. Por exemplo, contamos como bem formada uma frmula como ' $\forall x(1 < 2)$ '.

Do ponto de vista intuitivo, o conceito de satisfação aplica-se apenas a fórmulas abertas. Uma fórmula com n variáveis livres é satisfeita por n -uplas ordenadas de objetos: uma fórmula com três variáveis livres é satisfeita por ternos ordenados de objetos; uma fórmula com duas variáveis livres é satisfeita por pares ordenados de objetos; uma fórmula com uma variável livre é satisfeita por objetos singulares. Mas, do ponto de vista formal, podemos estender a aplicação do conceito de satisfação a fórmulas fechadas. Uma fórmula sem variáveis livres, isto é, uma sentença, é satisfeita pela “0-upla” de objetos, por assim dizer.

$\langle \rangle$ satisfaz a fórmula ‘ $7 < 8$ ’ $\Leftrightarrow 7 < 8$

É palpável que o conceito de satisfação pela “0-upla” de objetos captura a propriedade expressa no Esquema T, e, portanto, pode fazer as vezes do conceito de verdade. A partir dessas considerações intuitivas, então, fica claro que *o conceito de satisfação preenche os dois requisitos exigidos acima para contornar o problema de uma definição direta do conceito de verdade*. Aplica-se a todas as fórmulas, e, portanto, pode ser definido recursivamente de modo direto, e, quando aplicado em particular às fórmulas fechadas, conduz imediatamente ao conceito de verdade.

O uso de n -uplas ordenadas, entretanto, é tecnicamente inconveniente. Considerando a indexação das variáveis da linguagem-objeto, é mais adequado lançar mão de seqüências de objetos. Neste caso, temos duas alternativas: ou definimos o conceito de verdade por meio do conceito de satisfação por uma seqüência *finita* de objetos, ou definimo-lo por meio do conceito de satisfação por uma seqüência *infinita* (denumerável) de objetos. Em DSRN, Tarski esboça uma definição do conceito de verdade usando seqüências finitas, mas reconhece, no final do artigo, que o uso de seqüências infinitas simplifica mais a definição do conceito de satisfação (e, conseqüentemente, de verdade)⁷⁹. Em CTFL, as coisas se invertem: ele utiliza seqüências infinitas, e assinala a possibilidade de se proceder usando seqüências finitas⁸⁰. O uso de seqüências finitas assemelha-se ao uso intuitivo de n -uplas ordenadas: uma sentença é verdadeira, se, e somente se, é satisfeita pela seqüência vazia de objetos, e falsa, se nenhuma seqüência a satisfaz. Deixaremos de lado as seqüências finitas, e nos concentraremos na definição do conceito de verdade com base no conceito de satisfação por uma seqüência infinita de objetos.

⁷⁹ DSRN, pp. 116-7 e 141-2.

⁸⁰ CTFL, p. 195, nota 1.

Satisfação por uma seqüência denumerável de objetos

Intuitivamente, vejamos como a noção de satisfação por uma seqüência denumerável de objetos captura exatamente a propriedade básica do conceito de verdade. Para fins heurísticos, utilizemos a seguinte “definição” do conceito de satisfação: uma seqüência s satisfaz a fórmula $\varphi(x_{i_1}, \dots, x_{i_n})$ se, e somente se, vale a fórmula que resulta da substituição das variáveis $x_{i_1}, \dots, x_{i_n}$ pelos correspondentes termos $s_{i_1}, \dots, s_{i_n}$ da seqüência s , isto é, se, e somente se, $\varphi(s_{i_1}, \dots, s_{i_n})$. Ora, quando uma fórmula φ é fechada, não há termos da seqüência s que correspondam às variáveis livres de φ , simplesmente porque estas não existem. Assim, a fórmula que resulta da substituição das variáveis livres de φ pelos respectivos termos de s , é a própria fórmula φ . Então, uma seqüência s satisfaz a fórmula φ se, e somente se, vale a fórmula que resulta da substituição das variáveis livres de φ pelos correspondentes termos da seqüência s , isto é, se, e somente se, φ . Esquemáticamente, dada uma seqüência arbitrária s e uma fórmula fechada φ , temos:

$$s \text{ satisfaz } \varphi \Leftrightarrow \varphi$$

Tarski percebeu antecipadamente que a aplicação do conceito de satisfação a fórmulas fechadas captura exatamente a propriedade expressa no Esquema T, a propriedade descitatória. O esquema acima já sugere que o conceito de satisfação por uma seqüência s arbitrária é precisamente o conceito de verdade que estamos buscando. É intuitivamente claro (embora não formalmente imediato) que se alguma seqüência satisfaz φ , então toda seqüência satisfaz φ . A satisfatibilidade de φ não depende dos termos das seqüências, mas da própria φ – queremos dizer, depende do fato de φ valer ou não valer. Isso implica, é importante notar, que é indiferente definir o conceito de verdade em termos de satisfação por todas as seqüências, ou em termos de satisfação por alguma seqüência.

Em suma, podemos condensar o que foi dito acima, da seguinte maneira. Seja seguinte o “algoritmo”⁸¹ para construir a fórmula $\varphi[s]$:

- (i) Se φ contém variáveis livres $x_{i_1}, \dots, x_{i_n}$, substituí-las pelos termos $s_{i_1}, \dots, s_{i_n}$ da seqüência s ;
- (ii) Se φ não contém variáveis livres, deixar φ como tal.

Com base nessas instruções, definimos o conceito de satisfação assim:

⁸¹ É claro que não se trata, a rigor, de um algoritmo, pois a seqüência s que aparece na cláusula (i) não é necessariamente gerada de maneira efetiva.

$$s \text{ satisfaz } '\varphi' \Leftrightarrow \varphi[s]$$

Dado o algoritmo acima, é fácil ver que, sendo φ uma sentença (uma fórmula fechada), o esquema redonda no seguinte:

$$s \text{ satisfaz } '\varphi' \Leftrightarrow \varphi$$

para qualquer seqüência s arbitrariamente escolhida.

A satisfação de uma fórmula com n variáveis livres por uma seqüência s , depende apenas da n -upla de termos de s correspondentes a essas variáveis livres, sendo todos os outros termos irrelevantes. No caso extremo de uma fórmula sem variáveis livres, i.e., no caso de uma sentença, a satisfação não depende de termo algum da seqüência – todos os termos são irrelevantes. Pode ocorrer que uma fórmula aberta seja satisfeita por algumas seqüências, e não satisfeita por outras. Mas só existem duas possibilidades para uma sentença: ou ela é satisfeita por todas as seqüências, ou nenhuma seqüência a satisfaz. Isso sugere, então, a seguinte definição do conceito de verdade (e falsidade):

$$V(\alpha) \Leftrightarrow \text{toda (ou alguma) seqüência satisfaz } \alpha$$

$$F(\alpha) \Leftrightarrow \text{nenhuma seqüência satisfaz } \alpha$$

Tipos semânticos

Faremos um breve interlúdio aqui, para considerar um conceito que será de alguma valia neste capítulo, e, especialmente, no próximo. Trata-se do conceito de “tipo semântico” de uma fórmula, uma importante noção que Tarski utiliza ao longo de boa parte de CTFL. Com base nele, é possível entrever uma generalização do problema de uma definição recursiva direta do conceito de verdade, bem como da estratégia usada para contorná-lo. Isso nos proporcionará uma compreensão mais profunda do que está em jogo nesse problema.

Pois bem, denotemos por $v(\phi)$ o número de variáveis livres que ocorrem na fórmula ϕ . Dizemos que as fórmulas ϕ e γ possuem o mesmo tipo semântico, se, e somente se, ϕ e γ possuem o mesmo número de variáveis livres. Em símbolos:

$$\phi \equiv_t \gamma \Leftrightarrow_{df} v(\phi) = v(\gamma)$$

É fácil ver que $\equiv_t$ é uma relação de equivalência. Cada classe de equivalência determinada por uma fórmula ϕ *modulo* $\equiv_t$ diz-se o *tipo semântico* de ϕ . Em símbolos:

$$[\phi]_t =_{df.} \{\gamma \in \text{fbf} : \gamma \equiv_t \phi\}$$

Quando $v(\phi) = n$, dizemos que ϕ é de tipo semântico n . A partir das definições acima, é fácil ver que o número de tipos semânticos depende de quantas variáveis livres podem ocorrer numa fórmula. Como uma fórmula pode conter um número finito arbitrário de variáveis, segue que há um número infinito (contável) de tipos semânticos.

Observe que não há como construir uma definição recursiva de um tipo semântico n , para um n determinado (p.e., $n = 1$). De fato, como em qualquer definição recursiva, teríamos de fazer o seguinte: (i) partiríamos das fórmulas básicas do tipo semântico em questão; (ii) mostrariamos como compor fórmulas desse tipo, a partir de fórmulas mais simples desse mesmo tipo, por meio da aplicação de determinadas operações (negação, conjunção, generalização etc.). Mas é óbvio que isso não daria certo. Não daria certo, porque uma fórmula de um tipo semântico pode conter sub-fórmulas que pertencem a um tipo semântico arbitrário. Por exemplo, é fácil construir fórmulas de tipo semântico 1 que contêm uma sub-fórmula de tipo arbitrário: a fórmula $\forall x_1 \dots \forall x_n \alpha(x_1, \dots, x_n; y)$ contém a sub-fórmula $\alpha(x_1, \dots, x_n; y)$, com $n+1$ variáveis livres, sendo n arbitrário. Em suma, para construir uma definição recursiva de um dado tipo semântico, teríamos de pressupor como definidos todos os outros tipos semânticos.

Isso não significa que não podemos construir uma definição de tipo semântico de uma fórmula – de fato, começamos esta seção com uma tal definição! Mas essa definição pressupõe já a definição do conceito mais geral de fórmula, não importa de que tipo semântico. Este conceito, sim, é definível recursivamente. O conceito de tipo semântico de uma fórmula é obtido por uma especialização do conceito geral de fórmula.

O fato de não podermos definir recursivamente o conjunto das fórmulas de um determinado tipo semântico, tem uma consequência importante. Considerem-se as fórmulas de tipo semântico n , para um n qualquer. Seja P uma propriedade que, do ponto de vista intuitivo, aplica-se exclusivamente a tais fórmulas. Não há como construir uma definição recursiva direta do conjunto das fórmulas desse tipo que têm a propriedade P .

De que maneira podemos contornar essa dificuldade? Ora, lembremos que, embora não possamos definir recursivamente o tipo semântico de uma fórmula, podemos defini-lo indiretamente por meio de uma especialização do conceito geral de fórmula, que é definível recursivamente. A estratégia, então, é encontrar uma propriedade geral Q aplicável a fórmulas de qualquer tipo semântico (i.e., a todas as fórmulas), e, por

consequente, recursivamente definível, e que, quando aplicada às fórmulas do tipo semântico em questão, corresponda exatamente à propriedade P original. Em outras palavras, definimos recursivamente a propriedade Q , e obtemos a propriedade P como um caso especial, concernente ao tipo semântico que temos em vista.

Temos aqui uma explicação suficientemente geral da impossibilidade de se construir uma definição recursiva direta do conceito de verdade, bem como uma estratégia geral para contornar essa dificuldade. Já nos havíamos adiantado a uma explicação preliminar desse fato. Sentenças são fórmulas de tipo semântico 0, isto é, fórmulas sem variáveis livres. O conceito de verdade aplica-se exclusivamente a sentenças. Pelas razões expostas acima, não há como construir uma definição recursiva direta do conjunto das sentenças verdadeiras. A estratégia de Tarski é uma instância da estratégia geral delineada acima: recorremos ao conceito de satisfação, que se aplica a fórmulas de qualquer tipo, sendo, portanto, recursivamente definível, e que conduz ao conceito especial de verdade, quando aplicado a sentenças⁸².

Também temos aqui uma explicação geral da impossibilidade de se construir uma definição recursiva direta de qualquer conceito especial de satisfação⁸³, assim como um método para contornar esse problema. Dado um número natural n qualquer, o conceito de satisfação por uma n -upla ordenada de objetos aplica-se exclusivamente a fórmulas de tipo semântico n . No entanto, pelo que foi exposto acima, não há como construir uma definição recursiva direta do conceito especial de satisfação por n -uplas ordenadas. O método para contornar esse problema, mais uma vez, é uma instância da estratégia geral esboçada acima: recorremos ao conceito geral de satisfação por uma seqüência (infinita ou arbitrariamente finita), que se aplica a fórmulas de qualquer tipo, sendo, por conseguinte, recursivamente definível, e que nos leva ao conceito especial de satisfação por uma n -upla ordenada, quando aplicada a fórmulas de tipo semântico n .

A definição formal do conceito de satisfação por uma seqüência denumerável e do conceito de verdade

Findo o interlúdio acima e explicada a razão para a introdução de seqüências infinitas, vejamos como proceder formalmente a uma definição recursiva do conceito de satisfação. Para simplificar a exposição, vamos considerar a linguagem L da teoria elementar das ordens parciais, em que a única constante primitiva é o símbolo relacional ' $\leq$ '.

⁸² A propósito, convém conferir a primeira página de 'Arithmetical extensions of relational systems', em que Tarski alude brevemente à discussão deste interlúdio.

⁸³ No próximo capítulo, veremos que, em certos casos, seria interessante dispor de uma tal definição.

No caso básico em que α é uma fórmula atômica ' $x_i \leq x_j$ ', não temos dificuldades:

$$(1) s \text{ satisfaz } x_i \leq x_j \Leftrightarrow_{df.} s_i \text{ é menor ou igual a } s_j$$

No caso de fórmulas compostas construídas a partir dos conectivos proposicionais, também não há dificuldades. Sejam α e β fórmulas bem formadas de L. Então:

- (2) s satisfaz $\neg\alpha \Leftrightarrow_{df.} s$ não satisfaz α
- (3) s satisfaz $\alpha \wedge \beta \Leftrightarrow_{df.} s$ satisfaz α e s satisfaz β
- (4) s satisfaz $\alpha \vee \beta \Leftrightarrow_{df.} s$ satisfaz α ou s satisfaz β
- (5) s satisfaz $\alpha \rightarrow \beta \Leftrightarrow_{df.}$ se s satisfaz α , então s satisfaz β
- (6) s satisfaz $\alpha \leftrightarrow \beta \Leftrightarrow_{df.} s$ satisfaz α , se, e somente se, s satisfaz β

A definição de satisfação para fórmulas universalmente quantificadas pode parecer um tanto estranha à primeira vista. Assim, talvez seja interessante fazer algumas considerações heurísticas preliminares. Consideremos, à guisa de ilustração, a fórmula ' $\forall x_i (x_j \leq x_i)$ '. Sob que condições dizemos que uma seqüência s satisfaz a fórmula ' $\forall x_i (x_j \leq x_i)$ '? Vejamos, passo a passo.

- (i) s satisfaz ' $\forall x_i (x_j \leq x_i)$ ', se, e somente se, $\forall x_i (s_j \leq x_i)$.
- (ii) $\forall x_i (s_j \leq x_i)$, se, e somente se, para qualquer seqüência t , t satisfaz ' $s_j \leq x_i$ '.
- (iii) para qualquer seqüência t , t satisfaz ' $s_j \leq x_i$ ', se, e somente se, para qualquer seqüência t tal que $t_j = s_j$, t satisfaz ' $x_j \leq x_i$ '.

De (i)-(iii), segue, portanto:

- (iv) s satisfaz ' $\forall x_i (x_j \leq x_i)$ ', se, e somente se, para qualquer seqüência t tal que $t_j = s_j$, t satisfaz ' $x_j \leq x_i$ '.

Por conveniência técnica, reescrevemos (iv) da seguinte forma, sendo ' $t \cong_i s$ ' uma abreviação de ' t difere de s no máximo quanto ao i -ésimo termo':

- (v) s satisfaz ' $\forall x_i (x_j \leq x_i)$ ', se, e somente se, para qualquer seqüência $t \cong_i s$, t satisfaz $x_j \leq x_i$.

De um modo geral, portanto, temos, para qualquer seqüência s e qualquer fórmula α de L:

(7) s satisfaz $\forall x_i \alpha \Leftrightarrow_{df.}$ para qualquer seqüência $t \cong_i s$, t satisfaz α .

(8) s satisfaz $\exists x_i \alpha \Leftrightarrow_{df.}$ existe uma seqüência $t \cong_i s$, tal que t satisfaz α .

Podemos transformar essa definição recursiva numa definição explícita. Por razões de simplicidade, suponhamos que os conectivos proposicionais primitivos de L sejam ‘ $\wedge$ ’ e ‘ $\neg$ ’, e que o único símbolo quantificacional primitivo de L seja ‘ $\forall$ ’. Os outros conectivos e o quantificador existencial podem ser introduzidos por meio de tais símbolos. Considere a seguinte fórmula (aberta na variável S), em que Σ denota a classe de todas as seqüências do domínio de interpretação de L , fbf denota o conjunto das fórmulas bem formadas de L , e N o conjunto dos números naturais:

Para quaisquer $t, \gamma, \langle t, \gamma \rangle \in S$ se, e somente se, $\langle t, \gamma \rangle \in \Sigma \times fbf$, e

1. existem $i, j \in N$ tais que $\gamma = 'x_i \leq x_j'$ e t_i é menor ou igual a t_j , ou
2. existe $\alpha \in fbf$ tal que $\gamma = '\neg \alpha'$ e $\langle t, \alpha \rangle \notin S$, ou
3. existem $\alpha, \beta \in fbf$ tais que $\gamma = '\alpha \wedge \beta'$ e $\langle t, \alpha \rangle \in S$ e $\langle t, \beta \rangle \in S$, ou
4. existem i, α tais que $i \in N$ e $\alpha \in fbf$ e $\gamma = '\forall x_i \alpha'$ e, para todo $r \in \Sigma$ tal que $r \cong_i t$, $\langle r, \alpha \rangle \in S$.

Abreviemos a fórmula acima como ‘ $C(S)$ ’. Aplicando o método de Dedekind-Frege, transformamos a definição recursiva original na seguinte definição explícita:

$$s \text{ satisfaz } \varphi \Leftrightarrow_{df.} \forall S (C(S) \Rightarrow \langle s, \varphi \rangle \in S)$$

Equivalentemente:

$$s \text{ satisfaz } \varphi \Leftrightarrow_{df.} \langle s, \varphi \rangle \in \bigcap \{S : C(S)\}$$

Tal definição permite-nos eliminar o predicado ‘satisfaz’ de qualquer contexto em que ocorra. Sendo assim, o critério de eliminabilidade de Lesniewski é respeitado, e o programa reducionista de Tarski é realizado cabalmente.

Finalmente, definimos o conceito de verdade-em- L . Seja α uma sentença (i.e., uma fórmula fechada). Então:

$$V_L(\alpha) \Leftrightarrow_{df.} \text{ toda seqüência } s \in \Sigma \text{ satisfaz } \alpha$$

$$F_L(\alpha) \Leftrightarrow_{df.} \text{ nenhuma seqüência } s \in \Sigma \text{ satisfaz } \alpha$$

E isso é tudo. Está respondida a famosa pergunta que Pilatos uma vez dirigiu a Cristo: *Veritas? Quod est veritas?* Duas conseqüências da definição de Tarski devem ser mencionadas. A primeira é que ela é materialmente adequada, embora Tarski não tenha demonstrado isso em CTFL (e, aliás, em nenhum outro artigo!). Todas as instâncias do Esquema T para as sentenças da linguagem-objeto são dedutíveis na meta-teoria em que a definição é formulada. A segunda conseqüência é que, a partir da definição, várias outras propriedades que intuitivamente atribuímos ao conceito de verdade, como o princípio do terceiro excluído e a lei de não-contradição, podem igualmente ser deduzidas na meta-teoria. Isso já era de se esperar, já que, de fato, a definição é materialmente adequada.

É bom não esquecer que a definição de Tarski em CTFL diz respeito exclusivamente ao conceito de verdade-em-L, para L constante, e não variável. A tarefa de definir o conceito “translingüístico” de verdade-em-L, com L percorrendo as linguagens de uma determinada classe (portanto, L sendo variável), pertence ao que Tarski chama de ‘semântica geral’. O empreendimento da semântica geral, do ponto de vista definicional, só pode ser bem sucedido se a classe de linguagens que a variável L percorre não incluir todas as linguagens possíveis. Do contrário, a linguagem em que é formulada a semântica geral será uma linguagem semanticamente fechada. Essa é a principal razão por que Tarski não aceita a existência de uma linguagem universal. Uma outra dificuldade que Tarski associa à semântica geral concerne à necessidade de se definir o próprio conceito de linguagem. No entanto, em CTFL, como o próprio Tarski assinala, não há uma resposta formal precisa à pergunta: *O que é uma linguagem?* O leitor atento perceberá que Tarski, em CTFL, sequer tem uma resposta formal precisa à pergunta: *O que é uma linguagem formalizada?* Essas noções atravessam toda a monografia em nível puramente intuitivo. Explorar essas questões, no entanto, por mais interessantes que sejam, foge ao escopo geral desta dissertação. •

Capítulo 4 – *Para ler* The Concept of Truth

SUMÁRIO. Preliminares: fundamentos da metamatemática. A teoria das categorias semânticas de Husserl-Lesniewski. Tipologia Tarskiana das linguagens formalizadas. A ambigüidade semântica do conceito de satisfação. Linguagens de ordem infinita e a versão absoluta do teorema da indefinibilidade da verdade. Teoria axiomática da verdade. O pós-escrito de CTFL e a derivação da teoria das categorias semânticas.

