HENRIQUE ANTUNES ALMEIDA

REVISITANDO O TEOREMA DE FREGE

Campinas

2014

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

HENRIQUE ANTUNES ALMEIDA

REVISITANDO O TEOREMA DE FREGE

ORIENTADOR: WALTER ALEXANDRE CARNIELLI

Dissertação de Mestrado apresentada ao Instituto de Filosofia e Ciências Humanas para a obtenção do Título de Mestre em Filosofia.

Este exemplar corresponde à versão final da dissertação defendida pelo aluno Henrique Antunes Almeida, e orientada pelo Prof. Dr. Walter Alexandre Carnielli.

Campinas
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Filosofia e Ciências Humanas
Cecília Maria Jorge Nicolau - CRB 8/338

AL64r

Almeida, Henrique Antunes, 1989-
Revisitando o Teorema de Frege / Henrique Antunes Almeida. – Campinas, SP : [s.n.], 2014.

Orientador: Walter Alexandre Carnielli.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Frege, Gottlob, 1848-1925. 2. Abstração. 3. Aritmética . 4. Matemática - Filosofia. 5. Lógica simbólica e matemática. I. Carnielli, Walter Alexandre,1952-. II. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Revisiting Frege's Theorem
Palavras-chave em inglês:
Abstraction
Arithmetic
Mathematics - Philosophy
Symbolic and mathematical logic
Área de concentração: Filosofia
Titulação: Mestre em Filosofia
Banca examinadora:
Walter Alexandre Carnielli [Orientador]
Marco Antonio Caron Ruffino
Abílio Azambuja Rodrigues Filho
Data de defesa: 26-09-2014
Programa de Pós-Graduação: Filosofia

UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

A Comissão Julgadora dos trabalhos de Defesa de Dissertação de Mestrado, em sessão pública realizada em 26 de setembro de 2014, considerou o candidato HENRIQUE ANTUNES ALMEIDA aprovado.

Este exemplar corresponde à redação final da Dissertação defendida e aprovada pela Comissão Julgadora.

Prof. Dr. Walter Alexandre Carnielli

Prof. Dr. Marco Antonio Caron Ruffino

Prof. Dr. Abílio Azambuja Rodrigues Filho

Prof. Dr. Rodrigo de Alvarenga Filho

Prof. Dr. Giorgio Venturi

Resumo

Neste trabalho, abordamos o *Teorema de Frege* sob uma perspectiva técnica. Primeiramente, propomos uma caracterização geral de linguagens de segunda ordem que sejam adequadas para formalizar quaisquer *teorias fregeanas* – teorias que resultam da introdução de um ou mais *princípios de abstração* a um sistema dedutivo de lógica de segunda ordem; fornecemos uma semântica e um sistema dedutivo para essas linguagens e elaboramos alguns resultados metateóricos acerca desse sistema. Em segundo lugar, apresentamos uma exposição detalhada da prova do Teorema de Frege, enunciado como uma relação entre a *Aritmética de Frege* e a *Aritmética de Dedekind-Peano*. Por fim, provamos a equiconsistência entre essas teorias e a *Aritmética de Peano de Segunda Ordem*.

Palavras-chave: *Princípio de Abstração*, *Princípio de Hume*, *Aritmética de Frege*, *Teorema de Frege*

Abstract

In this work, we discuss *Frege's Theorem* under a technical perspective. First, we propose a general caracterization of second-order languages suitable to formalize all *Fregean theories* – theories that result from the introduction of one or more *abstraction principles* to a deductive system of second-order logic; we also furnish a semantics and a deductive system for these languages and establish a few metatheorical results about the system. Second, we present a detailed proof of Frege's Theorem, formulated as a relation between *Frege's Arithmetic* and *Dedekind-Peano Arithemtic*. Finally, we prove the equiconsistency between these theories and *Peano Second-Order Arithmetic*.