Preliminares: fundamentos da metamatemática

É essencial estabelecer, de antemão, o sistema fundacional matemático que servirá de lastro teórico para o desenvolvimento das investigações de caráter metodológico – em especial, das investigações semânticas. Faz muita diferença partir de uma teoria dos tipos (especialmente, uma em que não ocorrem tipos de ordem transfinita, como veremos) e partir de uma teoria axiomática dos conjuntos (como, por exemplo, ZF ou BGN). Praticamente todas as investigações metamatemáticas de Tarski do período anterior à Segunda Guerra Mundial, são desenvolvidas no âmbito da teoria dos tipos. Em diversos artigos, ele remete ao sistema fundacional dos *Principia Mathematica* de Russell e Whitehead, embora seja sempre a versão simples que é assumida como arcabouço teórico de suas pesquisas. Essa influência da teoria dos tipos é tão considerável, que se estende mesmo a artigos do período pós-guerra⁸⁴. Na maioria das vezes, contudo, Tarski cuida em assinalar que seus resultados são extensíveis, *mutatis mutandis*, às teorias não baseadas na teoria dos tipos (como as baseadas em alguma axiomatização da teoria dos conjuntos) – mesmo sabendo que a “extensibilidade *mutatis mutandis*” nem sempre é uma matéria tão simples.

Na primeira publicação de CTFL (1933), as investigações de Tarski desenvolvem-se no interior da chamada ‘teoria das categorias semânticas’ (que abreviaremos por TCS) – segundo ele, uma extensão da teoria simples dos tipos. Um pequeno pós-escrito, no entanto, foi acrescentado à tradução alemã (1936), em que Tarski revisa sua posição inicial e considera a aplicação de seus resultados a teorias não fundadas em TCS, embora ainda baseadas, de alguma forma, na teoria dos tipos (de fato, numa versão em que ocorrem tipos de ordem transfinita). Reservamos este capítulo para expor, de maneira mais ou menos esquemática, como TCS está inserida nas investigações de CTFL, e qual a

⁸⁴ Por exemplo, ‘A problem concerning the notion of definability’ (1949).

relação existente entre ela e o problema da definição do conceito de verdade. Passaremos em revista também, rapidamente, as razões e as conseqüências do relaxamento de TCS no pós-escrito de CTFL.

A teoria das categorias semânticas de Husserl-Lesniewski

É difícil reconstituir o que seja exatamente a teoria das categorias semânticas, se tomarmos como base apenas o texto de CTFL. Tudo o que temos é uma apresentação informal de alguns conceitos e princípios básicos de TCS, os quais Tarski introduz até onde lhe interessa no que diz respeito ao problema central de sua monografia. De qualquer modo, não sendo nossa intenção desenvolver um estudo aprofundado de TCS *per se*, vamo-nos contentar com o esboço oferecido por Tarski em CTFL.

O conceito de categoria semântica é atribuído a Husserl, mas, segundo Tarski, o desenvolvimento formal de TCS deve-se a Lesniewski⁸⁵. O conceito de categoria semântica é análogo ao conceito Russelliano de tipo (de fato, como dissemos há pouco, Tarski afirma que, do ponto de vista formal, TCS é apenas uma extensão da teoria simples dos tipos). Há apenas uma diferença genética entre essas duas noções: o conceito de tipo surgiu com finalidade profilática nas pesquisas sobre os fundamentos da matemática, ao passo que o conceito de categoria semântica é uma abstração da noção intuitiva de ‘partes da fala’, e está calcado em nossas intuições lingüísticas mais fundamentais sobre a significatividade das expressões (pelo menos assim pensava Tarski à época da primeira versão de CTFL). Vejamos as informações mais formais que Tarski fornece sobre TCS. Tais informações podem ser distribuídas em seis itens.

(1) O conceito de categoria semântica. Em primeiro lugar, Tarski define a seguinte relação: (i) duas expressões *a* e *b* pertencem à mesma categoria semântica, se, e somente se, *a* e *b* são inter-substituíveis numa fórmula bem formada ϕ *salva beneformatione* (ou, talvez fosse mais coerente dizer, *salva significatione*). Então, em seguida, ele introduz o conceito de categoria semântica, da seguinte maneira: (ii) sendo a relação de pertencer à mesma categoria semântica uma relação de equivalência, define-se a categoria semântica de uma expressão *a* como a classe de equivalência determinada por *a* *modulo* essa relação.

⁸⁵ Tarski não indica a fonte Husserliana de TCS, mas sabe-se que é em suas *Logische Untersuchungen* que Husserl introduz a noção de categoria semântica, ou *Bedeutungskategorie*. Quanto a Lesniewski, Tarski remete aos *Grundzüge eines neuen Systems der Grundlagen der Mathematik* (1929). Ele faz também a remissão a um livro de Ajdukiewicz, *Main principles of the methodology of the sciences and of formal logic* (1928, em polonês).

(2) O primeiro princípio da teoria das categorias semânticas. A fim de determinar se duas expressões pertencem à mesma categoria semântica, é necessário verificar a inter-substituibilidade em *todas* as fórmulas bem formadas em que uma delas ocorre, ou é suficiente verificar se esse é o caso apenas em *algumas* ou mesmo em *uma só* fórmula bem formada? Segundo Tarski, o uso ordinário da linguagem força-nos a esta última alternativa. Temos então o seguinte:

Primeiro Princípio de TCS. Se (i) existe pelo menos uma fórmula bem formada ϕ em que uma das expressões a ou b ocorre, e (ii) as expressões a e b são inter-substituíveis em ϕ *salva beneformatione* (ou *salva significatione*), então a e b pertencem à mesma categoria semântica.

(3) Uma lei geral concernente aos funtores formadores de sentenças. Em CTFL, Tarski emprega uma terminologia que hoje não está mais em voga, e que, aliás, mesmo àquela época, era uma idiossincrasia dos poloneses. Por influência de Kotarbinski e Ajdukiewicz, Tarski utiliza bastante o termo ‘funtor’. Funtores são expressões insaturadas que, uma vez preenchidos seus lugares de argumento, formam expressões nominais ou expressões sentenciais – no primeiro caso, temos um ‘funtor formador de nomes’, e no segundo caso, temos um ‘funtor formador de sentenças’. Operadores são exemplos de funtores formadores de nomes. Há dois tipos básicos de funtores formadores de sentenças, a saber, aqueles que possuem argumentos nominais, e aqueles que possuem argumentos sentenciais. Os primeiros, que Tarski chama de ‘funtores formadores de funções sentenciais primitivas’, são os símbolos de predicado (constantes ou variáveis) usados na construção de fórmulas atômicas. Os segundos são os conectivos sentenciais e os quantificadores⁸⁶. No âmbito de TCS, há um princípio importante envolvendo os funtores formadores de funções sentenciais primitivas (chamemo-lo de ‘lei dos funtores primitivos’), que Tarski enuncia do seguinte modo:

Lei dos Funtores Primitivos. Dois funtores primitivos pertencem à mesma categoria semântica, se, e somente se, (1) possuem a mesma aridade, e (2) dois argumentos quaisquer que ocupam a mesma posição no escopo dos funtores, pertencem à mesma categoria semântica.

(4) Ordem de uma categoria semântica. Denotemos por ‘ $o(a)$ ’ a ordem da categoria semântica de uma determinada expressão a . Temos então a seguinte definição: (i) $o(a) = 1$, se, e somente se, a é uma constante ou variável individual (na terminologia original, se

⁸⁶ A propósito, cf. CTFL, p. 161, nota 1.

a é o ‘nome de um indivíduo’ ou uma ‘variável representando o nome de um indivíduo’); $o(a) = n+1$, se, e somente se, a é uma constante ou variável de predicado tal que (1) seus argumentos são no máximo de ordem n , e (2) um desses argumentos é, de fato, de ordem n (na terminologia original, a é um ‘funtor formador de funções sentenciais primitivas’).

Essa é uma das principais definições concernentes a TCS, e de profundas (muito profundas) conseqüências metamatemáticas, como veremos mais adiante. Observe que, por definição, a ordem de uma categoria semântica é sempre finita. Desde já, fica eliminada a possibilidade de assinalarmos um valor transfinito à ordem de uma categoria.

(5) Tipos semânticos. Já antecipamos uma noção preliminar e simplificada de tipo semântico no capítulo anterior. Trata-se de uma noção muito importante em TCS, e, como veremos mais adiante, intimamente relacionada ao problema central de CTFL. Preservando a notação do capítulo anterior, temos o seguinte:

$\phi \equiv_t \psi \Leftrightarrow_{df.}$ existe uma bijeção do conjunto das variáveis livres de ϕ no conjunto das variáveis livres de ψ , tal que, a cada variável livre de ϕ , corresponde uma variável livre de ψ *de mesma categoria semântica*.

Por exemplo, se ϕ possui cinco variáveis livres, duas pertencentes a uma categoria semântica a e três pertencentes a uma categoria semântica b , e se ψ , igualmente, possui cinco variáveis livres, duas pertencentes à categoria semântica a e três pertencentes à categoria semântica b , então ϕ e ψ pertencem à mesma categoria semântica. Agora, sendo ‘ $\equiv_t$ ’ uma relação de equivalência, temos:

$$[\phi]_t =_{df.} \{\psi \in \text{fbf} : \psi \equiv_t \phi\}$$

(6) A contraparte ontológica de TCS. A teoria das categorias semânticas versa diretamente sobre a *significatividade de expressões*, e não sobre os objetos eventualmente relacionados a tais expressões. Ela estabelece leis sobre *expressões*, e não sobre *objetos*. De maneira oblíqua, no entanto, a teoria das categorias semânticas produz uma delimitação formal de sua ontologia. Como? Ao fixar as regras lingüísticas de significatividade das expressões, a teoria das categorias semânticas acaba por determinar o tipo de ontologia que lhe convém. Por exemplo, uma das conseqüências da lei dos funtores primitivos é que não existem classes heterogêneas de objetos. Ou melhor, falar de tais classes é um contra-senso do ponto de vista da teoria das categorias semânticas. Pela mesma razão, é um contra-senso lingüístico falar em seqüências heterogêneas de

objetos (pois isso pressupõe a possibilidade de se falar em classes heterogêneas de objetos como contradomínio de uma função com domínio em N). Eis outras consequências ontológicas de TCS, mencionadas por Tarski:

- (1) Todos os indivíduos pertencem à mesma “categoria semântica”, mas nenhuma classe ou relação pertence à “categoria semântica” dos indivíduos;
- (2) Duas classes pertencem à mesma “categoria semântica”, se, e somente se, todos os seus elementos pertencem a uma única e mesma “categoria semântica”;
- (3) Duas relações diádicas pertencem à mesma “categoria semântica”, se, e somente se, seus domínios (e respectivos contradomínios) pertencem à mesma “categoria semântica”;
- (4) Duas seqüências pertencem à mesma “categoria semântica”, se, e somente se, todos os seus termos pertencem à mesma “categoria semântica”;
- (5) Uma classe e uma relação, ou duas relações de aridades distintas, nunca pertencem à mesma “categoria semântica”;

As aspas em torno da expressão ‘categoria semântica’ servem para indicar que a estamos usando numa acepção derivada (“hipostasiada”, como diz Tarski), já que originalmente ela se aplica a expressões, e não a objetos relacionados a tais expressões.

Tipologia Tarskiana das linguagens formalizadas fundadas em TCS

Pois bem, uma vez tendo esboçado TCS em seus traços gerais (e não há mais em CTFL do que os seis itens mencionados na seção anterior⁸⁷), Tarski propõe uma classificação das linguagens formalizadas, que deve ser cuidadosamente levada em conta quando estiver em jogo o estabelecimento de sua semântica. As linguagens são divididas em quatro espécies, de acordo com o esquema abaixo:

- **Linguagens de espécie 1:** em tais linguagens, todas as variáveis pertencem a uma única e mesma categoria semântica;
- **Linguagens de espécie 2:** em tais linguagens, o número de categorias semânticas nas quais as variáveis estão incluídas, é maior do que 1, mas finito;
- **Linguagens de espécie 3:** em tais linguagens, há infinitas categorias semânticas distintas às quais as variáveis pertencem, mas a ordem

⁸⁷ À exceção de umas poucas observações em notas de rodapé.

dessas variáveis não pode exceder um número natural n previamente dado;

- **Linguagens de espécie 4:** tais linguagens contêm variáveis de ordem arbitrariamente alta (e, por conseguinte, um número infinito de categorias semânticas de variáveis).

As linguagens de espécies 1, 2 e 3 formam a classe das **linguagens de ordem finita**. De acordo com a ordem mais alta das variáveis que ocorrem numa determinada linguagem L de ordem finita, L pode ser de primeira ordem, de segunda ordem, de terceira ordem etc. As linguagens de espécie 4 formam a classe das **linguagens de ordem infinita**.

A linguagem do *cálculo de classes*⁸⁸, que Tarski utiliza na ilustração de seu método de definição do conceito de verdade, é uma linguagem de espécie 1. A única categoria semântica de variáveis que ela contém, é a referente às variáveis que percorrem classes de indivíduos. Trata-se de uma linguagem de segunda ordem, já que indivíduos são objetos de primeira ordem no âmbito de TCS. Uma outra linguagem que Tarski classifica como sendo de espécie 1, é a linguagem do *cálculo proposicional estendido*, cujas variáveis percorrem apenas proposições. Trata-se de uma linguagem de primeira ordem, se assumirmos que as proposições são os indivíduos que compõem o domínio de quantificação.

Como exemplo de linguagem de espécie 2, Tarski considera a linguagem da *lógica de relações diádicas*⁸⁹, cujas variáveis distribuem-se em duas categorias: aquelas que percorrem indivíduos, e aquelas que percorrem relações diádicas entre indivíduos. Trata-se de uma linguagem de segunda ordem, já que a ordem mais alta é a das variáveis relacionais, que percorrem objetos de segunda ordem.

A linguagem da *lógica de relações poliádicas*⁹⁰ é um exemplo de linguagem de espécie 3. Ela contém um número infinito (denumerável) de categorias semânticas de variáveis: variáveis de indivíduos, de classes de indivíduos, de relações diádicas entre indivíduos, e, de um modo geral, para todo $n \in \mathbb{N}$, variáveis de relações n -ádicas entre indivíduos. Trata-se de uma linguagem de segunda ordem, já que todas as variáveis relacionais percorrem relações entre indivíduos.

⁸⁸ O cálculo de classes formaliza uma álgebra de Boole cujo domínio é formado por classes de indivíduos. O sistema utilizado por Tarski em CTFL contém o símbolo de inclusão como único primitivo e é caracterizado por cinco axiomas próprios: (1) reflexividade da inclusão; (2) transitividade da inclusão; (3) existência de supremo para qualquer par de classes; (4) existência de ínfimo para qualquer par de classes; (5) existência de uma classe complementar para qualquer classe. A propósito, cf. CTFL, p. 179.

⁸⁹ Para a caracterização axiomática da lógica de relações diádicas, cf. CTFL, p. 223. Tarski dedicou um artigo inteiro a essa teoria: 'On the calculus of relations', JSL, v. 6, n. 3, 1941. No entanto, a abordagem desse artigo não toca em aspectos semânticos.

⁹⁰ Para a caracterização axiomática da lógica de relações poliádicas, cf. CTFL, p. 232.

Finalmente, Tarski exemplifica a espécie 4 com a linguagem da *teoria geral das classes*, cujas variáveis, em número infinito, percorrem indivíduos, classes de indivíduos, classes de classes de indivíduos, e assim por diante. Como não há limite superior natural para a ordem das variáveis, trata-se de uma linguagem de ordem infinita. É importante observar que a teoria geral das classes, segundo Tarski, constitui uma base suficiente para o estabelecimento de várias partes da matemática e, em particular, da aritmética dos números naturais. Em outras palavras, a teoria geral das classes é uma teoria projetada para servir de base fundacional da matemática (pelo menos até certa extensão)⁹¹.

A ambigüidade semântica do conceito de satisfação

O problema central de CTFL – construir uma definição formalmente correta e materialmente adequada do conceito de verdade – remete a três outros problemas, cuja solução constitui o cerne de toda a monografia de Tarski: *em primeiro lugar*, há o problema geral das antinomias semânticas (veja-se o Capítulo 1); *em segundo lugar*, há a impossibilidade de se construir uma definição recursiva direta do conceito de verdade (veja-se o Capítulo 3); finalmente, *em terceiro lugar*, há o problema da ambigüidade semântica do conceito de satisfação, do qual nos ocuparemos nesta seção.

Chegamos agora ao núcleo deste capítulo. Examinaremos que dificuldades a teoria das categorias semânticas opõe à solução do problema central de CTFL. O entrave básico é o que Tarski chama de ‘ambigüidade semântica’ da noção de satisfação. Intuitivamente, dizer que a noção de satisfação é semanticamente ambígua, nada mais significa do que dizer que existem diversas (na verdade, denumeravelmente infinitas) noções de satisfação. Veremos que o problema da ambigüidade do conceito de satisfação está intimamente relacionado com a noção já discutida de ‘tipo semântico’ de uma fórmula.

Recordemos que o tipo semântico de uma fórmula depende do número de suas variáveis livres e de suas respectivas categorias semânticas: duas fórmulas ϕ e ψ possuem o mesmo tipo semântico, se, e somente se, existe uma bijeção do conjunto das variáveis livres de ϕ no conjunto das variáveis livres de ψ , tal que, a cada variável livre de ϕ , corresponde uma variável livre de ψ de mesma categoria semântica. Pois bem, de que modo a noção de tipo semântico está relacionada com a ambigüidade semântica da noção de satisfação? Nossa resposta será dada em duas etapas. Consideraremos, em primeiro lugar, as linguagens baseadas em TCS cujas variáveis pertencem à mesma categoria semântica; em segundo lugar, consideraremos as linguagens baseadas em TCS cujas variáveis pertencem a categorias semânticas distintas. Em cada um desses casos, a noção

⁹¹ Para a caracterização axiomática da teoria geral das classes, cf. CTFL, p. 243.

de tipo semântico está relacionada com uma espécie de ambigüidade da noção de satisfação. Veremos que, no primeiro caso, a ambigüidade surge em virtude de uma fórmula poder conter um número arbitrário de variáveis livres. E veremos que, no segundo caso, a ambigüidade surge não apenas por causa do número arbitrário de variáveis livres que podem ocorrer numa fórmula, mas também por conta da possibilidade de tais variáveis pertencerem a categorias semânticas distintas.

1.º Caso

O caso das linguagens baseadas em TCS cujas variáveis pertencem a uma única categoria semântica (a saber, as linguagens de espécie 1), é tratado segundo o procedimento que foi delineado no capítulo anterior. Antecipamos ali (implicitamente) o modo como a ambigüidade da noção de satisfação está atrelada à noção de tipo semântico. A ambigüidade semântica da noção de satisfação é removida nessas linguagens pelo emprego de seqüências denumeráveis de objetos (supondo que estes sejam todos de mesma categoria semântica). Chamemos o procedimento ali empregue de ‘método das seqüências simples’.

2.º Caso

Vejamos o caso das linguagens baseadas em TCS cujas variáveis pertencem a categorias semânticas distintas, que é mais complicado. Esse é o caso, de acordo com a tipologia de Tarski, das linguagens de espécies 2, 3 e 4. Restrinjamos nossa atenção, inicialmente, às linguagens de espécies 2 e 3. Considere, por exemplo, as linguagens da lógica de relações diádicas e da lógica de relações poliádicas. Vimos que as variáveis da primeira estão divididas em duas categorias semânticas: há as variáveis que percorrem indivíduos (variáveis individuais), e há as variáveis que percorrem relações diádicas entre indivíduos (variáveis relacionais). Vimos também que as variáveis da segunda estão divididas em denumeravelmente infinitas categorias semânticas: há as variáveis que percorrem indivíduos, as variáveis que percorrem classes de indivíduos, relações diádicas entre indivíduos, relações ternárias entre indivíduos, e assim por diante. O seguinte vale para ambas as linguagens: em algumas fórmulas, as variáveis livres pertencem todas à mesma categoria semântica; em outras fórmulas, por outro lado, as variáveis livres pertencem a categorias semânticas distintas. Para fixar uma terminologia provisória, chamemos o tipo semântico de uma fórmula de ‘homogêneo’ ou ‘heterogêneo’, conforme as variáveis pertençam à mesma categoria ou a categorias distintas, respectivamente; além disso, se uma fórmula é de tipo semântico homogêneo, chamemo-lo de ‘individual’ ou ‘ n -relacional’, conforme as variáveis sejam todas individuais ou percorram relações n -ádicas, respectivamente.

Pois bem, a cada um dos tipos semânticos acima descritos, corresponde uma noção de satisfação distinta. Aos tipos semânticos homogêneos individuais, corresponde a noção de satisfação por uma seqüência denumerável de indivíduos; aos tipos semânticos homogêneos n -relacionais, corresponde a noção de satisfação por uma seqüência denumerável de relações n -ádicas entre indivíduos; finalmente, aos tipos semânticos heterogêneos, corresponde a noção de satisfação por uma seqüência denumerável de objetos de diversas naturezas (no caso, indivíduos ou relações de diversas aridades). Note que há um número denumeravelmente infinito de tipos semânticos heterogêneos (e, portanto, infinitas noções de satisfação), uma vez que não há limite para o número de ocorrências de variáveis de categorias distintas numa fórmula. O último caso é particularmente grave, porque, como vimos, conflita com o fato de que, no âmbito de TCS, simplesmente não pode haver seqüências heterogêneas de objetos: com efeito, tais seqüências são aquelas cujo contradomínio é uma classe que contém objetos de diversas naturezas, e tais classes não fazem parte da hierarquia de tipos em que TCS se funda⁹².

Tarski apresenta dois métodos para solucionar o problema da ambigüidade semântica da noção de satisfação (em linguagens de espécies 2 e 3), a saber, (i) o que ele chama de ‘método das seqüências múltiplas’ (*many-rowed sequences*), e (ii) o que ele chama de ‘método da unificação semântica das variáveis’. Não vamos entrar em detalhes técnicos sobre esses métodos. As seguintes informações são suficientes.

Seqüências múltiplas e linguagens de espécie 2. A aplicação do método das seqüências múltiplas à lógica de relações diádicas, é bastante simples. Em vez de seqüências simples, operamos com pares ordenados de seqüências $\langle s, S \rangle$ (seqüências ‘duplas’ – *two-rowed*), sendo s uma seqüência denumerável de indivíduos, e S uma seqüência denumerável de relações diádicas entre indivíduos. Com o uso de tais seqüências, obtemos um conceito não-ambíguo de satisfação, a saber, o conceito de satisfação de uma fórmula por seqüências duplas. Assim, o problema está resolvido para a lógica de relações diádicas.

Seqüências múltiplas e linguagens de espécie 3. O método das seqüências múltiplas não se aplica (pelo menos diretamente) à lógica de relações poliádicas. A razão é simples: uma vez que existem infinitas categorias semânticas de variáveis em tal linguagem, teríamos que lançar mão de seqüências múltiplas com um número infinito de termos, cada termo diferindo de todos os outros quanto à categoria semântica. Mas já sabemos que não há tais seqüências no âmbito de TCS.

⁹² Cf. CTFL, p. 219, *in fine*.

Unificação semântica e linguagens de espécie 2. A aplicação do método da unificação semântica das variáveis à lógica de relações diádicas, também é simples. Em primeiro lugar, a cada indivíduo a , associamos a relação diádica $a^* = \{\langle a, a \rangle\}$. Depois, substituímos cada variável individual por uma variável relacional diádica. Reduzimos, assim, as duas categorias semânticas de variáveis a uma só (isto equivale a uma unificação total das variáveis). Então, aplicamos o mesmo método usado nas linguagens de espécie 1, isto é, o método das seqüências simples. Uma exigência, no entanto, deve ser atendida na aplicação deste método: a ordem da categoria unificadora deve ser no mínimo igual à maior ordem das categorias que ocorrem na linguagem original.

Unificação semântica e linguagens de espécie 3. Agora, vejamos como remover a ambigüidade da noção de satisfação na lógica de relações poliádicas. Denominemos ‘grafo-seqüência’ de uma relação n -ádica R de indivíduos, ao conjunto de todas as seqüências finitas de n indivíduos para os quais R vale⁹³. Observe que duas seqüências quaisquer de indivíduos pertencem à mesma “categoria semântica” – mesmo que o número de termos de uma seja diferente do número de termos da outra. Por conseguinte, mesmo que duas relações entre indivíduos diverjam quanto ao número de termos, seus grafos-seqüência correspondentes pertencem à mesma “categoria semântica”. Esse fato é fundamental para solucionar o problema da ambigüidade semântica da noção de satisfação em linguagens de espécie 3. No caso da linguagem da lógica de relações poliádicas, o procedimento de Tarski é o seguinte. A cada relação n -ádica R , associamos seu correspondente grafo-seqüência. Depois, substituímos cada variável relacional por uma variável que percorre grafos-seqüência. Reduzimos, assim, as infinitas categorias semânticas de variáveis a apenas duas categorias semânticas (isto equivale a uma unificação *parcial* das variáveis). Então, aplicamos finalmente o mesmo procedimento de remoção de ambigüidade que foi usado nas linguagens de espécie 2 – o método das seqüências múltiplas ou, novamente, o método da unificação semântica (desta vez, *total*) das variáveis. A mesma exigência feita à aplicação deste último método às linguagens de espécie 2, deve ser atendida aqui: a ordem da categoria unificadora deve ser no mínimo igual à maior ordem das categorias que ocorrem na linguagem original.

⁹³ Na verdade, o conceito de grafo-seqüência de uma relação n -ádica só difere do conceito usual de grafo, porque, em vez de n -uplas ordenadas, consideramos seqüências finitas.

Linguagens de ordem infinita e a versão absoluta do teorema da indefinibilidade da verdade

Para as linguagens de ordem infinita, que são as linguagens de espécie 4 (e a linguagem da teoria geral das classes, em particular), o problema da ambigüidade semântica da noção de satisfação, mostra-se insolúvel – e portanto, *ipso facto*, o problema de uma definição satisfatória do conceito de verdade em tais linguagens. Nenhum dos métodos aventados acima podem ser aplicados de maneira eficaz para a remoção da ambigüidade do conceito de satisfação nessas linguagens. Por um lado, o método das seqüências múltiplas mostra-se inoperante pelas mesmas razões por que o é com respeito às linguagens de espécie 3, isto é, porque o número de categorias semânticas a que pertencem suas variáveis, é denumeravelmente infinito. Por outro lado, vimos que a aplicação do método da unificação semântica das variáveis às linguagens de espécie 3, só é possível, porque existe um limite natural superior à ordem de suas variáveis (lembramos que a escolha da categoria unificadora deve levar em conta esse limite). No entanto, em linguagens de espécie 4, esse limite simplesmente não existe (é por isso que Tarski as denomina de ‘linguagens de ordem *infinita*’). A escolha da categoria unificadora das variáveis, que, como vimos, deve ser de ordem no mínimo igual à maior ordem das variáveis da linguagem-objeto, revela-se, assim, impossível. A solução, que seria lançar mão de variáveis de ordem transfinita (neste caso, de ordem ω), não está aberta às linguagens baseadas em TCS, pois, como vimos mais atrás, estas não dispõem de tais variáveis⁹⁴.

Alguém poderia sugerir o seguinte: se a noção *geral* de satisfação é indefinível para linguagens de ordem infinita, uma noção *especial* dessa mesma noção não solucionaria o problema? Essa sugestão parte de uma idéia simples. Tarski afirma que se dispuséssemos da definição de uma noção especial de satisfação, como, por exemplo, satisfação por objetos isolados, poderíamos definir facilmente o conceito de sentença verdadeira. Fariamos duas coisas. Em primeiro lugar, tomaríamos uma fórmula ϕ com uma variável livre, que fosse satisfeita por qualquer objeto (poderia ser, digamos, ‘ $x = x$ ’). Então, depois, definiríamos o conceito de sentença verdadeira, do seguinte modo: s é uma sentença verdadeira, se, e somente se, todo objeto satisfaz ‘ $s \wedge \phi$ ’. Note que a fórmula ‘ $s \wedge \phi$ ’ contém apenas uma variável livre, e que nosso conceito de satisfação se aplica a fórmulas com uma única variável livre. Como todo objeto satisfaz ϕ , a satisfação de ‘ $s \wedge \phi$ ’ por todos os objetos depende não de ϕ , mas apenas de s . Segue então que uma tal

⁹⁴ Cf. CTFL, p. 244.

definição do conceito de verdade seria materialmente adequada (capturando, assim, a propriedade expressa no Esquema T). Mas, infelizmente, não é possível construir essa definição. Por quê? Porque já vimos (no capítulo anterior) que não é possível definir um conceito especial de satisfação para um dado tipo semântico, sem antes ter definido o conceito geral de satisfação, aplicável a todos os tipos semânticos. E o conceito geral de satisfação é indefinível em linguagens de ordem infinita.

Pode parecer que a insolubilidade do problema de uma definição do conceito de verdade em linguagens de espécie 4, deve-se antes a limitações específicas dos métodos disponíveis. Uma eventual descoberta de outro método poderia mostrar que esse problema tem solução. Se o entrave básico está relacionado à noção de satisfação, então talvez se pudesse pensar que a saída seria buscar um método que não a empregue. Isto é, deixemos de lado a noção de satisfação! Mas não é esse o caso. Qualquer tentativa de encontrar outro método definicional para o conceito de verdade em linguagens de ordem infinita está destinada a fracassar. Por quê? A resposta está naquilo que ficou conhecido como o ‘teorema da indefinibilidade da verdade’. Esse teorema (de fato, um ‘meta-meta-teorema’) pode ser formulado da seguinte maneira:

Teorema da indefinibilidade da verdade. Se assumirmos a consistência da meta-teoria, então o conceito de verdade na linguagem da teoria geral das classes (e, por extensão, em qualquer linguagem de ordem infinita), é indefinível a partir dos conceitos primitivos da meta-teoria⁹⁵.

Uma consequência metodológica do teorema da indefinibilidade da verdade, de extrema relevância, é que ele implica o primeiro teorema de incompletude de Gödel (vide Apêndice 2). No contexto de CTFL, essa consequência é informalmente considerada e enunciada assim: a teoria geral das classes não só é *incompleta*, mas também *incompletável*⁹⁶. Alternativamente falando: *o conjunto das sentenças verdadeiras na linguagem da teoria geral das classes não é axiomatizável*. Uma outra consequência metodológica, igualmente relevante, será discutida na próxima seção – trata-se da teoria axiomática da verdade. Para finalizar, lembremos que Tarski demonstra um teorema que ameniza o caráter limitativo do teorema da indefinibilidade da verdade. Segundo esse teorema, embora o conceito de verdade em linguagens de ordem infinita seja absolutamente indefinível, ele pode ser “fragmentariamente” definido para aquelas porções em que ocorrem variáveis de ordem no máximo igual a k , para um número

⁹⁵ Cf. teorema I do §5 de CTFL, p.247. Por razões de simplicidade, nossa formulação acima é ligeiramente diferente da formulação original, mas o conteúdo é o mesmo. Observe que, em contraste com a versão hoje conhecida (vide Apêndice 2), a versão original do teorema da indefinibilidade da verdade é dada em termos de definibilidade *sintática* e não em termos de definibilidade *semântica*.

⁹⁶ Cf. CTFL, último parágrafo da p. 254. Note que não há remissão explícita a Gödel.

natural k arbitrariamente escolhido⁹⁷. Em outras palavras, existe uma sequência infinita de definições fragmentárias do conceito de verdade em linguagens de ordem infinita, que converge, por assim dizer, para uma “definição-limite” completa do conceito de verdade em tais linguagens.

O quadro abaixo resume esquematicamente o que foi discutido nessas duas últimas seções.

Linguagens	Método para definir a noção de satisfação
Espécie 1	seqüências simples
Espécie 2	seqüências múltiplas / unificação das variáveis
Espécie 3	unificação das variáveis
Espécie 4	inexistente (teorema da indefinibilidade)

Teoria axiomática da verdade

Recapitulando: no âmbito de TCS, não há uma linguagem essencialmente mais rica do que as linguagens de ordem infinita. Esse fato segue trivialmente da inexistência de variáveis de ordem transfinita no estoque de símbolos das linguagens baseadas em TCS. O teorema da indefinibilidade da verdade, em sua primeira versão, mostra que, se tentarmos definir o conceito de verdade para linguagens de ordem infinita, caímos em contradição. Por conseguinte, a via definicional está fechada para a semântica de tais linguagens. A consequência metodológica mais imediata do teorema da indefinibilidade da verdade, então, é que a única saída para estabelecer os fundamentos da semântica de linguagens de ordem infinita, é a via axiomática. Mas qual o procedimento de construção de uma teoria axiomática da verdade em tais linguagens?

Seja L uma linguagem de ordem infinita (por exemplo, a linguagem da teoria geral das classes). Denominamos um sistema de axiomas em que o conceito de verdade em L é assumido como primitivo semântico, de ‘teoria da verdade em L ’ – em símbolos, TV_L . Inicialmente, TV_L consiste dos axiomas da meta-teoria acrescidos de todas as instâncias do Esquema T para as sentenças de L . Tarski mostra que, assumida a consistência da meta-teoria original, tal extensão também é consistente⁹⁸. A principal limitação de TV_L , no entanto, é a não-demonstrabilidade em seu interior das mais importantes propriedades do conceito de verdade (portanto, a incompletude de TV_L). Por exemplo, os princípios de não-contradição e terceiro excluído não são demonstráveis em TV_L . Tarski aponta duas

⁹⁷ Cf. teorema II do §5 de CTFL, p. 255.

⁹⁸ CTFL, teorema III do §5, pp. 256-257.

alternativas para incrementar o poder dedutivo de TV_L . A primeira alternativa explora a *base axiomática* de TV_L : expandimos o conjunto de seus axiomas por meio do acréscimo de algumas propriedades não-demonstráveis em seu interior. Além de todas as instâncias do Esquema T, podemos adicionar a TV_L , por exemplo, os seguintes axiomas: (1) os princípios de não-contradição e terceiro excluído; (2) o princípio segundo o qual as conseqüências de sentenças verdadeiras são também verdadeiras; (3) um axioma garantindo a verdade dos axiomas da teoria com suporte lingüístico em L; entre outros. Há, contudo, segundo Tarski, alguns problemas com uma tal extensão de TV_L . Em primeiro lugar, cada expansão do conjunto original dos axiomas de TV_L é puramente acidental e arbitrária, porque depende do estágio atual de nosso conhecimento sobre as propriedades do conceito de verdade. Pode haver propriedades importantes que nem sequer divisamos ainda. Em segundo lugar, há o problema da categoricidade de cada extensão dos axiomas de TV_L : nenhuma delas é categórica, e, portanto, nenhuma delas caracteriza de maneira não-ambígua a extensão do conceito de verdade em L. Isso frustra um dos requisitos fundamentais para Tarski de qualquer formalização do conceito de verdade pela via axiomática⁹⁹. Por conseguinte, a primeira alternativa ao problema da completude de TV_L não é uma boa solução.

A segunda alternativa explora o *aparelho inferencial* de TV_L : além de suas regras comuns (finitárias) de inferência, acrescentamos-lhe o que Tarski chama de ‘regra de indução infinita’ (na verdade, uma versão da regra ω). Seja $s_1, s_2, \dots, s_n, \dots$ uma enumeração de todas as sentenças de L. Se cada uma das sentenças a seguir é demonstrável em TV_L : $P(s_1), P(s_2), \dots, P(s_n), \dots$, sendo $P(x)$ uma fórmula metalingüística qualquer, *então* a generalização $\forall i (i \in \omega \rightarrow P(s_i))$ é demonstrável em TV_L . De acordo com Tarski, essa alternativa é muito mais vantajosa do que a alternativa anterior. Pelo menos, sob dois aspectos. Em primeiro lugar, os princípios de não-contradição e terceiro excluído (entre outras propriedades relevantes e conhecidas do conceito de verdade, como as mencionadas há pouco) passam a ser demonstráveis em TV_L , não sendo necessário expandir o conjunto original de seus axiomas. Em segundo lugar, e mais importante ainda, a teoria que resulta de TV_L pela introdução da regra de indução infinita é categórica – por conseguinte, satisfaz o requisito de determinação não-ambígua da extensão do conceito de verdade em L. Entretanto, a segunda alternativa não deixa de ter seus problemas – e problemas graves. Há três dificuldades básicas. A primeira é que, apesar de TV_L tornar-se categórica, ela continua incompleta. A segunda diz respeito à natureza infinitária da regra de indução infinita, que, segundo Tarski, não se ajusta à

⁹⁹ Cf. Capítulo 2.

concepção corrente (finitária) do método dedutivo¹⁰⁰. A terceira e mais importante dificuldade é a seguinte: se os fundamentos das investigações metamatemáticas sobre a semântica de linguagens de ordem infinita continuarem circunscritos a TCS, então uma demonstração da consistência de TV_L munida com a regra ω é um problema em aberto.

Por conta dos problemas associados a TV_L , Tarski mostra-se bastante antipático em relação ao uso do método axiomático em semântica. É interessante lembrar as objeções que ele levanta em ESS contra a axiomatização da semântica. São quatro as objeções. As duas primeiras são repetições de críticas feitas já em CTFL, críticas que acabamos de discutir. Apenas as duas últimas são críticas novas. As objeções são as seguintes, *verbatim*:

- (1) “A escolha dos axiomas sempre possui um caráter accidental, dependendo de fatores não essenciais (como o estágio atual de nosso conhecimento); vários critérios a que poderíamos recorrer para sanar esse problema, acabam por mostrar-se inaplicáveis;
- (2) O problema da consistência adquire importância especial no caso de uma teoria axiomática da verdade, por conta das antinomias semânticas;
- (3) O método axiomático na construção de uma teoria semântica não parece natural do ponto de vista psicológico, se neste método o papel dos conceitos primitivos – assim, dos conceitos cujo significado deveria parecer evidente – é desempenhado por conceitos que levaram a tantos mal-entendidos no passado;
- (4) Finalmente, se fosse provado que o método axiomático em semântica é o único possível, e isso não se devesse a apenas um estágio transitório, surgiriam certas dúvidas de um ponto de vista filosófico geral; de fato, seria bastante difícil conciliar esse método com os postulados da unidade da ciência e do fisicalismo (uma vez que os conceitos da semântica não seriam nem conceitos lógicos, nem conceitos físicos)”¹⁰¹.

Lembre-se que Tarski muitas vezes usa o termo ‘lógica’ em sentido amplo, i.e., incluindo não só o cálculo proposicional clássico e o cálculo de predicados de primeira ordem, como também algum sistema fundacional matemático, como, por exemplo, alguma versão da teoria dos tipos ou da teoria dos conjuntos¹⁰². É nesse sentido que está sendo usado termo na última objeção.

¹⁰⁰ Sobre esse ponto, veja SOC ω CC, p. 295.

¹⁰¹ ESS, pp. 405-406.

¹⁰² A propósito, cf. SCTFS, nota 12.

É interessante examinar a quarta objeção a uma teoria semântica axiomática (em particular, uma teoria da verdade), tendo em vista a primeira versão do teorema da indefinibilidade da verdade em CTFL. Como já sabemos, o conceito de verdade em linguagens de ordem infinita, no âmbito da teoria das categorias semânticas, é *absolutamente* indefinível. Em especial, o conceito de verdade na linguagem da teoria geral das classes, arcabouço fundacional de pelo menos parte da matemática clássica em CTFL, é *absolutamente* indefinível. Sendo assim, o conceito de verdade matemática é extremamente problemático do ponto de vista filosófico¹⁰³, na extensão em que tal conceito depende da teoria das categorias semânticas. Se o conceito de verdade matemática não é nem um conceito lógico (*sensu lato*), nem um conceito físico, então o conceito de verdade matemática é um conceito metafísico – por conseguinte, um conceito espúrio, um pseudo-conceito. Essa é uma consequência filosófica muito curiosa do teorema da indefinibilidade da verdade, se levarmos em consideração o fato de que CTFL foi muito bem acolhido pelos positivistas do Círculo de Viena. Mas é claro que esse acolhimento só foi possível às custas de uma interferência drástica na versão original de CTFL. Esse é o tema da próxima (e última) seção.

O pós-escrito de CTFL e a derrogação da teoria das categorias semânticas

Surpreendentemente, e em dissonância com um certo folclore acadêmico em torno do nome de Tarski, a primeira versão de CTFL tem um caráter extremamente *negativo* com relação à semântica, seja no que diz respeito à linguagem ordinária, seja no que diz respeito às linguagens formalizadas. No primeiro caso, todos já sabem, esse caráter decorre da constatação do fechamento semântico das linguagens naturais, como vimos no Capítulo 1. No último caso, como pouca gente lembra, esse caráter decorre de duas razões interligadas: (i) a indefinibilidade absoluta do conceito de verdade matemática; (ii) o problema da consistência de uma teoria axiomática da verdade matemática. Quando escrevemos ‘verdade matemática’, estamo-nos referindo ao conceito de verdade na linguagem da teoria geral das classes, que, como já assinalamos, pode ser visto como uma espécie de sistema fundacional matemático em CTFL. O item (i) é filosoficamente devastador, se prestarmos atenção à quarta objeção levantada por Tarski a uma teoria axiomática da verdade – a verdade matemática não é nem um conceito lógico (*sensu lato*), nem um conceito físico. O item (ii) é metodologicamente preocupante, se

¹⁰³ Segundo os critérios, é claro, do positivismo lógico, então dominante na filosofia austríaca e polonesa da década de trinta.

lembrarmos que o problema da consistência de uma teoria axiomática da verdade matemática, é um problema em aberto na primeira versão de CTFL. Mas a semântica de linguagens formalizadas é interessante especialmente quando diz respeito à matemática clássica. Então, o resultado da primeira versão da monografia de Tarski, trocando em miúdos, é o seguinte: a semântica de linguagens formalizadas, *quando é interessante*, é filosoficamente espúria ou extremamente problemática do ponto de vista metodológico (porque, afinal de contas, não sabemos se uma teoria axiomática da verdade, circunscrita ao ambiente de TCS, é consistente ou não).