Keywords: *Abstraction Principle*, *Hume's Principle*, *Frege's Arithmetic*, *Frege's Theorem*

Sumário

Introdução **1**

1 Lógica de Segunda Ordem **9**

1.1 Sintaxe 13

1.2 Semântica 22

1.3 Um Sistema Dedutivo 26

1.3.1 Axiomas Lógicos 26

1.3.2 Metateoremas, Regras Derivadas e Demonstrações 30

1.3.3 O Operador de Abstração 33

1.4 Teorias, Definições e Interpretações 42

2 O Teorema de Frege **53**

2.1 Aritmética de Frege 54

2.2 Aritmética de Dedekind-Peano 57

2.3 A Prova do Teorema 60

2.3.1 Relação Ancestral e os Princípios de Indução Generalizada 61

2.3.2 Demonstração de **ADP**1, **ADP**2, **ADP**7 e **ADP**5 66

2.3.3 Demonstração de **ADP**3 e **ADP**4 70

2.3.4 Demonstração de **ADP**6 72

3 Consistência **81**

3.1 Aritmética de Peano de Segunda Ordem 83

3.2 Uma Prova de Consistência Relativa de **AF** 86
3.2.1 Uma Prova Modelo-Teórica da Consistência de **AF** 87
3.2.2 O Operador η 88
3.2.3 Demonstração de Fin_1 93
3.2.4 Finalização da Prova 101

Conclusão **103**

Apêndice A **105**
A1. O Operador de Abstração 105
A2. Teorema da Interpretação 108

Apêndice B **117**
B1. Demonstração de *Anc*1-12 117
B2. Demonstração de **ADP**4.1.1 119

Apêndice C **123**

Bibliografia **126**

a Patrícia

Agradecimentos

Agradeço a todos aqueles que direta ou indiretamente contruibuíram para a execução e conclusão deste trabalho. Ao meu orientador, professor Walter A. Carnielli, que me concedeu a liberdade na escolha do tema de pesquisa e cuja solicitude e confiança foram fundamentais durante todo o percurso. Ao professor Marco A. C. Ruffino, pelos conselhos sempre úteis e pela sugestão valiosa de que eu me inscrevesse no curso de Mestrado da UNICAMP em 2011. Aos professores Abílio A. Rodrigues Filho, Elaine G. Pimentel e Túlio R. X. de Aguiar, que desempenharam um papel muito importante na minha formação durante os anos de graduação; e a Richard L. Epstein, com quem tive a oportunidade de conviver durante a minha estadia na ARF Library, no segundo semestre de 2013, e cujo encorajamento obstinado foi indispensável na etapa final do trabalho.

Agradeço aos professores Marcelo E. Coniglio e Itala M. L. D'Ottaviano, pelos cursos ministrados, e aos colegas de pós-graduação, por me acolherem amistosamente em Campinas. Em particular, a Edgar L. B. de Almeida e Leandro O. Suguitani, com quem aprendi muito e que sempre estiveram dispostos a discurtir algum tópico relativo ao trabalho. Agradeço também a Guilherme A. Cardoso, que, durante a graduação, foi um interlocutor estimulante nos estudos ainda incipientes em lógica e filosofia.

Agradeço a minha família, Omar, Silvana e Rafael, por todo o apoio; e a Patrícia, pelo companheirismo e dedicação inestimáveis.

Introdução

Durante grande parte de sua carreira, Gottlob Frege sustentou a chamada *tese logicista* em filosofia da matemática, enunciada em seu *Die Grundlagen der Arithmetik* (1884) como a tese de que as verdades da aritmética elementar e da análise real são analíticas, no sentido de que elas poderiam ser provadas unicamente a partir de leis lógicas gerais e definições apropriadas. Com o intuito de demonstrar o logicismo nessa formulação particular, Frege empreendeu uma tentativa de derivar os axiomas dessas teorias matemáticas no sistema formal apresentado em seu *Grundgesetze der Arithmetik* (vol. I, 1893; vol. II, 1903), elaborado a partir do sistema prévio do *Begriffsschrift* (1879) e que, salvo muitas especificidades, consistia em um sistema de *lógica de segunda ordem* (*LSO*) suplementado por um axioma destinado a regulamentar o comportamento do que Frege denominou *percursos de valores de funções*. Esse axioma, a *Lei Básica V*, e o sistema como um todo mostraram-se, contudo, inconsistentes por ocasião da descoberta do *paradoxo de Russell*, comunicado a Frege por Bertrand Russell em junho de 1902, quando o segundo volume do *Grundgesetze* ainda estava no prelo. Ao reconhecer imediatamente a inconsistência de seu sistema, Frege adicionou um apêndice ao novo volume no qual ele creditava a sua descoberta a Russell e expunha a derivação formal do paradoxo. Após ensaiar algumas soluções mal-sucedidas para reparar a Lei Básica V, ele abandonou de uma vez por todas a tese logicista, por volta de 1906.