A maneira de neutralizar o potencial destrutivo de CTFL foi abandonar a teoria das categorias semânticas. Tarski não diz em absoluto, explicitamente, por que abandonou a sua posição original pró-TCS (também não aduz, na primeira versão, as razões pró-TCS). É bem possível que tenha sido uma saída simplesmente técnica. E uma saída relativamente fácil: substitui-se TCS por uma teoria simples dos tipos em que ocorrem tipos de ordem transfinita. E pronto. Mas acreditamos que, acima de tudo, essa foi uma maneira de aplacar a força destrutiva da primeira versão de CTFL. A derrogação de TCS tem basicamente dois efeitos positivos imediatos. Em primeiro lugar, obtém-se uma relativização do teorema da indefinibilidade da verdade – o conceito de verdade matemática não é mais um conceito absolutamente indefinível. Com a admissão de tipos de ordem transfinita, a condição de riqueza essencial da metalinguagem em relação à linguagem-objeto, pode ser sempre satisfeita. Em segundo lugar, Tarski percebe que, derogada TCS, os métodos desenvolvidos em CTFL podem contrabalançar *positivamente* os resultados negativos obtidos por Gödel (e por ele mesmo), principalmente o primeiro teorema de incompletude da aritmética elementar. Isso é enfatizado não só no pós-escrito de CTFL, mas também em ESS e, muitos anos depois, em TP. Esse passa a ser o aspecto mais estimável, segundo Tarski, de seus métodos. A verdade de sentenças indecidíveis numa dada teoria, é decidida numa teoria mais forte, com o auxílio das técnicas desenvolvidas em CTFL. Mas esse préstimo só é possível às custas do abandono da teoria das categorias semânticas. A partir de então, CTFL mostra seu caráter positivo, com o qual hoje estamos habituados, e que justifica em parte o “folclore acadêmico” em torno do nome de Tarski, a que nos referimos no início desta seção. •

Capítulo 5 – *CrITÉrio de verdade*

SUMÁRIO. Preliminares: definição *versus* critério. A noção de critério de verdade nos artigos informais. A noção de critério de verdade em *The Concept of Truth*. A noção de critério de verdade numa classe de estruturas. Mais algumas indicações histórico-conceituais.

Preliminares: definição *versus* critério

É usual fazer-se uma distinção entre ‘definição’ e ‘critério’ de verdade. Tarski explicitamente diferencia as duas noções, como teremos ocasião de ver mais adiante. Essa distinção é apenas um caso particular da distinção entre ‘definição’ e ‘critério’ em geral. Intuitivamente, o contraste é fácil de ser explicado. Por exemplo, seja L uma linguagem adequada para a aritmética elementar dos números naturais. Considere um conjunto C de números naturais, definido por uma determinada fórmula ϕ de L . Por um ‘critério para C ’, entende-se um procedimento efetivo que nos permite decidir se, dado um número natural n , $n \in C$ ou $n \notin C$, ou, o que dá no mesmo, se n satisfaz ou não a fórmula que define C , i.e., a fórmula ϕ . Em outras palavras, um critério é simplesmente um teste mecânico de decisão. É claro que o mero fato de C estar definido em L por meio de ϕ , não é garantia suficiente da existência de um algoritmo de decisão para C . Há conjuntos de números naturais que, embora definíveis em L , são indecidíveis (como o famoso Conjunto da Parada¹⁰⁴, da teoria da recursão).

Não há dificuldades, então, em se considerar a distinção particular entre ‘definição’ e ‘critério’ de verdade. Seja ML a metalinguagem de uma dada linguagem-objeto LO . Já vimos que em ML definimos uma série de coleções específicas de expressões de LO , como o conjunto das fórmulas bem formadas de LO , o conjunto das sentenças de LO demonstráveis numa dada teoria T , entre várias outras coleções. Em especial, definimos em ML , por meio de uma fórmula ϕ , o conjunto V das sentenças verdadeiras de LO com respeito a uma determinada teoria T . Então, por um ‘critério de verdade’, entende-se um procedimento algorítmico que nos permite decidir se, dada uma sentença s de LO , $s \in V$ ou $s \notin V$, ou, o que dá no mesmo, se s satisfaz ou não a definição ϕ , ou, ainda, se s é ou

¹⁰⁴ Aproximadamente, o Conjunto da Parada é definido da seguinte maneira: seja Z um programa de Turing qualquer, e seja n um código de Z (dada uma enumeração recursiva conveniente de todos os programas); o número n pertence ao Conjunto da Parada, se, e somente se, a função unária f computada de acordo com Z está definida para n .

não verdadeira. Um critério de verdade é, assim, um ‘teste mecânico’ da verdade das sentenças de LO com respeito à teoria T. Analogamente ao que dissemos no parágrafo anterior, o simples fato de nós haveremos definido o conjunto V das verdades de LO, não nos garante a existência de um critério de decisão para V.

Não é o objetivo de Tarski, em CTFL, discutir formalmente a noção de critério de verdade, bem como o correspondente problema da existência de critério de verdade para a linguagem de uma teoria. O objetivo central de CTFL, como vimos, é fornecer um método geral para a construção de uma definição satisfatória (i.e., materialmente adequada e formalmente correta) do conceito de verdade em linguagens formalizadas. Entretanto, o problema da existência de um critério de verdade, embora esteja em segundo plano na monografia de Tarski, não é de somenos importância do ponto de vista metamatemático (já que se trata do problema da decisão de uma teoria, considerado sob um enfoque semântico), e em diversos momentos é mencionado em conexão com o problema da definição de verdade¹⁰⁵. Neste capítulo, empreenderemos basicamente duas tarefas: em primeiro lugar, discutiremos a noção de critério de verdade, tal como aparece nos artigos de divulgação da concepção semântica da verdade, e, evidentemente, em CTFL; em segundo lugar, trasladaremos a noção de critério de verdade, de seu contexto original *simpliciter*, para um contexto modelo-teorético. No final, retornaremos brevemente aos trabalhos originais de Tarski, a fim de dar algumas indicações para ulterior aprofundamento no assunto.

A noção de critério de verdade nos artigos informais

Três são os artigos de divulgação das investigações semânticas de Tarski: ‘The Establishment of Scientific Semantics’ (ESS, 1935), ‘The Semantic Conception of Truth and the Foundations of Semantics’ (SCTFS, 1944), e ‘Truth and Proof’ (TP, 1969). Como é apresentada nesses trabalhos a noção de critério de verdade?

Bem, em ESS, curiosamente, Tarski não discute o assunto. Em SCTFS e TP, a noção é brevemente examinada, mas sob uma perspectiva completamente estranha àquela presente em CTFL, que analisaremos mais adiante. Em ambos os artigos, é notável que a noção de critério de verdade é examinada sempre em conexão com um princípio metodológico de inspiração antimetafísica, especialmente elaborado para o controle do vocabulário de uma teoria científica (formal ou empírica). Para facilitar nossa discussão, enunciemos esse princípio da seguinte maneira:

¹⁰⁵ Cf. CTFL, pp. 197, 208-9, 222, 238, 238 (nota), 240-1, 254 (nota).

Princípio do Critério para Definições (PCD). Seja ϕ uma fórmula da linguagem de uma teoria T , que define um dado conjunto C de objetos do domínio de um modelo de T . Dizemos que ϕ é *criteriologicamente satisfatória*, se, e somente se, ϕ fornece um critério de decisão para C – i.e., se, com base em ϕ , podemos decidir efetivamente se um dado objeto pertence ou não ao conjunto C . O Princípio do Critério para Definições consiste na seguinte regra: *apenas definições criteriologicamente satisfatórias são aceitáveis*.

Em SCTFS, a passagem em que Tarski brevemente se reporta à noção de critério de verdade, está dissociada de qualquer discussão de natureza metamatemática, e se encontra num contexto bem específico, em que ele examina a possibilidade de a concepção semântica da verdade estar imbuída de elementos metafísicos¹⁰⁶. Para alguns, mesmo que o vocabulário da linguagem de uma dada teoria esteja isento de termos primitivos metafísicos, estes podem ser introduzidos por meio de definições que não sejam criteriologicamente satisfatórias; como a definição de verdade para a linguagem de uma teoria não é, em geral, criteriologicamente satisfatória, a concepção semântica da verdade possui um certo caráter metafísico. Em outras palavras, a concepção semântica da verdade possui elementos metafísicos, porque relaxa o PCD. É precisamente a esse tipo de crítica que Tarski está a se referir no excerto abaixo:

There are, however, some who believe that, even if no metaphysical terms occur among the primitive terms of a language, they may be introduced by definitions; namely, by those definitions which fail to provide us with general criteria for deciding whether an object falls under the defined concept. It is argued that term 'true' is of this kind, since no universal criterion of truth follows immediately from the definition of this term, and since it is generally believed (and in a certain sense can even be proved) that such a criterion will never be found¹⁰⁷.

Em TP, já não está em pauta (pelo menos explicitamente) o comprometimento da concepção semântica da verdade com elementos metafísicos, mas a noção de critério de verdade novamente aparece discutida em conexão com o PCD. De início, Tarski observa que a definição do conceito de verdade não fornece um critério de decisão para a verdade de sentenças particularmente consideradas:

¹⁰⁶ A passagem se encontra na segunda parte, item 19, 'Alleged metaphysical elements in semantics'.

¹⁰⁷ SCTFS, p. 63. É num contexto semelhante a esse que Popper discute a noção de critério de verdade numa certa altura de 'Comments on Tarski's Truth Definition', pp. 398-9.

Whatever may be achieved by constructing an adequate definition of truth for a scientific language, one fact seems to be certain: the definition does not carry with it a workable criterion for deciding whether particular sentences in this language are true or false (and indeed it is not designed at all for this purpose)¹⁰⁸.

E então, pouco depois, Tarski se refere aos defensores do PCD:

Some philosophers and methodologists of science are inclined to reject every definition that does not provide a criterion for deciding whether any given particular object falls under the notion defined or not. In the methodology of empirical sciences such a tendency is represented by the doctrine of operationalism; philosophers of mathematics who belong to the constructivist school seem to exhibit a similar tendency¹⁰⁹.

Seja em SCTFS, seja em TP, a posição de Tarski é clara com respeito ao PCD: é um princípio tão restritivo que, se o aceitarmos seriamente, seremos obrigados a renunciar a grande parte de nosso conhecimento – o que é totalmente correto, a começar, p.e., pela definição de ‘demonstrável’ numa axiomatização do cálculo de predicados de primeira ordem. Criticando o ‘operacionalismo’ em metodologia das ciências e o ‘construtivismo’ em filosofia da matemática, Tarski escreve, em TP:

In both cases, however, the people who hold this opinion appear to be in a small minority. A consistent attempt to carry out the program in practice (that is, to develop a science without using undesirable definitions) has hardly ever been made. It seems clear that under this program much of contemporary mathematics would disappear, and theoretical parts of physics, chemistry, biology, and other empirical sciences would be severally mutilated¹¹⁰.

E, em SCTFS, à objeção de que a concepção semântica da verdade viola o PCD, Tarski responde em termos bastante próximos aos que transcrevemos acima:

¹⁰⁸ TP, p. 69.

¹⁰⁹ *Idem*, p. 70.

¹¹⁰ *Ibidem*, p. 70.

[...] it should be noticed that the notion of truth does not differ in this respect from many notions in logic, mathematics, and theoretical parts of various empirical sciences, e.g., in theoretical physics¹¹¹.

Em CTFL, que passaremos a examinar agora, a noção de critério de verdade aparece num contexto diferente. Não mais se a discute em conexão com o PCD. O novo contexto é estritamente metamatemático.

A noção de critério de verdade em *The Concept of Truth*

A primeira vez em que Tarski menciona em CTFL a noção de critério de verdade, ocorre quando ele observa que a definição (semântica) do conceito de verdade na linguagem do cálculo de classes (definição 23 do § 3) não fornece um critério geral de verdade:

Def. 23 alone gives no general criterion for the truth of a sentence¹¹².

Mas já observamos que definições nem sempre vêm acompanhadas de critérios de decisão, e isso não constitui um defeito. De fato, Tarski escreve:

At least when it is regarded from the methodological viewpoint this is not a defect of the definition in question; in this respect it does not differ at all from the greater part of the definitions which occur in the deductive sciences¹¹³.

De um modo geral, o fato de a definição semântica do conceito de verdade não fornecer um método efetivo de decisão, está em parte vinculado à natureza infinitária do próprio método de Tarski de definição de verdade. Lembremos que o conceito de verdade é definido em termos do conceito de satisfação por uma seqüência denumerável de objetos: uma sentença é verdadeira, se é satisfeita por todas as seqüências. Ora, se o domínio de interpretação de uma linguagem contém pelo menos dois objetos distintos, então existe uma quantidade não-denumerável de seqüências de objetos construídas sobre tal domínio (precisamente, $2^{\aleph_0}$). Não há como testar efetivamente, segundo tal método, se uma dada sentença é verdadeira ou não, porque é óbvio que não podemos percorrer num número finito de passos uma quantidade não-denumerável de seqüências de objetos

¹¹¹ SCTFS, p. 63.

¹¹² CTFL, p. 197.

¹¹³ *Idem*, p. 197 (nota).

e decidir se todas elas satisfazem essa sentença (além disso, mesmo num caso particular, não podemos decidir se uma dada seqüência satisfaz ou não uma certa sentença, se a seqüência não for algoritmicamente gerada). Dissemos acima que esse problema se deve ‘em parte’ à natureza infinitária da definição semântica de verdade. Isso porque mesmo nos casos em que a definição semântica de verdade pode ser substituída por uma definição estrutural – portanto, de caráter finitário –, como no caso da definição de verdade lógica (em primeira ordem), não há garantia alguma de que possamos decidir efetivamente quais sentenças são verdadeiras, e quais são falsas¹¹⁴. A definibilidade estrutural da verdade garante apenas a enumerabilidade recursiva das verdades. A existência de um critério de verdade, formalmente falando, depende única e exclusivamente da recursividade do conjunto das verdades, como afirma Tarski na nota mais esclarecedora de CTFL sobre o problema:

*[...] We say that there is a general structural criterion of truth when, and only when, the set of all true sentences is (general) recursive*¹¹⁵.

Essa nota foi acrescida à tradução de Woodger, de 1956. À época da primeira publicação (em polonês) de CTFL, em 1933, Tarski ainda não podia delinear formal e precisamente a noção de critério de verdade, simplesmente porque as bases formais da teoria da recursão, assim como suas primeiras aplicações metamatemáticas, aguardariam mais um pouco para serem definitivamente estabelecidas (embora Gödel e Herbrand já tivessem dado os primeiros passos nessa direção). Mais exatamente, aguardariam até a publicação dos artigos de Church e Turing sobre o assunto, ambos de 1936¹¹⁶. Antes disso, a noção de critério de verdade ainda era uma noção puramente intuitiva – a de um ‘teste mecânico’ da verdade de sentenças. Depois de enunciada a Tese de Church-Turing, tornou-se possível defini-la em termos de alguma versão formal de ‘calculabilidade efetiva’ – como a versão referida por Tarski, a de recursividade.

À parte a recursividade do conjunto das verdades, é interessante notar que Tarski marca a palavra ‘critério’ com os dois qualificativos ‘geral’ e ‘estrutural’. O primeiro é auto-sugestivo: o critério de verdade diz-se ‘geral’, quando, dada qualquer sentença ϕ da linguagem L de uma teoria T , podemos decidir efetivamente se ϕ é verdadeira ou falsa com respeito à teoria T . Agora, o qualificativo ‘estrutural’ conduz-nos ao que Tarski chama de ‘propriedades estruturais’. A noção de ‘propriedades estruturais’ figura entre aquelas noções que Corcoran, na introdução à edição revista de LSM, menciona como

¹¹⁴ Cf. CTFL, p. 238 (nota).

¹¹⁵ *Idem*, p. 254 (nota 1, *in fine*).

¹¹⁶ Church, ‘An unsolvable problem of elementary number theory’ (1936); Turing, ‘On computable numbers, with an application to the Entscheidungsproblem’ (1936).

passíveis de melhor delineamento formal¹¹⁷. Seu conteúdo, entretanto, é bastante claro, e Tarski o apresenta informalmente com as seguintes palavras:

*[Structural properties]: properties concerning the form and order of succession of the individual parts of the expression*¹¹⁸.

Pois bem, o critério de verdade diz-se ‘estrutural’, se, dada uma sentença ϕ qualquer da linguagem L de uma teoria T , podemos decidir efetivamente, *com base nas propriedades estruturais de ϕ* , se ϕ pertence ou não ao conjunto das sentenças verdadeiras de L com respeito à teoria T .

Vejamos um caso simples para ilustrar. Consideremos o cálculo proposicional clássico. Há alguns métodos bastante conhecidos, por meio dos quais podemos decidir se uma dada forma proposicional é ou não uma tautologia – p.e., o método das tabelas de verdade e o método dos tablôs semânticos. Todos eles exploram a estrutura sintática da forma proposicional (i.e., suas propriedades estruturais), ao lado do assinalamento de valores às formas componentes. Mas há um outro método de decisão, em que se prescinde totalmente do assinalamento de valores, e em que apenas as propriedades estruturais são consideradas. Trata-se do método da forma normal conjuntiva: para decidir se uma dada forma proposicional é ou não uma tautologia, reduzimo-la à forma normal conjuntiva e verificamos se cada um dos conjuntos é da forma ‘ $\phi \vee \neg\phi$ ’ – caso em que a forma é tautológica.

Em 1936, como é bem conhecido, Church apresentou a solução negativa geral para o problema da decisão da lógica de primeira ordem (i.e., do conjunto de todas as verdades lógicas elementares), e da aritmética elementar dos números naturais (i.e., do conjunto de todas as verdades aritméticas elementares no modelo *standard* $\langle \mathbb{N}; +, \cdot, s, 0 \rangle$). Quando CTFL foi publicado, contudo, nenhum resultado negativo geral para o problema da decisão de uma teoria específica havia sido ainda alcançado. Mas a solução positiva geral para o cálculo proposicional clássico era conhecida desde a monografia de Post (1921)¹¹⁹, e algumas soluções positivas gerais já tinham sido dadas para algumas teorias matemáticas específicas. Além disso, bem se sabe, empreendeu-se uma série de soluções positivas parciais para o problema da decisão da lógica de primeira ordem. Em todas elas,

¹¹⁷ LSM, xxvi, nota 16.

¹¹⁸ CTFL, p. 163. Com respeito ao conceito de propriedades estruturais, podemos ser levados a pensar erroneamente da seguinte maneira: sendo uma expressão um objeto puramente sintático, *qualquer* propriedade que se lhe aplique deve ser forçosamente estrutural. Mas isso é falso. A propriedade de ser verdadeira (satisfeita por todas as seqüências enumeráveis de indivíduos de um dado domínio) aplica-se a sentenças, que são objetos puramente sintáticos, mas não se trata, evidentemente, de uma propriedade estrutural.

¹¹⁹ Post, ‘Introduction to a general theory of elementary propositions’ (1921). Wittgenstein, independentemente, também chegou a esse mesmo resultado no ‘Tractatus logico-philosophicus’ (1921).

apela-se para uma forma normal determinada. Mostra-se que o valor de cada sentença redutível a um dado padrão sintático é mecanicamente testável, e que, por conseguinte, o conjunto das sentenças assim normalizáveis é um conjunto decidível.

Por conveniência, havemos por bem considerar em separado, a seguir, o problema da existência de critério de verdade lógica, e o problema da existência de critério de verdade para teorias matemáticas específicas.

(i) Critério geral e critérios parciais de verdade lógica (em primeira ordem)

Em virtude do teorema de completude de Gödel, o problema da existência de um critério geral estrutural de verdade lógica (em primeira ordem), confunde-se com o *Entscheidungsproblem* da teoria elementar da quantificação. O problema (que, como acabamos de apontar, foi demonstrado recursivamente insolúvel por Church) ainda estava em aberto quando da primeira publicação de CTFL. Mas, em nota à tradução posterior de Woodger, Tarski assinala a inexistência de um tal critério para o conceito de ‘função sentencial universalmente válida’ no *engere Funktionenkalkül* de Hilbert e Ackermann¹²⁰, i.e., para o conceito de verdade lógica em primeira ordem¹²¹.

Numa das passagens em que Tarski discute o conceito relativo de verdade (verdade num domínio específico de indivíduos), ele alude a alguns casos em que podemos decidir efetivamente se determinadas sentenças, com certas propriedades estruturais, são ou não verdades lógicas.

[...] *In recent years numerous results have been obtained which enable us to infer from the correctness of certain sentences in special individual domains or from their structural properties their correctness in every domain and thus their truth*¹²².

Tarski está a se referir, entre outros, a alguns resultados obtidos por Bernays, Schönfinkel e Ackermann em 1928 – os dois primeiros, conjuntamente, em ‘Zum Entscheidungsproblem der mathematischen Logik’, e o último em ‘Über die Erfüllbarkeit gewisser Zählusdrücke’. Com efeito, em nota acrescida à tradução de Woodger, ele esclarece:

Certain results of Bernays, Schönfinkel, and Ackermann [...] allow us to correlate a particular natural number k with sentences of a special

¹²⁰ Hilbert & Ackermann, *Grundzüge der theoretischen Logik* (1928).

¹²¹ CTFL, p. 254 (nota).

¹²² *Idem*, p. 241.

structure in such a way that from the correctness of a given sentence in the domain with k elements (thus – as we already know – from purely structural properties of the sentence) its correctness in every domain follows¹²³.

Dos resultados de Bernays e Schönfinkel, constam, por exemplo, os três seguintes:

(a) Se ϕ é uma sentença que pode ser reduzida à forma normal prenex de tipo ' $\forall x_1 \dots \forall x_m M(x_1, \dots, x_m)$ ' e ϕ é verdadeira em algum domínio com m indivíduos, então ϕ é uma verdade lógica.

(b) Se ϕ é uma sentença que pode ser reduzida à forma normal prenex de tipo ' $\exists x_1 \dots \exists x_m M(x_1, \dots, x_m)$ ' e ϕ é verdadeira em algum domínio com *um único* indivíduo, então ϕ é uma verdade lógica.

(c) Se ϕ é uma sentença que pode ser reduzida à forma normal prenex de tipo ' $\forall x_1 \dots \forall x_m \exists y_1 \dots \exists y_n M(x_1, \dots, x_m, y_1, \dots, y_n)$ ' e ϕ é verdadeira em algum domínio com m indivíduos, então ϕ é uma verdade lógica.

Dos resultados obtidos por Ackermann, consta o referente às sentenças de primeira ordem que podem ser reduzidas à forma normal prenex, cujo prefixo consiste de apenas um quantificador existencial¹²⁴. Em todos os casos citados, pode-se testar mecanicamente se a sentença em questão é ou não uma verdade lógica.

(ii) Critérios de verdade para teorias matemáticas específicas

Em CTFL, Tarski se refere a duas teorias que possuem critério geral estrutural de verdade: o cálculo de classes¹²⁵ e o cálculo proposicional com quantificadores¹²⁶ (que Tarski chama de 'cálculo sentencial estendido'). Em relação a este último, ele apenas adianta, sem pormenores, que um critério geral de verdade pode ser modelado a partir do método das matrizes (tabelas de verdade)¹²⁷. De qualquer modo, este caso não nos interessa aqui, já que não se trata de uma teoria matemática específica. Em relação ao cálculo de classes, o critério geral de verdade é obtido por meio da aplicação do método de eliminação dos quantificadores, a respeito do qual já fornecemos algumas informações no primeiro capítulo. No final do § 3, Tarski indica que existe um critério geral estrutural de verdade para a linguagem do cálculo de classes. Ele não o exhibe, mas faz a remissão a

¹²³ *Ibidem*, p. 242 (nota, *in fine*).

¹²⁴ Tanto os resultados de Bernays e Schönfinkel, como os de Ackermann, mencionados por Tarski, são citados por Kneale & Kneale, *O Desenvolvimento da Lógica*, pp. 734-6, de onde os extraímos.

¹²⁵ CTFL, pp. 208-9 e 238-9.

¹²⁶ *Idem*, pp. 222 e 238-9.

¹²⁷ Cf. TARSKI & LUKASIEWICZ. 'Investigations into the sentential calculus', in LSM, pp. 54-9.

um lema atribuído essencialmente ao já citado artigo de 1919 de Skolem¹²⁸, em que se mostra a decidibilidade, por meio do método de eliminação dos quantificadores, da teoria das álgebras de conjuntos¹²⁹. Em ‘Foundations of the Calculus of Systems’ (FCS, 1935-6), Tarski chama o cálculo de classes de ‘álgebra atomística da lógica’, e remete ao mesmo artigo de Skolem e a CTFL, para o exame de suas principais propriedades metateóricas¹³⁰. E isso é tudo. Agora, é claro, todas as teorias axiomáticas estudadas durante os seminários de Varsóvia (geometria elementar, álgebra elementar dos números reais, teoria das ordens lineares discretas e densas etc.), às quais o método de eliminação dos quantificadores havia sido aplicado com êxito, poderiam ter sido mencionadas por Tarski já em CTFL, uma vez que ele dispunha então de critérios de verdade para cada uma delas.

A noção de critério de verdade numa classe de estruturas

No final do Capítulo 1, assinalamos que o que Tarski definiu em CTFL foi o conceito de verdade-numa-linguagem, e não o conceito de verdade-numa-estrutura, e vimos também que este último conceito só foi definido por Tarski duas décadas depois da publicação de sua famosa monografia. Dissemos que, embora não seja fácil identificar as razões para esse fato, é relativamente simples passar do ambiente original *simpliciter* para o ambiente padrão da teoria de modelos. Reservamos esta seção para examinar como transcrever a noção original de critério de verdade num contexto modelo-teorético atualizado¹³¹. Antes, contudo, façamos uma pequena digressão, que servirá de justificativa para o enfoque que adotaremos.

Sabemos que, em primeira ordem, as relações ‘ $\vdash$ ’ e ‘ $\models$ ’ (respectivamente, consequência sintática e consequência semântica) são relações co-extensivas. Então, a noção de consequência lógica em primeira ordem pode ser tratada indistintamente em termos sintáticos ou semânticos. De fato, pela versão forte do teorema de completude de Gödel, sendo ‘ $\vdash$ ’ a relação de consequência sintática numa axiomatização adequada do cálculo de predicados (o sistema de Hilbert-Ackermann, por exemplo), temos:

$$\Sigma \models \phi \Rightarrow \Sigma \vdash \phi$$

¹²⁸ Cf. capítulo 1.

¹²⁹ O lema a que Tarski se refere é o Lema K, usado na demonstração do Teorema 28 do § 3 de CTFL. Por esse teorema, é estabelecida a completude de uma certa extensão do cálculo de classes.

¹³⁰ FCS, p. 378 (nota). Sobre o cálculo de classes como uma álgebra booleana com átomos, cf. TARSKI, ‘On the foundations of boolean algebra’, LSM, p. 341.

¹³¹ Esta é uma seção relativamente independente das demais seções deste capítulo. O leitor, se quiser, pode passar da seção anterior para a próxima seção, sem risco de solução de continuidade na leitura.

Sendo a axiomatização correta (i.e., se vale $\Sigma \vdash \phi \Rightarrow \Sigma \models \phi$), é óbvio o seguinte:

$$\Sigma \models \phi \Leftrightarrow \Sigma \vdash \phi$$

Consideremos o caso em que uma teoria elementar T é axiomática. Dado um conjunto Σ (finito ou infinito) de sentenças, escolhido como base axiomática de T , é irrelevante se caracterizamos T como o conjunto das conseqüências semânticas de Σ , ou como o conjunto das conseqüências sintáticas de Σ , i.e.,

$$\text{Se } T = \{\phi : \Sigma \models \phi\} \text{ e } T' = \{\phi : \Sigma \vdash \phi\}, \text{ então } T = T'.$$

Há apenas uma diferença de enfoque entre as duas caracterizações acima: na primeira, a ênfase recai sobre a classe dos modelos da teoria estudada, ao passo que, na segunda, a ênfase recai sobre a estrutura sintático-dedutiva de T .

De qualquer modo, o ponto de partida em ambas as caracterizações é o mesmo: parte-se de um conjunto fixo de sentenças, e consideram-se todas as suas conseqüências lógicas. Por falta de expressão melhor, podemos chamar tal forma de caracterizar uma teoria de *preponderantemente lingüística*. Mas há uma outra maneira de se caracterizar uma teoria. Em vez de partirmos simplesmente de sentenças, partimos de uma dada estrutura relacional, ou de uma dada classe de estruturas relacionais, e então consideramos o conjunto de todas as sentenças que são verdadeiras a seu respeito. Se, por exemplo, A é uma certa estrutura, e o símbolo ' $\models$ ' agora denota a relação de satisfatibilidade, a teoria T de A é caracterizada assim:

$$T = \{\phi : A \models \phi\}$$

O lado saliente dessa espécie de caracterização de uma teoria, é o lado ontológico – em contraposição à espécie anterior, em que o lado saliente é o lingüístico. À falta de designação melhor, chamemos tal caracterização de *preponderantemente ontológica*. Além disso, pelo menos dois outros aspectos marcam o contraste entre as duas caracterizações: (i) uma teoria caracterizada segundo o enfoque ontológico é sempre completa, o que não é verdadeiro com relação a uma teoria caracterizada segundo o enfoque lingüístico; (ii) existem teorias que só podem ser caracterizadas segundo o

enfoque ontológico, como é o caso da aritmética elementar dos números naturais (e, em geral, qualquer outra teoria não-axiomatizável).

O enfoque que escolhemos aqui para reproduzir, *mutatis mutandis*, a noção Tarskiana de critério de verdade, é o preponderantemente ontológico, por uma razão bastante simples. Uma vez que o que está em pauta é a recursividade ou não-recursividade do conjunto das *verdades* da linguagem de uma dada teoria, só podemos caracterizar uma teoria, no caso presente, como sendo o conjunto de todas as verdades a respeito de uma dada estrutura (ou classe de estruturas). *O problema da decisão para teorias caracterizadas segundo o enfoque ontológico, é sempre o problema da existência de critério de verdade.*

Pois bem, formalizemos então a noção de critério de verdade num contexto modelo-teorético. Assumiremos familiaridade com as noções básicas da teoria de modelos. Seja $\text{Est}(\tau)$ uma classe de estruturas de tipo de similaridade τ . Seja K uma subclasse de $\text{Est}(\tau)$. Seja L uma linguagem de primeira ordem adequada ao tipo τ . Seja L_s o conjunto de todas as sentenças de L . Consideremos as seguintes definições:

Definição 1. Por “teoria de K na linguagem L ”, denotada por $\mathfrak{I}_L(K)$, entendemos o conjunto de todas as sentenças de L que são verdadeiras em qualquer estrutura pertencente a K . Em símbolos:

$$\mathfrak{I}_L(K) =_{\text{def.}} \{ \phi \in L_s \mid \forall A \in K : A \models \phi \}$$

No caso particular em que K consiste de apenas uma estrutura A , temos:

$$\mathfrak{I}_L(\{A\}) =_{\text{def.}} \{ \phi \in L_s : A \models \phi \}$$

Admitamos que toda a sintaxe de L tenha sido aritmetizada, i.e., admitamos que existe uma injeção $c : E \longrightarrow \mathbb{N}$ (em que E denota o conjunto das expressões¹³² de L , e $\mathbb{N}$, o conjunto dos números naturais), tal que (i) dada uma expressão e qualquer de L , podemos encontrar algoritmicamente seu número código $c(e)$, e, (ii) existe um procedimento efetivo tal que, dado um número natural n qualquer, podemos decidir se n está ou não na imagem de c , e, em caso afirmativo, podemos decodificá-lo, i.e., encontrar a expressão e tal que $c(e) = n$. Sendo assim, associada univocamente a cada relação R de expressões de L , existe uma relação R^* de códigos das expressões de R . Em particular,

¹³² A rigor, por “expressões de L ”, estamos entendendo, de um modo geral, qualquer símbolo isolado do alfabeto de L , seqüências finitas de tais símbolos, e seqüências finitas de tais seqüências de símbolos.

associado ao conjunto $\mathfrak{I}_L(K)$ das sentenças de L verdadeiras em qualquer estrutura de K , existe o conjunto $\mathfrak{I}_L(K)^*$ de seus respectivos códigos, e associado ao conjunto $\mathfrak{I}_L(\{A\})$ das sentenças de L verdadeiras em A , existe o conjunto $\mathfrak{I}_L(\{A\})^*$ de seus respectivos códigos. A aritmetização da sintaxe de L permite-nos examinar alguns problemas (em especial, de decidibilidade) à luz da teoria das funções recursivas.

Definição 2. O problema da decisão para $\mathfrak{I}_L(K)^*$ diz-se recursivamente solúvel, se, e somente se, $\mathfrak{I}_L(K)^*$ é recursivo.

Em particular, o problema da decisão para $\mathfrak{I}_L(\{A\})^*$ diz-se recursivamente solúvel, se, e somente se, $\mathfrak{I}_L(\{A\})^*$ é recursivo.

Definição 3. Dizemos que existe um critério geral de verdade para $\mathfrak{I}_L(K)$, se, e somente se, o problema da decisão para $\mathfrak{I}_L(K)^*$ é recursivamente solúvel.

Em particular, dizemos que existe um critério geral de verdade para $\mathfrak{I}_L(\{A\})$, se, e somente se, o problema da decisão para $\mathfrak{I}_L(\{A\})^*$ é recursivamente solúvel.

Das definições 2 e 3, segue imediatamente que existe um critério geral de verdade para $\mathfrak{I}_L(K)$, quando, e somente quando, $\mathfrak{I}_L(K)^*$ é recursivo. Em particular, das definições 2 e 3, segue imediatamente que existe um critério geral de verdade para $\mathfrak{I}_L(\{A\})$, quando, e somente quando, $\mathfrak{I}_L(\{A\})^*$ é recursivo. Este último caso pode ser visto como a versão imediata, formulada em termos modelo-teóricos, da noção original tarskiana de “critério geral estrutural de verdade”, como extraída da nota que reproduzimos alguns parágrafos atrás¹³³.

Vejamos agora dois exemplos (familiares) de teorias que possuem critério de verdade, no sentido explicado acima, e um exemplo (também familiar) de teoria que não possui tal critério. Precisamos de algumas definições prévias. Seja L uma linguagem de primeira ordem, cujo alfabeto consiste dos seguintes símbolos: $\forall, \exists, ', \neg, \rightarrow, =, \leq, +, \bullet, s, 0, 1, (,)$. O símbolo “ s ” denota a função sucessor. Cada variável individual é formada pela concatenação da letra “ x ” com um número finito de justaposições do símbolo “ $'$ ”: x', x'', x''' etc. A definição dos conceitos de termo, fórmula etc., segue o usual.

¹³³ Uma vez que um conjunto numérico qualquer diz-se recursivo somente se sua função característica é recursiva, podemos considerar as expressões “critério” e “função característica recursiva” como co-extensivas.

Definição 4. Por *Aritmética*, denotada por “ $\mathfrak{G}_L(\{N\})$ ”, entendemos o conjunto de todas as sentenças de L verdadeiras na estrutura $N = \langle \mathbb{N}; \leq, +, \bullet, s, 0, 1 \rangle$, em que N é o modelo *standard* da aritmética de primeira ordem. Sendo L_s o conjunto de todas as sentenças de L , temos, em símbolos:

$$\mathfrak{G}_L(\{N\}) =_{\text{def.}} \{ \phi \in L_s : N \models \phi \}$$

Definição 5. Por *Aritmética de Presburger*, denotada por “ $\mathfrak{G}_{\text{Pb}}(\{N\})$ ”, entendemos o conjunto de todas as sentenças de L em que não ocorre o operador “ $\bullet$ ”, que são verdadeiras em N . Sendo $L_s(\text{Pb})$ o conjunto de todas as sentenças de L em que não ocorre o operador “ $\bullet$ ”, temos, em símbolos:

$$\mathfrak{G}_{\text{Pb}}(\{N\}) =_{\text{def.}} \{ \phi \in L_s(\text{Pb}) : N \models \phi \}$$

Definição 6. Por *Aritmética de Skolem*, denotada por “ $\mathfrak{G}_{\text{Sk}}(\{N\})$ ” entendemos o conjunto de todas as sentenças de L em que não ocorrem os operadores “ $+$ ” e “ s ”, que são verdadeiras em N . Sendo $L_s(\text{Sk})$ o conjunto de todas as sentenças de L em que não ocorrem os operadores “ $+$ ” e “ s ”, temos, em símbolos:

$$\mathfrak{G}_{\text{Sk}}(\{N\}) =_{\text{def.}} \{ \phi \in L_s(\text{Sk}) : N \models \phi \}$$

Proposição 1 (*Presburger*, 1929). O problema da decisão para $\mathfrak{G}_{\text{Pb}}(\{N\})^*$ é recursivamente solúvel.

Corolário. Existe um critério geral de verdade para $\mathfrak{G}_{\text{Pb}}(\{N\})$.

Proposição 2 (*Skolem*, 1930). O problema da decisão para $\mathfrak{G}_{\text{Sk}}(\{N\})^*$ é recursivamente solúvel.

Corolário. Existe um critério geral de verdade para $\mathfrak{G}_{\text{Sk}}(\{N\})$.

Proposição 3 (*Church*, 1936). O problema da decisão para $\mathfrak{G}_L(\{N\})^*$ não é recursivamente solúvel.

Corolário. Não existe critério geral de verdade para $\mathfrak{G}_L(\{N\})$.

Observação. Como $\mathfrak{P}_b(\{N\})^* \subset \mathfrak{P}_L(\{N\})^*$ e $\mathfrak{P}_{Sk}(\{N\})^* \subset \mathfrak{P}_L(\{N\})^*$, podemos dizer, com base nas Proposições 1 e 2, que existem critérios “parciais” de verdade para $\mathfrak{P}_L(\{N\})$.

Mais algumas indicações histórico-conceituais

Para finalizar, gostaríamos de traçar um quadro histórico-conceitual para futuro aprofundamento no assunto deste capítulo. Um estudo técnico minucioso das investigações de Tarski sobre o problema da existência de critério de verdade para uma teoria, deve levar em consideração suas duas principais contribuições à teoria da decidibilidade¹³⁴: (i) a já citada aplicação do método de eliminação dos quantificadores para provas de decidibilidade de teorias, e (ii) o chamado ‘método de interpretação’ para provas de *ind*decidibilidade de teorias. Note bem: pelo primeiro método, obtém-se uma solução positiva para o problema da decisão de uma teoria, ao passo que, pelo segundo, obtém-se uma solução negativa. Já dissemos que o método de eliminação dos quantificadores ocupava o centro das pesquisas nos seminários ministrados por Tarski na Universidade de Varsóvia, durante os anos 1927-9. Alguns dos principais resultados obtidos por Tarski nesses anos, foram reunidos numa monografia que seria publicada em 1940, *The Completeness of Elementary Algebra and Geometry*. Entretanto, por conta da II Guerra Mundial, a publicação foi cancelada. Oito anos depois, com o apoio de McKinsey, Tarski reescreveu a monografia e a publicou sob o título *A Decision Method for Elementary Algebra and Geometry*. A mudança de ‘completeness’ para ‘decision method’ nos títulos, reflete, não uma mudança de conteúdo, mas apenas uma mudança de ênfase nas pesquisas de Tarski. De fato, como já assinalamos, o método de eliminação dos quantificadores aplica-se à demonstração de completude de uma teoria, mas colateralmente fornece um procedimento efetivo de decisão para essa mesma teoria. Antes da II Guerra Mundial, Tarski dedicou-se menos ao problema da decisão do que ao problema da completude de teorias¹³⁵. A inversão deu-se no final da década de 40, com a publicação da monografia acima mencionada, e coroou-se com a publicação de *Undecidable Theories*, em 1953, de que falaremos a seguir.

Undecidable Theories divide-se em três artigos, o primeiro dos quais, ‘A general method in proofs of undecidability’, contém a outra grande contribuição de Tarski à teoria da decidibilidade, o método da interpretação. Os dois últimos artigos mostram

¹³⁴ O que chamamos aqui de ‘teoria da decidibilidade’ é simplesmente a teoria da recursão aplicada à lógica.

¹³⁵ Para um exame dessa mudança de ênfase, e também para um exame histórico-conceitual do método de eliminação dos quantificadores, cf. DONER, J. & HODGES, W. ‘Alfred Tarski’s and decidable theories’, JSL, v. 53, n.º 1, 1988, pp. 20-35 (em especial, pp. 29-32).

aplicações desse método. Vejamos, muito superficialmente, em que ele consiste. De início, Tarski observa¹³⁶ que, antes da elaboração do método da interpretação, dois métodos básicos estavam disponíveis para provas de indecidibilidade de teorias: (a) o ‘método direto’, em que a técnica central envolvida é a diagonalização, e (b) o ‘método indireto’, em que a técnica central envolvida é a redução do problema da decisão de uma teoria T_1 ao problema da decisão de uma teoria T_2 , já demonstrada a insolubilidade deste último. No método indireto, a redução assume uma das duas formas: (b₁) dada uma teoria indecidível T_2 , mostra-se que a teoria T_1 pode ser obtida de T_2 removendo-se um número finito de axiomas desta última, e mantendo-se todas as constantes de sua linguagem; (b₂) dada uma teoria essencialmente indecidível T_2 (i.e., uma teoria desprovida de extensão consistente decidível), mostra-se que T_2 é interpretável na teoria T_1 . O método da interpretação de Tarski, na verdade, é uma versão modificada e estendida do método indireto, e, como ele mesmo observa, combina elementos de (b₁) e (b₂). Tarski descreve-o da seguinte maneira:

[...] in order to establish the undecidability of a theory T_1 , it suffices to show that some essentially undecidable theory T_2 can be interpreted, not necessarily in T_1 , but (what is much easier) in some consistent extension of T_1 – provided only that T_2 is based upon a finite axiom system¹³⁷.

O segundo artigo de *Undecidable Theories*, ‘Undecidability and essential undecidability in arithmetic’, escrito em parceria com Robinson e Mostowski, fornece uma série de resultados importantes obtidos por meio da aplicação do método de interpretação a subteorias da aritmética elementar dos números naturais. No último artigo, ‘Undecidability of the elementary theory of groups’, como se depreende do próprio título, Tarski emprega o método de interpretação para solucionar negativamente o problema da decisão da teoria dos grupos em primeira ordem¹³⁸. •

¹³⁶ UT, pp. 3-4.