Embora o sistema do *Grundgesetze* (*Gg*) seja inconsistente, os desenvolvimentos formais elaborados por Frege no livro contêm um resultado de ordem técnica que adquiriu grande relevância filosófica a partir da segunda métade do século passado, o qual passou a ser posteriormente denominado *Teorema de Frege*. Em uma formulação contemporânea, esse teorema

estabelece que os *axiomas de Dedekind-Peano* para a aritmética elementar podem ser derivados em um sistema dedutivo de *LSO* suplementado por um único axioma (não lógico) que governa as condições de identidade entre números cardinais de conceitos, hoje amplamente conhecido como *Princípio de Hume* (*PH*):

Para todos os conceitos F e G,

o número de Fs é igual ao número de Gs se e somente se F é equinumérico a G

Formalmente,

$$\forall F \forall G(\#F = \#G \leftrightarrow F \approx G)$$

em que # é um símbolo de função, denominado *operador de cardinalidade*, que se aplica a variáveis de segunda ordem unárias, resultando termos individuais; e a fórmula $F \approx G$ expressa a relação, definível em termos do vocabulário lógico da *LSO*, entre F e G quando existe uma correlação um a um entre os Fs e os Gs – por ora, a relação de *equinumericidade* pode ser entendida como relação entre F e G quando existe uma função injetiva que correlaciona os objetos que são F aos objetos que são G.

O Teorema de Frege já havia sido informalmente formulado nas seções 70-83 do Capítulo 4 do *Grundlagen* (*Gl*), no qual Frege apresenta um programa a ser tecnicamente desenvolvido posteriormente (no sistema do *Grundgesetze*) e cuja execução estava destinada a estabelecer a tese logicista para o caso particular da aritmética dos números naturais. Nessas seções, Frege apresenta um esboço detalhado de como os axiomas de Dedekind-Peano poderiam ser provados a partir do Princípio de Hume, sem fazer uso da Lei Básica V[1]. Embora estivesse presente no *Grundlagen* e no *Grundgesetze*, no entanto, o teorema somente foi explicitamente

[1]Em *Gl*, §68 Frege define o número cardinal de um conceito F explicitamente como a extensão do conceito de segunda ordem 'conceito equinumérico a F' e, ao que tudo indica, faz um apelo tácito a uma versão da Lei Básica V para provar *PH* a partir dessa definição (em *Gl*, §73). Com o Princípio de Hume em mãos, ele utiliza-o para provar os axiomas de Dedekind-Peano, sem mais recorrer a extensões de conceitos. Esse ponto foi apontado pela primeira vez por Charles Parsons em [29] e, posteriormente, elaborado por Boolos em [4].

reconhecido pela tradição filosófica no final do século passado: em 1983, Crispin Wright [40] apresentou um desenvolvimento semiformal das provas desses axiomas em uma teoria de segunda ordem em que *PH* figura como o único axioma não lógico – denominada posteriormente *Aritmética de Frege* (**AF**)[2] – e tentou reproduzir sem sucesso a derivação do paradoxo de Russell nessa teoria, conjecturando que ela seria, de fato, uma teoria consistente. A conjectura de Wright foi então confirmada pouco tempo depois: em 1984, John Burgess [10] formulou uma prova modelo-teórica da Aritmética de Frege; e, em 1986, George Boolos [3] formulou uma prova de consistência em relação a uma versão de segunda ordem da *Aritmética de Peano*.

A redescoberta do Teorema de Frege e as provas de consistência da teoria homônima motivaram o surgimento de uma proposta de reabilitação do logicismo fregeano, na qual Bob Hale e Crispin Wright reivindicam que, para além de um resultado meramente técnico, ele desempenha um papel central na explicação da natureza *a priori* de nosso conhecimento aritmético: segundo eles, o teorema demonstra que os axiomas Dedekind-Peano e os teoremas que deles se seguem são analíticos, na medida em que eles podem ser provados na lógica de segunda ordem com base em um princípio, *PH*, que define implicitamente o conceito de número cardinal. No ensaio 'On the Philosophical Significance of Frege's Theorem', por exemplo, Wright afirma:

> (...) O Teorema de Frege irá (...) garantir a verdade do seguinte: que as leis fundamentais da aritmética podem ser derivadas em um sistema de lógica de segunda ordem aumentado por um princípio [*PH*] cujo papel é *explicar*, senão exatamente definir, a noção geral de identidade entre números cardinais (...); e que essa explicação procede em termos de uma noção [a relação de equinumericidade entre conceitos] que pode ser definida em termos

[2] O nome 'Aritmética de Frege' foi cunhado por Boolos em [4]; e o nome 'Teorema de Frege', cunhado em [5], p. 209. Boolos também cunhou o nome 'Princípio de Hume' em [4], p. 186, devido ao fato de Frege creditar o princípio ao filósofo escocês David Hume ao enunciá-lo em *Gl*, §63, citando a seguinte passagem de *A Treatise of Human Nature* (Livro I, Parte III, Seção I, §5): "When two numbers are so combin'd, as that the one has always a unite answering to every unite of the other, we pronounce them equal.".