¹³⁷ UT, p. 5.

¹³⁸ Para um exame mais detalhado do método da interpretação, cf. McNULTY, G. F. ‘Alfred Tarski and undecidable theories’, JSL, v. 51, n.º 4, 1986, pp. 890-8 (especialmente, pp. 890-2).

Considerações finais

Retrospecto

Chegamos ao fim desta dissertação. Nossa tendência, como deve ter ficado claro ao longo dos cinco capítulos, foi enfatizar a dimensão lógico-matemática, mais do que a dimensão filosófica, das investigações semânticas desenvolvidas por Tarski. Isso não foi fruto de uma escolha arbitrária ou resultado de um fato acidental. Nosso propósito, como anunciamos já na Introdução, consistiu em oferecer uma apresentação geral dessas investigações. E o curso dominante de tais investigações, nós constatamos, é preponderantemente lógico-matemático. Mas isso não significa, é claro, que em certos momentos não intervenham questões filosóficas, nem que não haja também alguma motivação extra-matemática na origem das pesquisas semânticas de Tarski. O fato é que a dimensão filosófica simplesmente não é desenvolvida num grau de elaboração comparável aos aspectos mais técnicos dessas pesquisas. Pois bem, vejamos sumariamente o que foi visto, a fim de termos uma visão sinóptica da dissertação.

Duas foram as principais razões que levaram Tarski a definir o conceito de verdade e a estabelecer os fundamentos da semântica de linguagens formalizadas. Em primeiro lugar, havia o uso informal de noções semânticas em importantes resultados de natureza metamatemática. Entre tais resultados, figuravam, em destaque, os teoremas de Löwenheim-Skolem, o teorema de completude de Gödel, e a classificação Tarskiana das extensões completas e dos modelos elementarmente equivalentes de uma teoria dedutiva. Em segundo lugar, havia também a ameaça à consistência da semântica, oferecida por paradoxos famosos como o de Eubúlides-Epimênides (relacionado ao conceito de verdade) e os de Richard e Berry (relacionados ao conceito de definibilidade). O problema das antinomias semânticas parecia muito mais grave do que o uso indiscriminadamente informal de noções como definibilidade e verdade.

Foi nesse contexto que surgiu, assim, o programa Tarskiano de formalização da semântica, cuja tarefa principal consistia em formalizar o conceito de verdade. Há quem defenda, como Hartry Field, que o contexto era ainda mais geral, uma vez que Tarski subscrevia, à época, o programa fisicalista de redução de todos os conceitos de natureza semântica a conceitos físicos ou lógico-matemáticos¹³⁹. Já vimos a relação desse

¹³⁹ Field, Hartry. 'Tarski's Theory of Truth'. *Journal of Philosophy*, 69 (1972), pp. 347-75. O famoso argumento de Field é que Tarski não conseguiu realizar cabalmente esse programa. Segundo ele, a definição de Tarski pressupõe tacitamente uma noção não reduzida, o conceito de 'denotação primitiva', que assinala indivíduos a constantes individuais, relações a constantes relacionais, e grafos a constantes funcionais. Há problemas com o argumento de Field, mas não vamos discuti-los.

programa com as críticas de Tarski a uma teoria axiomática da verdade (e a uma axiomatização da semântica em geral). De qualquer maneira, do ponto de vista metodológico, a formalização deveria ser levada a cabo ou por meio de definições no interior de algum sistema fundacional matemático, como a teoria dos tipos de Russell-Whitehead ou a axiomática de Zermelo-Fraenkel da teoria dos conjuntos, ou por meio do estabelecimento de uma teoria axiomática da verdade, assumindo-se o conceito de verdade como primitivo semântico e fixando-se axiomáticamente suas principais propriedades intuitivamente aceitas (descitação, terceiro excluído, não-contradição etc.). Por uma série de razões técnicas e filosóficas, essa última via parecia insatisfatória para Tarski, que acabou por privilegiar o enfoque definicional de formalização da semântica.

Nesse enfoque, o problema central consistia em definir *satisfatoriamente* o conceito de verdade, ou, em outras palavras, construir uma definição *formalmente correta* e *materialmente adequada* desse conceito. O critério de correção formal concerne a todos os tipos de definição, não importa com que propósito específico a definição seja confeccionada. Trata-se simplesmente de obedecer às regras formais de introdução (e eliminação) de novos termos a partir dos termos primitivos (ou de termos previamente definidos), a fim de evitar o surgimento de contradições no interior da teoria. No entanto, se a definição é construída com fins explicativos de uma noção já utilizada em nível pré-teorético (como é o caso da definição do conceito de verdade), é necessário satisfazer, além do critério de correção formal, o critério de adequação material, pelo qual se exige que a definição preserve as principais propriedades intuitivamente atribuídas ao *definiendum*. Um dos grandes feitos de Tarski foi mostrar que a única propriedade intuitiva do conceito de verdade com que nos devemos preocupar, é a propriedade expressa no Esquema T, a propriedade descitatória. O restante das propriedades intuitivamente assinaladas à noção de verdade, segue facilmente de uma definição formalmente correta que preserve essa única propriedade.

A definição do conceito de verdade deve seguir a profilaxia recomendada por Tarski para a prevenção do surgimento de antinomias semânticas – em especial, do paradoxo de Eubúlides-Epimênides. Estabelecemos uma hierarquia $L_0, L_1, L_2, \dots$, de linguagens formais semanticamente abertas, tal que (i) dados dois níveis quaisquer m e n , se $m < n$, então L_m é uma parte própria de L_n (ou L_m pode ser traduzida em L_n), e (ii) dados dois níveis quaisquer m e n , se $m < n$, então L_n é essencialmente mais rica do que a linguagem L_m . A definição do conceito de verdade na linguagem de um determinado nível, só pode ser construída numa linguagem de nível superior pertencente a essa hierarquia, sob pena de inconsistência. Tomada essa precaução geral, ainda há alguns problemas de natureza técnica a serem resolvidos. O principal deles é a impossibilidade de se construir uma definição recursiva direta do conceito de verdade em linguagens com

quantificadores. A solução, se a linguagem-objeto não contém um número finito de sentenças e se a metalinguagem é finitária, é lançar mão do conceito de satisfação por uma seqüência (finita ou infinita) de objetos. Esse conceito é suficientemente geral e se aplica a todas as fórmulas da linguagem-objeto, abertas ou fechadas, de modo que é possível defini-lo por um procedimento recursivo direto. Quando aplicado a sentenças, isto é, a fórmulas fechadas, o conceito de satisfação conduz imediatamente ao conceito de verdade. A propriedade de uma fórmula ser satisfeita por todas as seqüências infinitas de objetos (ou pela seqüência finita vazia), captura exatamente a propriedade expressa no Esquema T, e o problema de uma definição satisfatória do conceito de verdade está, assim, resolvido.

Toda investigação de natureza metamatemática pressupõe um sistema fundacional matemático no qual ela é desenvolvida. Quando Tarski publicou a primeira versão de CTFL, em 1933, o sistema fundacional pressuposto era uma determinada extensão da teoria simples dos tipos, a chamada ‘teoria das categorias semânticas’ de Husserl-Lesniewski (TCS). No âmbito dessa teoria, Tarski estabeleceu uma tipologia das linguagens formalizadas, que precisava ser cuidadosamente observada na hora de aplicar-lhes os métodos definicionais desenvolvidos para fins semânticos. Segundo essa tipologia, as linguagens dividem-se em dois grandes grupos – as linguagens de ordem finita (de espécies 1, 2, 3) e as linguagens de ordem infinita (de espécie 4). Nas linguagens de ordem finita, Tarski mostrou que é possível definir satisfatoriamente o conceito de verdade – variando-se apenas a complexidade dos métodos definicionais, de acordo com a espécie de linguagem investigada. Nas linguagens de ordem infinita, no entanto, Tarski mostrou que é impossível construir uma definição do conceito de verdade (ou de qualquer outra noção semântica). A razão dessa impossibilidade reside no fato de as linguagens baseadas em TCS não conterem variáveis de ordem transfinita. A prova decisiva de tal impossibilidade foi dada por aquilo que mais tarde ficaria conhecido como o ‘teorema da indefinibilidade da verdade’. Uma das conseqüências metamatemáticas mais importantes desse teorema, é que ele implica o primeiro teorema de completude de Gödel. Outra conseqüência relevante é que, no âmbito de TCS, apenas a via axiomática é possível para a semântica de linguagens formalizadas de ordem infinita. Mas, por conta de uma série de problemas técnicos e filosóficos engendrados por uma teoria axiomática da verdade, Tarski resolveu abandonar TCS, e acrescentou um pós-escrito à tradução alemã de CTFL, que foi publicada em 1936. Nesse pós-escrito, Tarski relaxou TCS, mas ainda continuou com uma teoria dos tipos – só que, desta vez, com tipos de ordem transfinita. A definição padrão, modelo-teorética, em termos de estruturas relacionais, com a qual estamos hoje acostumados, só viria a ser dada mais tarde, quase vinte anos

depois, num artigo escrito em parceria com Robert Vaught, AERS, em 1957. Esse longo hiato entre CTFL e AERS tem dado azo a alguma discussão recente.

Na Introdução desta dissertação, dissemos que nos restringiríamos ao nó da bifurcação lógico-filosófica que surge do trabalho original de Tarski, com os ramos da semântica formal das linguagens naturais e da análise dos paradoxos semânticos. Pois bem, entendemos que o que foi dito nos parágrafos anteriores, e que foi visto mais detalhadamente no curso dos quatro primeiros capítulos, cobre o essencial quanto a esse nó. Contudo, ainda assim, há detalhes que às vezes passam despercebidos, mas que mereceriam um exame mais pormenorizado. O quinto e último capítulo desta dissertação abordou justamente um desses detalhes que parecem relevantes. De fato, nem sempre a distinção entre *critério* de verdade e *definição* de verdade (entre critério e definição, de uma maneira geral) foi estabelecida satisfatoriamente. Muitas vezes, ambas as noções são confundidas, e muitas vezes introduzem-se noções vagas e de natureza psicológica para caracterizar um critério de verdade – como a noção de evidência de uma proposição (uma herança certamente Cartesiana). Em CTFL, encontramos um tratamento, se não suficientemente geral, pelo menos preciso, rigoroso e anti-psicologista, da noção de critério de verdade. Tarski identifica aí a noção de critério de verdade com o conceito de decidibilidade do conjunto das verdades de uma teoria. A origem desse tratamento em Tarski remonta aos seminários de Varsóvia, em que estavam em foco diversas questões relacionadas ao método de eliminação dos quantificadores – e que resultaram na descoberta de critérios de verdades para diversas teorias, incluindo os casos famosos da geometria elementar e da álgebra elementar dos números reais. A noção de critério de verdade, como a noção mesma de verdade, é uma dessas noções que se encontram na interface da filosofia e da matemática – de um lado, interessa de perto à epistemologia (sob que condições podemos ter certeza de que estamos diante de uma verdade?), e, de outro, interessa de perto aos recônditos mais técnicos da teoria da recursão.

Prospecto

Ao longo de nossa pesquisa, foi-se reforçando cada vez mais a impressão de que ainda há muito terreno virgem a ser explorado nos escritos de Tarski sobre o conceito de verdade – e sobre a semântica, de uma maneira geral. A leitura feita por filósofos e lógicos sempre foi muito seletiva nos tópicos. Mas isso é assim mesmo, porque, se há conexões que precisam ser desenhadas, esse é o múnus da historiografia da lógica. Traçar o mapa dessas conexões, no entanto, não é tarefa simples. Há duas escalas envolvidas – uma escala “micro-historiográfica” e uma escala “macro-historiográfica” – que precisam ser delicadamente coordenadas. E, dado o estatuto eclético dos escritos semânticos de

Tarski, temos que estar também permanentemente atentos às duas dimensões aí envolvidas – a dimensão matemática e a dimensão filosófica – mesmo que o tom prevalente seja lógico-matemático. Num estudo em escala micro, a ênfase recai quase que exclusivamente sobre os escritos do próprio Tarski (seja um só, sejam vários). Em escala macro, levamos em consideração a interação das idéias de Tarski com outros autores (por exemplo, Tarski-Hilbert-Gödel ou Tarski-Lesniewski-Kotarbinski). Se privilegiarmos a dimensão filosófica, uma boa fonte de estudos é a tradição austro-polonesa, que envolve nomes como Bolzano, Brentano, Meinong, Husserl, Twardowski, Lesniewski, Kotarbinski, Lukasiewicz, Ajdukiewicz, Carnap e os positivistas do Círculo de Viena etc¹⁴⁰. Se privilegiamos a dimensão matemática, uma boa fonte de estudos é Russell e Whitehead (por conta da enorme influência exercida pelos *Principia Mathematica*), Schröder, Löwenheim, Skolem, Herbrand, Gödel, Hilbert e os membros da escola de Göttingen, Lesniewski, Lindenbaum, Lukasiewicz etc. Cada uma das dimensões possui sua própria constelação de temas. Do lado filosófico, por exemplo, são temas recorrentes, além da natureza da verdade, a questão do caráter absoluto ou relativo da verdade, a questão dos portadores primários da verdade (crenças? proposições? sentenças?), a questão do caráter temporal ou atemporal da verdade (eternidade, sempiternidade, efemeridade), a questão dos critérios de verdade etc. É fácil ver que esta dissertação se desenvolve em escala micro-historiográfica e privilegia a dimensão lógico-matemática.

Bem, qualquer que seja a combinação escolhida de escalas (micro/macro) e dimensões (matemática/filosófica), parece necessário, do ponto de vista metodológico, selecionar um aspecto a ser focalizado. Para dar uma idéia da riqueza do terreno a ser explorado, vejamos um estudo que se pode fazer em escala micro e de dimensão puramente matemática (portanto, ainda na linha desta dissertação). Uma abordagem que nos parece interessante é situar as investigações semânticas que Tarski desenvolve em CTFL, no âmbito mais geral da **teoria da definição** – afinal, o programa Tarskiano de formalização da semântica privilegia o enfoque *definicional*, muito mais do que o enfoque *axiomático*. Como se sabe, a teoria da definição desdobra-se em dois setores, quais sejam, a teoria da definibilidade *sintática* e a teoria da definibilidade *semântica*. Tarski deu contribuições importantes a ambas as áreas. Na coletânea *Logic, Semantics, Metamathematics*, por exemplo, há três artigos dedicados à teoria semântica da definibilidade: ‘On Definable Sets of Real Numbers’ (DSRN, 1931), ‘Logical Operations and Projective Sets’ (LOPS, 1931, escrito em parceria com Kuratowski) e ‘On the Limitations of the Means of Expression of Deductive Theories’ (LMEDT, 1934-5, escrito

¹⁴⁰ Aqui, um bom guia seria Wolenski, Jan. ‘Theories of Truth in Austrian and Polish Philosophy’.

em parceria com Lindenbaum). Há também um artigo voltado para a teoria sintática da definibilidade: ‘Some Methodological Investigations on the Definability of Concepts’ (SMIDC, 1934). Observem-se as datas desses artigos – elas revelam que os anos de 1931 a 1935 compreendem um período de intensa pesquisa no campo da teoria da definição. Justamente no meio desse período, em 1933, Tarski publica sua monografia sobre o conceito de verdade.

Não há dúvida de que há conexões entre as investigações que Tarski desenvolve nesses artigos, e as investigações específicas que ele desenvolve sobre o conceito de verdade em CTFL. Considere, por exemplo, DSRN, o artigo mais evidente no flanco da teoria semântica da definibilidade. Vimos que o método básico de definição da noção de verdade, baseado no conceito de satisfação por uma seqüência finita ou infinita de objetos, bem como a profilaxia contra os paradoxos semânticos, já estão delineados em DSRN – que, como dissemos, é o primeiro trabalho de Tarski em semântica.

Agora, considere o flanco da teoria sintática da definibilidade. Quais as possíveis afinidades entre SMIDC e CTFL? Bem, a motivação intuitiva básica da teoria sintática da definibilidade, é responder ao seguinte problema: sob que condições uma dada constante c é definível em termos do conjunto $\{c_1, \dots, c_n\}$ (ou $\{c_1, \dots, c_n, \dots\}$) das constantes primitivas da linguagem de uma certa teoria T ? Ora, no contexto de CTFL, temos um caso especial desse problema: sob que condições o conceito de verdade, relativo à linguagem de uma dada teoria, é definível em termos das noções primitivas da metalinguagem, isto é, da linguagem da meta-teoria? Claramente, a resposta a essa pergunta nos conduzirá ao conceito de riqueza essencial da metalinguagem em relação à linguagem-objeto, e esse conceito (riqueza essencial em geral) é abordado em SMIDC. Talvez não seja difícil rastrear outras afinidades. Por exemplo, ao discutir a completude de um sistema de termos primitivos, Tarski se reporta ao conceito de categoricidade. Numa pequena digressão, ele afirma que um sistema não-categorico de sentenças não determina precisamente o significado dos conceitos que nelas ocorrem. Ora, vimos que, em CTFL, uma das propriedades que Tarski reclama para que uma teoria axiomática da verdade seja considerada satisfatória, é justamente a categoricidade. Uma teoria não-categorica da verdade não determinaria de maneira não-ambígua a extensão do conceito de verdade.

Além da teoria da definição, podemos estender o mapeamento das conexões de CTFL com outros artigos, focalizando **outros aspectos**. Por exemplo, ‘Some Observations on the Concepts of ω -Consistency and ω -Completeness’ (SOC ω CC, 1933) também é um artigo que certamente interessa ao estudo de CTFL. Aí encontramos novamente a distinção entre níveis lingüísticos, e mais uma vez a teoria metalingüística da concatenação. Mas o principal interesse nesse artigo recai sobre a discussão da regra ω , que, como vimos, é de extrema relevância para uma teoria axiomática da verdade,

principalmente no âmbito de uma teoria dos tipos em que só há tipos de ordem finita. Outro exemplo: um estudo de ‘Foundations of the Calculus of Systems’ (FCS, 1935-6) parece valioso, se lembrarmos que é nesse artigo que Tarski registra seus resultados sobre a classificação das extensões completas de uma teoria, bem como de seus modelos elementarmente equivalentes – resultados esses que, como vimos, motivaram as investigações semânticas que culminaram em CTFL. É recomendável ler os dois ancestrais de FCS: ‘On Some Fundamental Concepts of Metamathematics’ (1930) e ‘Fundamental Concepts of the Methodology of the Deductive Sciences’ (1930). Ainda mais: seria desnecessário chamar a atenção para o cotejo que se pode fazer entre CTFL e ‘On the Concept of Logical Consequence’ (CLC, 1936), para um exame da caracterização da noção de consequência lógica em termos da noção de verdade. Mas encerremos os exemplos aqui, para evitar o enfadamento.

A rede dessas conexões é bastante complexa. Não sabemos se tais conexões estavam claramente presentes à consciência de Tarski nesse período, ou mesmo se lhe pareciam interessantes. O fato é que elas existem, e parece não haver na literatura – até onde sabemos – um tratamento sistemático e explícito delas. Esta dissertação pretendeu ser, entre outras coisas, um exercício preliminar, um primeiro passo em direção à realização desse projeto historiográfico maior.

Apêndice 1 – *O paradoxo de Eubúlides-Epimênides*

A análise Tarskiana do paradoxo de Eubúlides-Epimênides

Vejamos a análise que Tarski faz do paradoxo de Eubúlides-Epimênides em SCTFS. A versão que ele apresenta, e que vamos agora reproduzir, é devida a Lukasiewicz.

Considere-se a seguinte sentença:

- (1) A sentença escrita à página *i*, linhas 8-10, do Apêndice 1 da dissertação intitulada *Alfred Tarski e o Conceito de Verdade em Linguagens Formalizadas*, do aluno Leonardo Weber Castor de Lima, não é verdadeira.

Por razões de economia tipográfica, chamemos essa sentença de *S*.

Agora, exibamos as condições necessárias e suficientes para que *S* seja considerada verdadeira, a partir do Esquema T:

- (2) *S* é verdadeira se, e somente se, a sentença escrita à página *i*, linhas 8-10, do Apêndice 1 da dissertação intitulada *Alfred Tarski e o Conceito de Verdade em Linguagens Formalizadas*, do aluno Leonardo Weber Castor de Lima, não é verdadeira.

Ora, observando qual é de fato a sentença escrita à página *i*, linhas 8-10, do Apêndice 1 da dissertação intitulada *Alfred Tarski e o Conceito de Verdade em Linguagens Formalizadas*, do aluno Leonardo Weber Castor de Lima, e lembrando que convencionamos denotá-la por *S*, podemos enunciar o seguinte:

- (3) A sentença escrita à página *i*, linhas 8-10, do Apêndice 1 da dissertação intitulada *Alfred Tarski e o Conceito de Verdade em Linguagens Formalizadas*, do aluno Leonardo Weber Castor de Lima, é idêntica a *S*.

Agora, usando a lei de Leibniz (princípio de substituição dos idênticos) no lado direito da equivalência 2, obtemos a seguinte sentença:

($\perp$) S é verdadeira se, e somente se, S não é verdadeira.

Essa sentença, evidentemente, é um absurdo.

A solução de um paradoxo exige, em primeiro lugar, o exame das assunções implícitas que são relevantes para a sua produção, e, em segundo lugar, a eliminação das assunções que realmente conduzem a ele. Vejamos os resultados a que Tarski chega, após examinar o paradoxo de Eubúlides-Epimênides.

Três coisas foram assumidas implicitamente:

I. Assumiu-se que a linguagem na qual o paradoxo é construído, contém, além de suas expressões:

- Os nomes de suas expressões (que podem ser facilmente obtidos pelo procedimento usual de se encerrarem estas com aspas, ou então por meio de descrições estruturais);
- O termo semântico "verdadeiro", referente a sentenças nela construídas;
- Todas as sentenças que determinam o uso adequado do termo semântico "verdadeiro" (isto é, todas as instâncias do Esquema T).

II. Assumiu-se que, na linguagem em que o paradoxo é construído, valem as leis ordinárias da lógica (p.e., o princípio de bivalência, a lei do terceiro excluído, a lei de não-contradição, a lei de substituição dos idênticos etc.).

III. Assumiu-se que, na linguagem em que o paradoxo é construído, sentenças empíricas, como (3), podem ser formuladas e asseridas.

Uma linguagem em que vigem as propriedades admitidas na assunção I, é chamada de 'semanticamente fechada'.

Segundo Tarski, a assunção III é dispensável, pois o paradoxo de Eubúlides-Epimênides pode ser reformulado sem o auxílio de sentenças empíricas como a n.º 3, embora de um modo um pouco mais complicado (*vide* próximo tópico). As assunções básicas, portanto, sobre as quais repousa o paradoxo em exame, são as assunções I e II.

Se se pretende excluir a possibilidade de construção desse paradoxo – e tal deve ser uma das pretensões de qualquer definição satisfatória da noção de verdade –, é preciso excluir uma das suposições. Como é insensato, para Tarski, recusar a suposição II, a saber, a validade das leis ordinárias da lógica clássica, resta a exclusão da suposição I, isto é, o uso de uma linguagem semanticamente fechada.

O paradoxo de Eubúlides-Epimênides sem premissas empíricas

Em várias versões, a construção do paradoxo de Eubúlides-Epimênides envolve o recurso a expressões demonstrativas (como ‘esta sentença é falsa’ ou ‘o que eu estou dizendo é falso’) ou a sentenças cuja asserção depende de algum fato empírico (como a inspeção visual de que uma sentença escrita em tal e tal lugar, é a mesma sentença que denotamos por um determinado nome). Pode parecer, assim, que o uso de demonstrativos ou de sentenças empíricas é essencial à construção desse paradoxo – e de um paradoxo semântico em geral. Por conseguinte, pode-se pensar que linguagens formais, por não recorrerem a tais expedientes, são imunes ao paradoxo de Eubúlides-Epimênides. A versão de Lukasiewicz, apresentada no tópico anterior, mostra que é possível reconstruir esse paradoxo sem utilizar expressões demonstrativas. No entanto, ainda faz uso de sentenças empíricas. Em SCTFS¹⁴¹, Tarski afirma que é possível uma reformulação em linguagens naturais sem recorrer a ‘premissas empíricas’. A versão – completamente dedutiva – faz uso de uma noção semântica que Tarski chama de ‘auto-aplicabilidade’ de sentenças. Vejamo-la.

Definição 1. Seja (S) uma sentença qualquer da língua portuguesa, iniciada com as palavras ‘toda sentença’. Associamos à sentença (S) a sentença (S)*, construída da seguinte forma: substituímos a palavra ‘toda’ pelo artigo ‘a’, e inserimos após a palavra ‘sentença’ a própria sentença (S) encerrada entre aspas, mantendo o restante inalterado.

Por exemplo, se (S) é a sentença ‘toda sentença compõe-se de sujeito e predicado’, a sentença (S)* é ‘a sentença ‘toda sentença compõe-se de sujeito e predicado’ compõe-se de sujeito e predicado’.

Definição 2. Dizemos que (S) é auto-aplicável, se, e somente se, (S)* é verdadeira. Dizemos que (S) é não-auto-aplicável, se, e somente se, (S)* é falsa.

¹⁴¹ SCTFS, nota 11.

Consideremos a sentença abaixo:

(S) Toda sentença é não-auto-aplicável.

Sua correlata estrelada é, de acordo com a definição 1:

(S)* A sentença ‘toda sentença é não-auto-aplicável’ é não-auto-aplicável.

A sentença ‘toda sentença é não-auto-aplicável’ é auto-aplicável? Ou é não-auto-aplicável? Mostraremos o seguinte:

Paradoxo da auto-aplicabilidade. A sentença ‘toda sentença é não-auto-aplicável’ é auto-aplicável, se, e somente se, a sentença ‘toda sentença é não-auto-aplicável’ é não auto-aplicável.

Em primeiro lugar, suponhamos que vale a auto-aplicabilidade de (S):

(1) A sentença ‘toda sentença é não-auto-aplicável’ é auto-aplicável.

Dessa hipótese e da definição de auto-aplicabilidade, segue que:

(2) ‘A sentença ‘toda sentença é não-auto-aplicável’ é não-auto-aplicável’ é verdadeira.

Então, deduzimos facilmente de (2), com base no Esquema T, o seguinte:

(3) A sentença ‘toda sentença é não-auto-aplicável’ é não-auto-aplicável.

Conclui-se, portanto, de (1)-(3), que:

(4) Se a sentença ‘toda sentença é não-auto-aplicável’ é auto-aplicável, então a sentença ‘toda sentença é não-auto-aplicável’ é não-auto-aplicável.

Agora, em segundo lugar, suponhamos que vale a não-auto-aplicabilidade de (S):

(5) A sentença ‘toda sentença é não-auto-aplicável’ é não-auto-aplicável.

Dessa hipótese e da definição de não-auto-aplicabilidade, segue que:

(6) ‘A sentença ‘toda sentença é não-auto-aplicável’ é não-auto-aplicável’ é falsa.

Dada uma sentença S qualquer, sabemos que S é falsa, se, e somente se, S não é verdadeira. Mas S não é verdadeira, se, e somente se, $\sim S$ é verdadeira. Então, por transitividade, S é falsa, se, e somente se, $\sim S$ é verdadeira. Assim, desta última propriedade e de (6), deduzimos que:

(7) ‘A sentença ‘toda sentença é não-auto-aplicável’ não é não-auto-aplicável’ é verdadeira.

Então, deduzimos facilmente de (7), com base no Esquema T, o seguinte:

(8) A sentença ‘toda sentença é não-auto-aplicável’ não é não-auto-aplicável.

Obviamente, dada uma sentença S qualquer, S é auto-aplicável, se, e somente se, S não é não-auto-aplicável. Aplicando essa propriedade a (8), decorre o seguinte:

(9) A sentença ‘toda sentença é não-auto-aplicável’ é auto-aplicável.

Conclui-se, portanto, de (5)-(9), que:

(10) Se A sentença ‘toda sentença é não-auto-aplicável’ é não-auto-aplicável, então a sentença ‘toda sentença é não-auto-aplicável’ é auto-aplicável.

Portanto, de (4) e (10), deduzimos, finalmente:

($\perp$) A sentença ‘toda sentença é não-auto-aplicável’ é auto-aplicável, se, e somente se, a sentença ‘toda sentença é não-auto-aplicável’ é não auto-aplicável.

Quod est absurdum.

Apêndice 2 – *Duas demonstrações do teorema da indefinibilidade da verdade, e sua relação com o primeiro teorema de incompletude de Gödel*

Preliminares: algumas informações históricas

Em sua versão original, o teorema de Tarski da indefinibilidade da verdade (CTFL, teorema I-§5) estabelece que o conceito de verdade em linguagens de ordem infinita baseadas em TCS é *absolutamente* indefinível (mais especificamente, isso é demonstrado para a linguagem da teoria geral das classes). No pós-escrito de CTFL (que foi introduzido na tradução alemã de 1936), Tarski estende seus resultados a linguagens não baseadas em TCS; com o relaxamento de TCS, a indefinibilidade passa a ser apenas *relativa*. Entretanto, ainda no pós-escrito, as linguagens que Tarski leva em consideração são baseadas na teoria dos tipos (com tipos de ordem transfinita). É apenas em ‘On Undecidable Statements in Enlarged Systems of Logic and the Concept of Truth’ (1939) que Tarski explicita as condições de adaptação de seus resultados às linguagens de teorias não fundadas na teoria dos tipos. De qualquer modo, seja no pós-escrito de CTFL, seja em ‘On Undecidable Statements etc.’, não há uma nova demonstração do teorema da indefinibilidade da verdade. A demonstração que hoje nos é familiar, referente à aritmética em primeira ordem, baseia-se num teorema mais geral, demonstrado por Tarski, Robinson e Mostowski, na segunda parte de *Undecidable Theories* (1953, teorema I). Neste segundo Apêndice, passaremos em revista a demonstração desse teorema, e mostraremos como o teorema de Tarski da indefinibilidade da verdade é facilmente dele deduzido. Em seguida, evidenciaremos a pedra de toque embutida nesses resultados, o chamado ‘teorema do ponto fixo’, e mostraremos como o teorema de Tarski nela se apóia. Finalmente, mostraremos que o primeiro teorema de incompletude de Gödel é uma simples consequência do teorema de Tarski.

Primeira demonstração: a prova em *Undecidable Theories*

Antes de mais nada, vejamos algumas definições. Seja T uma teoria de primeira ordem com igualdade. Denotemos por L_T a linguagem de T . Estabeleçamos o seguinte.

Definição 1. Seja R uma relação n -ária em ω . Dizemos que R é **expressável** em T , se, e somente se, existe uma fórmula $R(x_1, \dots, x_n)$ de L_T , tal que:

- (i) Para todo $m_1, \dots, m_n$ em ω , se $\langle m_1, \dots, m_n \rangle \in R$, então $\vdash_T R(\mathbf{m}_1, \dots, \mathbf{m}_n)$
- (ii) Para todo $m_1, \dots, m_n$ em ω , se $\langle m_1, \dots, m_n \rangle \notin R$, então $\vdash_T \sim R(\mathbf{m}_1, \dots, \mathbf{m}_n)$

Definição 2. Seja uma função $f : \omega^n \longrightarrow \omega$. Dizemos que f é **representável** em T , se, e somente se, existe uma fórmula $F(x_1, \dots, x_n, y)$ de L_T , tal que:

- (i) Para todo $m_1, \dots, m_n$ e n em ω , se $f(m_1, \dots, m_n) = k$, então $\vdash_T F(\mathbf{m}_1, \dots, \mathbf{m}_n, \mathbf{k})$
- (ii) $\vdash_T \forall x \forall y ((F(\mathbf{m}_1, \dots, \mathbf{m}_n, x) \wedge F(\mathbf{m}_1, \dots, \mathbf{m}_n, y)) \rightarrow x = y)$

Escrevendo (ii) em notação abreviada, temos:

- (ii)' $\vdash_T \exists_1 x F(\mathbf{m}_1, \dots, \mathbf{m}_n, x)$

As cláusulas (i) e (ii) acima são equivalentes à seguinte:

- (iii) Para todo $m_1, \dots, m_n$ e n em ω , se $f(m_1, \dots, m_n) = k$, então $\vdash_T \forall x (F(\mathbf{m}_1, \dots, \mathbf{m}_n, x) \leftrightarrow x = \mathbf{k})$

Terminologia: Tarski, Robinson e Mostowski, em *Undecidable Theories* (Parte II), usam o epíteto comum ‘definibilidade de relações e funções em T ’ para o que aqui estamos chamando separadamente de ‘expressabilidade de relações em T ’ e ‘representabilidade de funções em T ’. Este segundo uso aparece, por exemplo, em Mendelson (*Introduction to Mathematical Logic*) e em Boolos (*Computability and Logic*).

Definição 3. Admitamos que L_T foi convenientemente aritmetizada. Denotamos por $\# \varepsilon$ o número de Gödel da expressão ε de L_T , e por ‘ ε ’ o numeral de L_T correspondente ao número $\# \varepsilon$. Agora, seja $\Psi(x)$ uma fórmula de L_T em que a variável x ocorre livre. A diagonalização de $\Psi(x)$, denotada por $\Psi(x)_{\text{diag}}$, é a sentença $\Psi(\Psi(x))$, i.e., o resultado da substituição da variável livre x em $\Psi(x)$ pelo numeral ‘ $\Psi(x)$ ’ (que denota o número de Gödel de $\Psi(x)$, a saber, $\# \Psi(x)$). A função diagonal $D(x)$ é definida como segue:

$$D(\# \Psi(x)) = \# \Psi(x)_{\text{diag}}.$$

Vejamos agora o teorema geral a que nos referimos nas considerações preliminares.

Teorema (Tarski-Robinson-Mostowski, 1953). Seja T uma teoria de primeira ordem com igualdade (ou ‘com formalização *standard*’, na terminologia de Tarski). Denotemos por L_T a linguagem de T . Seja V o conjunto dos números de Gödel das sentenças de L_T válidas em T . Seja D a função diagonal.

Se T é consistente, então D e V não são ambos definíveis em T .

Prova:

Suponha, por absurdo, que D e V são ambos definíveis em T . Então, existem duas fórmulas $D(y, x)$ e $V(x)$ de L_T , tais que, para todo m, n em ω :

1. Se $D(m) = n$, então $\vdash_T \forall x (D('m', x) \leftrightarrow x = 'n')$
2. Se $m \in V$, então $\vdash_T V('m')$
3. Se $m \notin V$, então $\vdash_T \sim V('m')$

Considere a seguinte fórmula:

$$\varphi : \forall x (D(y, x) \rightarrow \sim V(x))$$

Seja sua respectiva forma diagonalizada:

$$\varphi_{\text{diag}} : \forall x (D(' \varphi ', x) \rightarrow \sim V(x))$$

Obviamente, pela definição da função D , temos:

$$4. D(\# \varphi) = \# \varphi_{\text{diag}}$$

Então, de 1 e 4, é imediato:

$$5. \vdash_T \forall x (D(' \varphi ', x) \leftrightarrow x = ' \varphi_{\text{diag}} ')$$

Mostraremos que tanto $\sim V(' \varphi_{\text{diag}} ')$, como $V(' \varphi_{\text{diag}} ')$, podem ser provadas em T , o que obviamente não pode ser mantido com a hipótese de consistência de T . Vejamos, em primeiro lugar, como demonstrar que $\vdash_T \sim V(' \varphi_{\text{diag}} ')$. Suponhamos que φ_{diag} é válida em T , i.e.:

$$(a) \vdash_T \forall x (D(' \varphi ', x) \rightarrow \sim V(x))$$

Eliminando $\forall$ em 5, obtemos a seguinte instância:

$$(b) \vdash_T D(' \varphi ', ' \varphi_{\text{diag}} ') \leftrightarrow ' \varphi_{\text{diag}} ' = ' \varphi_{\text{diag}} '$$

Ora, dado que é um truísmo $' \varphi_{\text{diag}} ' = ' \varphi_{\text{diag}} '$, segue, por eliminação de $\leftrightarrow$ em (b):

$$(c) \vdash_T D(' \varphi ', ' \varphi_{\text{diag}} ')$$

Agora, se eliminarmos $\forall$ da hipótese (a), obtemos a seguinte instância:

$$(d) \vdash_T D(' \varphi ', ' \varphi_{\text{diag}} ') \rightarrow \sim V(' \varphi_{\text{diag}} ')$$

Sendo assim, deduzimos de (c) e (d), por *modus ponens*:

$$(e) \vdash_T \sim V(' \varphi_{\text{diag}} ')$$

Então, de (a)-(e), concluímos:

$$(f) \text{ Se } \varphi_{\text{diag}} \text{ é válida em T, então } \sim V(' \varphi_{\text{diag}} ') \text{ é válida em T.}$$

Agora, suponhamos que φ_{diag} não é válida em T, i.e.:

$$(g) \nvdash_T \forall x (D(' \varphi ', x) \rightarrow \sim V(x))$$

Então, pela definição de V, deduzimos de (g):

$$(h) \# \varphi_{\text{diag}} \notin V$$

Por definibilidade de V em T, segue de 3 e (h):

$$(i) \vdash_T \sim V(' \varphi_{\text{diag}} ')$$

Então, de (g)-(i), concluímos:

$$(j) \text{ se } \varphi_{\text{diag}} \text{ não é válida em T, então } \sim V(' \varphi_{\text{diag}} ') \text{ é válida em T.}$$

Portanto, tendo em vista a tautologia $((\alpha \rightarrow \beta) \wedge (\sim \alpha \rightarrow \beta)) \rightarrow \beta$, segue de (f) e (j):

$$6. \vdash_T \sim V(' \varphi_{\text{diag}} ')$$

Vejamos, então, em segundo lugar, como demonstrar que $\vdash_T V(' \varphi_{\text{diag}} ')$. De 6, pela teoria da igualdade, segue:

$$7. \vdash_T x = ' \varphi_{\text{diag}} ' \rightarrow \sim V(x)$$

Agora, eliminando $\forall$ e $\leftrightarrow$ de 5, obtemos:

$$8. \vdash_T D(' \varphi ', x) \rightarrow x = ' \varphi_{\text{diag}} '$$

De 8 e 7, por transitividade de $\rightarrow$, deduzimos:

$$9. \vdash_T D(' \varphi ', x) \rightarrow \sim V(x)$$

Então, introduzindo $\forall$ em 9, temos:

$$10. \vdash_T \forall x (D(' \varphi ', x) \rightarrow \sim V(x))$$

Lembrando que $\forall x (D(' \varphi ', x) \rightarrow \sim V(x))$ é φ_{diag} , deduzimos, de 10 e definição de V:

$$11. \# \varphi_{\text{diag}} \in V$$

Finalmente, por definibilidade de V em T, segue de 3 e 11:

$$12. \vdash_T V(' \varphi_{\text{diag}} ')$$

Assim, de 6 e 12, segue:

$$13. \vdash_T \sim V(\ulcorner \varphi_{\text{diag}} \urcorner) \text{ e } \vdash_T V(\ulcorner \varphi_{\text{diag}} \urcorner)$$

Mas 13 contradiz a hipótese de consistência de T . Logo, a suposição de que V e D são ambos definíveis em T está incorreta, e, por conseguinte, se T é consistente, então V e D não são ambos definíveis em T , como queríamos demonstrar. ■

Teorema de Tarski (Indefinibilidade da Verdade). Seja T a aritmética elementar, i.e., $T = \{\varphi \in L_T : \langle \omega; +, \bullet, s, 0 \rangle \models \varphi\}$, sendo L_T uma linguagem de primeira ordem com o simbolismo lógico-aritmético usual (com igualdade). O conjunto V dos naturais associados às sentenças de L_T válidas em T , não é definível em T .

Prova:

Como T é consistente (uma vez que, para qualquer estrutura A , e, em particular, $\langle \omega; +, \bullet, s, 0 \rangle$, não existe uma sentença φ tal que $A \models \varphi$ e $A \models \sim \varphi$), e a função D é definível em T , segue imediatamente do teorema anterior que V não é definível em T . ■

Segunda demonstração: auto-referência, teorema do ponto fixo e indefinibilidade da verdade

Vamos demonstrar o teorema que subjaz à técnica de construção de sentenças auto-referentes. Seja K uma teoria de primeira ordem com igualdade, cuja linguagem possui os mesmos símbolos não-lógicos da linguagem de S (a teoria formal dos números naturais, na axiomatização de Mendelson [ou outra que lhe seja equivalente]¹⁴²), e em que toda função recursiva é representável. Seja $A(x)$ uma fórmula de L_K , com a variável livre x . Chamamos de “ponto fixo” de $A(x)$ a uma sentença s de L_K tal que $\vdash_K s \leftrightarrow A(\ulcorner s \urcorner)$.

Teorema do Ponto Fixo. Toda fórmula $A(x)$ de L_K possui um ponto fixo.

Prova:

Mostraremos como construir um ponto fixo para a fórmula $A(x)$. Uma vez que, por hipótese, toda função recursiva é representável em K , segue que a função diagonal $D(x)$ é

¹⁴² MENDELSON, E. *Introduction to Mathematical Logic*, Cap. 3.

representável em K por uma fórmula $D(y, x)$ de L_K , tal que as seguintes condições são satisfeitas:

1. Se $D(m) = n$, então $\vdash_K D('m', 'n')$

2. $\vdash_K \exists_1 x D('m', x)$

Dadas as fórmulas $A(x)$ e $D(y, x)$, construíamos a seguinte fórmula:

$\varphi : \forall x (D(y, x) \rightarrow A(x))$

Como convenciamos, o número de Gödel de φ é $\# \varphi$. Agora, diagonalizando φ , obtemos a seguinte sentença:

$\varphi_{\text{diag}} : \forall x (D(' \varphi ', x) \rightarrow A(x))$

Como convenciamos, o número de Gödel de φ_{diag} é $\# \varphi_{\text{diag}}$. Mas, sendo φ_{diag} a diagonalização de φ , decorre:

3. $D(\# \varphi) = \# \varphi_{\text{diag}}$

Então, de 1 e 3, segue:

4. $\vdash_K D(' \varphi ', ' \varphi_{\text{diag}} ')$

A sentença φ_{diag} é um ponto fixo de $A(x)$, i.e., $\vdash_K \varphi_{\text{diag}} \leftrightarrow A(' \varphi_{\text{diag}} ')$. Senão, vejamos.

$[\rightarrow] \vdash_K \forall x (D(' \varphi ', x) \rightarrow A(x)) \rightarrow A(' \varphi_{\text{diag}} ')$

Seja, por hipótese:

5. $\forall x (D(' \varphi ', x) \rightarrow A(x))$

Por eliminação do quantificador universal, segue:

6. $D(' \varphi ', ' \varphi_{\text{diag}} ') \rightarrow A(' \varphi_{\text{diag}} ')$

De 4 e 6, vem:

7. $A(' \varphi_{\text{diag}} ')$

Assim,

8. $\forall x (D(' \varphi ', x) \rightarrow A(x)) \vdash_K A(' \varphi_{\text{diag}} ')$

Pelo Teorema da Dedução, concluímos:

9. $\vdash_K \forall x (D(' \varphi ', x) \rightarrow A(x)) \rightarrow A(' \varphi_{\text{diag}} ')$

$[\leftarrow] \vdash_K A(' \varphi_{\text{diag}} ') \rightarrow \forall x (D(' \varphi ', x) \rightarrow A(x))$

Seja, por hipótese:

10. $A(' \varphi_{\text{diag}} ')$

De 10, segue:

$$11. x = \phi_{\text{diag}} \rightarrow A(x)$$

Agora, de 2 e 4, deduzimos:

$$12. \vdash_K D(\phi, x) \rightarrow x = \phi_{\text{diag}}$$

Por transitividade da implicação material, temos, a partir de 12 e 11:

$$13. \vdash_K D(\phi, x) \rightarrow A(x)$$

Por introdução do quantificador universal, segue de 13:

$$14. \vdash_K \forall x(D(\phi, x) \rightarrow A(x))$$

Assim,

$$15. A(\phi_{\text{diag}}) \vdash_K D(\phi, x) \rightarrow A(x)$$

Pelo Teorema da Dedução, concluímos:

$$16. \vdash_K A(\phi_{\text{diag}}) \rightarrow \forall x(D(\phi, x) \rightarrow A(x))$$

Finalmente, de 9 e 16, é imediato:

$$17. \vdash_K \forall x(D(\phi, x) \rightarrow A(x)) \leftrightarrow A(\phi_{\text{diag}})$$

Ou seja, lembrando que ϕ_{diag} é $\forall x(D(\phi, x) \rightarrow A(x))$:

$$\vdash_K \phi_{\text{diag}} \leftrightarrow A(\phi_{\text{diag}})$$

Em outros termos, ϕ_{diag} é um ponto fixo de $A(x)$. Isso completa a prova. ■

Obs.: Intuitivamente, ϕ_{diag} diz de si mesma que satisfaz a propriedade expressa por $A(x)$; com efeito, ' $\forall x(D(\phi, x) \rightarrow A(x))$ ' diz "qualquer que seja x , se x é (o código de) a diagonal de ϕ , então x satisfaz a propriedade expressa por $A(x)$ ",¹⁴³ ora, ' $\forall x(D(\phi, x) \rightarrow A(x))$ ' é a diagonal de ϕ , i.e., ϕ_{diag} ; então, ' $\forall x(D(\phi, x) \rightarrow A(x))$ ' diz de si mesma que satisfaz $A(x)$, ou seja, diz "esta sentença satisfaz $A(x)$ " ou " $A(x)$ é verdadeira a meu respeito". Assim, um ponto fixo de uma fórmula $A(x)$ nada mais é do que uma sentença auto-referente, uma sentença que diz de si mesma que satisfaz $A(x)$. Note que, ao contrário de sentenças auto-referentes epimenídicas, cuja auto-referência podemos chamar de 'anômala', a auto-referência de um ponto fixo de uma fórmula $A(x)$ de L_K é inócua (no sentido de que não afeta a consistência de K) – e podemos chamá-la de 'normal'. Um fato interessante relacionado a essas observações é expresso no teorema de Tarski da indefinibilidade da verdade aritmética. Podemos parafraseá-lo dizendo: se o

¹⁴³ Observe que a auto-referência aqui é indireta, tal como o é em sentenças auto-referentes do tipo 'a sentença impressa à página tal, linha tal etc. tem a propriedade P' ou, equivalentemente, 'qualquer que seja a sentença x , se x é a sentença impressa à página tal, linha tal etc. tem a propriedade P'.

conceito de verdade aritmética (em primeira ordem) fosse aritmeticamente definível (em primeira ordem), então existiriam pontos fixos epimenídicos (i.e., sentenças m tais que $N \models m \leftrightarrow \sim V('m')$). Por conseguinte, poderíamos deduzir uma contradição (de maneira análoga à construção do Paradoxo de Epimênides), e, assim, a aritmética seria inconsistente. É o que vamos demonstrar abaixo.

Teorema de Tarski (Indefinibilidade da Verdade). Seja V o conjunto dos números de Gödel das fórmulas de L_K verdadeiras em $N = \langle \omega; +, \bullet, s, 0 \rangle$, sendo ω o conjunto dos números naturais, $+$, $\bullet$, s , as operações usuais de adição, multiplicação e sucessão em ω , e 0 o elemento neutro da adição. Em símbolos, $V = \{ \# \varphi : N \models \varphi \}$. V não é aritmeticamente definível, i.e., não existe fórmula $V(x)$ de L_K tal que (i) se $\# \varphi \in V$, então $N \models V(' \varphi')$, e (ii) se $\# \varphi \notin V$, então $N \not\models V(' \varphi')$.

Prova:

Suponhamos, por absurdo, que V é aritmeticamente definível; assim, existe uma fórmula $V(x)$ de L_K tal que satisfaz as seguintes condições:

1. se $\# \varphi \in V$, então $N \models V(' \varphi')$
2. se $\# \varphi \notin V$, então $N \not\models V(' \varphi')$

Segue imediatamente da definição de V :

3. $\# \varphi \in V$ se, e somente se, $N \models \varphi$

Considere a fórmula $\sim V(x)$ de L_K . Então, pelo Teorema do Ponto Fixo, $\sim V(x)$ possui um ponto fixo – digamos, γ .

4. $N \models \gamma \leftrightarrow \sim V(' \gamma')$

Agora, ou $\# \gamma \in V$, ou $\# \gamma \notin V$. Seja, por hipótese:

5. $\# \gamma \in V$

Então, de 5 e 1, segue:

6. $N \models V(' \gamma')$

Mas, de 6 e 4, deduzimos:

7. $N \models \sim \gamma$

Ocorre que temos a seguinte propriedade de $\models$:

8. Se $N \models \sim \gamma$, então $N \not\models \gamma$

Assim, segue imediatamente de 7 e 8:

$$9. N \models \gamma$$

De 3 e 9, concluímos:

$$10. \# \gamma \notin V$$

Então, de 5-10, segue:

$$11. \text{ Se } \# \gamma \in V, \text{ então } \# \gamma \notin V$$

Agora, assumamos o seguinte:

$$12. \# \gamma \notin V$$

Então, de 2 e 12, vem:

$$13. N \models V(\gamma')$$

É uma propriedade de $\models$:

$$14. \text{ Se } N \models V(\gamma'), \text{ então } N \models \sim V(\gamma')$$

Logo, decorre de 13 e 14:

$$15. N \models \sim V(\gamma')$$

Então, de 4 e 15, deduzimos:

$$16. N \models \gamma$$

E, de 16 e 3, concluímos:

$$17. \# \gamma \in V$$

Assim, de 12-17, segue:

$$18. \text{ Se } \# \gamma \notin V, \text{ então } \# \gamma \in V$$

Portanto, de 11 e 18, obtemos:

$$19. \# \gamma \in V \text{ se, e somente se, } \# \gamma \notin V \perp$$

Sendo assim, nossa hipótese de que V é aritmeticamente definível, conduziu-nos a um absurdo, e, por conseguinte, V não é aritmeticamente definível. Isso encerra a prova.

■

Obs.: Intuitivamente, como toda fórmula $A(x)$ de L_K possui um ponto fixo, se V fosse definível aritmeticamente por uma fórmula $V(x)$ de L_K , a fórmula $\sim V(x)$ teria um ponto fixo s (um ponto fixo epimenídico), e, por conseguinte, dadas algumas propriedades básicas do conceito de verdade em N , facilmente deduziríamos uma sentença contraditória em K , de maneira análoga à dedução de uma sentença epimenídica em linguagem natural.

O teorema da indefinibilidade da verdade implica o primeiro teorema de incompletude de Gödel

O primeiro teorema de incompletude de Gödel é facilmente deduzido do teorema de Tarski da indefinibilidade da verdade. Informalmente, a prova é a seguinte. Considere uma axiomatização adequada da teoria dos números naturais, p.e., o sistema S de Mendelson. É demonstrável que o conjunto T_S dos (números associados aos) teoremas de S é definível aritmeticamente. Mas, pelo teorema de Tarski, o conjunto V dos números associados às sentenças de L_S verdadeiras no modelo *standard* da aritmética elementar, não é aritmeticamente definível. Por conseguinte, $T_S \neq V$. Então existe uma sentença s tal que $s \in V$ e $s \notin T_S$. Como $s \in V$, segue que $\sim s \notin V$. Agora, é claro que $T_S \subset V$ (i.e., todo teorema de S é verdadeiro). Então, $\sim s \notin T_S$. Portanto, $s \in V$, $s \notin T_S$ e $\sim s \notin T_S$, i.e., s é uma sentença indecidível em S , e, *ipso facto*, S é incompleto. Isso encerra a prova. ■

Índice de abreviações

AERS – ‘Arithmetical Extensions of Relational Systems’

CLC – ‘The Concept of Logical Consequence’

CTFL – ‘The Concept of Truth in Formalized Languages’

DSRN – ‘On Definable Sets of Real Numbers’

ESS – ‘The Establishment of Scientific Semantics’

FCS - ‘Foundations of the Calculus of Systems’

SCTFS – ‘The Semantic Conception of Truth and the Foundations of Semantics’

SMIDC – ‘Some Methodological Investigations on the Definability of Concepts’

SOC ω CC – ‘Some Observations on the Concepts of ω -Consistency and ω -Completeness’.

TP – ‘Truth and Proof’

USES – ‘On Undecidable Statements in Enlarged Systems of Logic and the Concept of Truth’

UT – *Undecidable Theories*

Referências I

- BIRABEN, R.E. *Tese de Church: Algumas Questões Histórico-conceituais*. Coleção CLE, v. 16, Unicamp, Campinas, 1996.
- BOOLOS, G.; JEFFREY, R. *Computability and Logic*. Cambridge University Press, 1974.
- CHANG, C. C.; KEISLER, H. J. *Model Theory*, 3rd ed., Elsevier Science Publishers, 1990.
- DEDEKIND, R. ‘Continuity and Irrational Numbers’. In: *Essays on the Theory of Numbers*. Transl. Wooster Woodruff Beman, New York, Dover, 1963.
- DEDEKIND, R. ‘The Nature and Meaning of Numbers’. In: *Essays on the Theory of Numbers*. Transl. Wooster Woodruff Beman, New York, Dover, 1963.
- DONER, J.; HODGES, W. ‘Alfred Tarski’s and Decidable Theories’, JSL, v. 53, n.º 1, 1988.
- ETCHEMENDY, J. ‘Tarski on Truth and Logical Consequence’, JSL, v. 53, n.º 1, 1988.
- FEFERMAN, S. ‘Tarski’s Conceptual Analysis of Semantical Notions’, 2002. *Sémantique et épistémologie* (A. Benmakhlouf, ed.) Editions Le Fennec, Casablanca (2004) [distrib. J. Vrin, Paris] 79-108.
- FIELD, H. ‘Tarski’s Theory of Truth’. *Journal of Philosophy*, 69 (1972).
- HEIJENOORT, J. *From Frege to Gödel – A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, 1967.
- HERBRAND, J. ‘Investigations in Proof Theory’ (1930). In: van Heijenoort, J. *From Frege to Gödel – A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, 1967.

HODGES, W. 'Truth in a Structure', *Proceedings of the Aristotelian Society*, NS 86, 1986.

KNEALE, W.; KNEALE, M. *O Desenvolvimento da Lógica*, 3.^a ed., trad. M. S. Lourenço, Calouste Gulbenkian, Lisboa, 1991.

KREISEL, G. 'Gödel's Excursions into Intuitionistic Logic', *apud* Wolenski, J. 'From Intentionality to Formal Semantics (from Twardowski to Tarski)', *Erkenntnis*, 56, 2002

McNULTY, G. F. 'Alfred Tarski's and Undecidable Theories', *JSL*, v. 51, n.º 4, 1986.

MENDELSON, E. *Introduction to Mathematical Logic*. D. Van Nostrand Company, 1979.

MOSTOWSKI, A. 'Alfred Tarski'. *The Encyclopedia of Philosophy*, New York, 1967, v. 8, pp. 77-81.

PADOA, A. 'Logical Introduction to any Deductive Theory' (1900). In: van Heijenoort, J. *From Frege to Gödel – A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, 1967.

PASCAL, B. *De L'esprit Géométrique*. GF-Flammarion, Paris, 1985.

POPPER, K. 'Some Philosophical Comments on Tarski's Theory of Truth'. In: Leon Henkin *et al.* (eds.), *Proceedings of the Tarskian Symposium, Proceedings of Symposia in Pure Mathematics*, v. XXV, Providence RI, American Mathematical Society, 1974.

SCANLAN, M. 'Who Were the American Postulate Theorists?', *JSL*, v. 56, n. 3, 1991.

SUPPES, P. *Introduction to Logic*. D. Van Nostrand Company, 1957.

SUPPES, P. *Axiomatic Set Theory*. Dover, New York, 1972.

TARSKI, A. *Introduction to Logic and to the Methodology of Deductive Sciences*. Dover, New York, 1995.

TARSKI, A. 'The Concept of Truth in Formalized Languages'. In: *Logic, Semantics, Metamathematics*, trad. J. H. Woodger, J. Corcoran (ed.), Hackett, 2nd ed., 1983.

TARSKI, A. 'On Definable Sets of Real Numbers', *idem*.

TARSKI, A. 'Logical Operations and Projective Sets', *ibidem*.

TARSKI, A. 'Some Observations on the Concepts of ω -Consistency and ω -Completeness', *ibidem*.

TARSKI, A. 'The Establishment of Scientific Semantics', *ibidem*. TARSKI, A.. 'Foundations of the Calculus of Systems', *ibidem*.

TARSKI, A.. 'On the Foundations of Boolean Algebra', *ibidem*.

TARSKI, A. 'The Concept of Logical Consequence', *ibidem*.

TARSKI, A. 'On Some Fundamental Concepts of Metamathematics', *ibidem*.

TARSKI, A. 'Fundamental Concepts of the Methodology of the Deductive Sciences', *ibidem*.

TARSKI, A. 'Investigations into the Sentential Calculus', *ibidem*.

TARSKI, A. 'Some Methodological Investigations on the Definability of Concepts', *ibidem*.

TARSKI, A. 'The Semantic Conception of Truth and the Foundations of Semantics', *Philosophy and Phenomenological Research*, iv, 1944.

TARSKI, A.. 'Truth and Proof', *Scientific American*, 220, (6), 1969.

TARSKI, A. 'On Undecidable Statements in Enlarged Systems of Logic and the Concept of Truth'. JSL, iv, 1939.

TARSKI, A. 'A Problem Concerning the Notion of Definability'. JSL, xiii, 1948.

TARSKI, A. *A Decision Method for Elementary Algebra and Geometry*. RAND, Santa Monica, Calif., 1948.

TARSKI, A.; KURATOWSKI C. ‘Logical Operations and Projective Sets’. In: *Logic, Semantics, Metamathematics*.

TARSKI, A.; LINDENBAUM, A. ‘On the Limitations of the Means of Expression of Deductive Theories’. In: *Logic, Semantics, Metamathematics*.

TARSKI, A.; VAUGHT, R. ‘Arithmetical Extensions of Relational Systems’. *Compositio Math.*, 13, pp. 81-102, 1958.

TARSKI, A.; ROBINSON, R.; MOSTOWSKI, A. *Undecidable Theories*, Amsterdam, North-Holland, 1953.

VAUGHT, R. ‘Alfred Tarski’s Work in Model Theory’, *JSL*, v. 51, n.º 4, 1986.

VAUGHT, R. ‘Model Theory Before 1945’, in Leon Henkin *et al.* (eds.), *Proceedings of the Tarskian Symposium, Proceedings of Symposia in Pure Mathematics*, v. XXV, Providence RI, American Mathematical Society, 1974.

WOLENSKI, J. ‘From Intentionality to Formal Semantics (from Twardowski to Tarski)’, *Erkenntnis*, 56, 2002.

WOLENSKI, J. ‘Theories of Truth in Austrian and Polish Philosophy’. Acesso em: <http://www.fmag.unict.it/~polphil/PolPhil/LvovWarsaw/WolTruth.html>

Referências II*

ACKERMANN, W.; HILBERT, D. *Grundzüge der theoretischen Logik* (1928), Springer, Berlin. English transl., *Principles of Mathematical Logic*, Chelsea, New York, 1950.

BERNAYS, P.; SCHÖNFINKEL, M. ‘Zum Entscheidungsproblem der mathematischen Logik’ (1928). *Mathematische Annalen*, XCIX, pp. 342-372.

CHURCH, A. 'An Unsolvable Problem of Elementary Number Theory' (1936). *Amer. J. Math.*, 58, pp. 345-363.

GÖDEL, K. 'Über die Vollständigkeit des Logikkalküls' (1929). In: *Kurt Gödel Collected Works*, v. I, Oxford University Press, New York, 1986.

GÖDEL, K. 'Die Vollständigkeit der Axiome des logischen Funktionenkalküls' (1930). *Monatshefte für Mathematik und Physik*, v. 37, pp. 349-360. In: *Kurt Gödel Collected Works*, v. I, Oxford University Press, New York, 1986.

HENKIN, L. 'The Completeness of First-Order Functional Calculus'. *JSL*, v. 14, n.º 3, 1949.

KEMENY, J. 'Models of Logical Systems'. *JSL*, v. 13, n.º 1, 1948.

LANGFORD, C. H. 'Some Theorems on Deducibility' (1926). *Annals of Mathematics*, s. 2, v. 28, pp. 16-40.

LANGFORD, C. H. 'Theorems on Deducibility' (1927). *Annals of Mathematics*, s. 2, v. 29, pp. 459-471.

LÖWENHEIM, L. 'Über Möglichkeit im Relativkalkül' (1915). *Mathematische Annalen*, v. 76, pp. 447-470.

POST, E. L. 'Introduction to a General Theory of Elementary Propositions' (1921). *Amer. J. Math.*, 43, pp. 163-185.

PRESBURGUER, M. 'Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt' (1930). *Comptes-rendus du I Congrès des Mathématiciens des Pays Slaves, Warszawa, 1929. Warsaw, 1930*, pp. 92-101.

RAMSEY, F. P. 'The Foundations of Mathematics' (1926). *Proc. London Math. Soc.*, 2, v. 25, pp. 338-384.

RUSSELL, B. 'Mathematical Logic as Based on the Theory of Types' (1908). In: van Heijenoort, J. *From Frege to Gödel – A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, 1967.

SKOLEM, T. 'Untersuchungen über die Axiome des Klassenkalküls und über Produktations- und Summationsprobleme, welche gewisse Klassen von Aussagen betreffen' (1919). *Skifter utgit av Videnskapsselskapet i Kristiania. I, Matematisk-Naturvidenskabelig Klasse*, n. 3.

SKOLEM, T. 'Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen' (1920). *Skifter utgit av Videnskapsselskapet i Kristiania. I, Matematisk-Naturvidenskabelig Klasse*, n. 4.

SKOLEM, T. 'Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre' (1922). *Wissenschaftliche Vorträge gehalten auf dem fünften Kongress der Skandinavischen Mathematiker in Helsingfors vom 4. bis 7. Juli 1922*, Akademische Buchhandlung, Helsinki, 1923, pp. 217-232.

SKOLEM, T. 'Über die mathematische Logik' (1928). *Norsk. Mat. Tidsk.* 10, 125-142. In: van Heijenoort, J. *From Frege to Gödel – A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, 1967.

SKOLEM, T. 'Über einige Grundlagenfragen der Mathematik' (1929). *Skifter utgit av Det Norske Videnskaps-Akademi I Oslo, I. Matematisk-naturvidenskapelig klasse* 1929, n. 4.

TURING, A. 'On Computable Numbers, with an Application to the Entscheidungsproblem' (1936-7). *Proc. London Math. Soc.*, (2), 42, pp.230-265; correction, *ibid.* 43, pp. 544-546.

VEBLEN, O. 'A System of Axioms for Geometry' (1904). *Transactions of American Mathematical Society*, V, pp. 343-381.

WITTGENSTEIN, L. *Tractatus Logico-Philosophicus*. Trad. Luiz Henrique Lopes dos Santos. São Paulo: Edusp, 1993.

ZERMELO, E. ‘Untersuchungen über die Grundlagen der Mengenlehre I’ (1908).
Mathematische Annalen, v. 65, pp. 261-281.

(*) artigos apenas citados ou consultados indiretamente.