> dos conceitos da lógica de segunda ordem. Se esse princípio explanatório, juntamente com 'definições implícitas' em geral, pode ser considerado como *analítico*, então isso deve ser suficiente ao menos para demonstrar a analiticidade da aritmética. ([41], p. 279)

Grande parte dos esforços de Hale e Wright para defender o *logicismo aritmético neo-fregeano* foi direcionada no sentido de fornecer uma elaboração da noção de *definição implícita* mencionada na passagem acima que, por um lado, acomode o Princípio de Hume e que, por outro, a descreva como um procedimento definicional que não envolva pressupostos epistemológicos significativos, de modo que definições implícitas pudessem ser adequadamente caracterizadas como verdades analíticas: sob o pressuposto de que a analiticidade é transmissível através da relação de consequência lógica e sob a tese, sustentada pelos neo-fregeanos, de que a relação de consequência lógica de segunda ordem é, de fato, uma relação de consequência *lógica*, o Teorema de Frege e a caracterização de *PH* como uma definição implícita implicam na analiticidade dos axiomas de Dedekind-Peano e, *a fortiori*, na analiticidade de todos os teoremas aritméticos que deles se seguem. Esquematicamente:

(P1) Teorema de Frege

(P2) *PH* é uma definição implícita do conceito de número cardinal e, portanto, uma verdade analítica.

(P3.1) A analiticidade é transmissível através da relação de consequência lógica.

(P3.2) A relação de consequência lógica de segunda ordem é uma relação de consequência *lógica*.

(P3) A analiticidade é transmissível através da relação de consequência lógica de segunda ordem.

(C) Os axiomas de Dedekind-Peano são verdades analíticas.

(C') Todos os teoremas aritméticos são verdades analíticas.

Embora o logicismo neo-fregeano tenha sido originalmente formulado para o caso particular da aritmética elementar ou teoria dos números naturais, Hale e Wright propuseram

uma extensão programática de suas teses para outras teorias matemáticas: o *programa logicista neo-fregeano* em filosofia da matemática prevê a fundamentação de demais teorias matemáticas em princípios que compartilham a mesma forma que *PH* e que, assim como *PH*, estão destinados a definir implicitamente os conceitos fundamentais dessas teorias:

> Frege acreditou, ao menos durante a maior parte de sua carreira, que as leis fundamentais da aritmética elementar – a teoria dos números naturais (cardinais finitos) – e da análise real são analíticas, no sentido explicado no *Grundlagen* §3 – ou seja, que elas poderiam ser provadas com base em leis lógicas gerais, juntamente com definições adequadas. Uma vez que sua defesa dessa tese dependia de se considerar que essas teorias concerniam um domínio de objetos que independentemente existentes (...) – os números cardinais finitos e os números reais – sua posição consistia em uma versão Platonista de logicismo. O [neo-fregeanismo] sustenta que ele estava substancialmente certo nessas duas componentes de sua filosofia, mas mantém uma visão mais otimista do que Frege de explicações contextuais de conceitos matemáticos fundamentais – tais como os conceitos de número cardinal e número real – mediante [princípios] hoje amplamente denominados princípios de abstração. ([23], p. 166–167)

Os *princípios de abstração* mencionados acima são fórmulas da forma

$$\forall\alpha\forall\beta(@\alpha = @\beta \leftrightarrow \alpha\ Equiv\ \beta)$$

em que α e β são variáveis do mesmo tipo, @ é um símbolo de função que se aplica a termos da mesma categoria sintática que essas variáveis para formar termos individuais; e *Equiv* é um símbolo de predicado binário destinado a denotar uma relação de equivalência entre os itens no campo de variação de α e β. O programa neo-fregeano de Hale e Wright propõe, portanto, que alguns desses princípios consistem em definições implícitas ou explicações contextuais de conceitos fundamentais de teorias matemáticas; e que algumas teorias formais – a serem denominadas *teorias fregeanas* nesta dissertação – resultantes da introdução de um ou mais princípios de abstração a um sistema dedutivo de *LSO* são adequadas para fornecer uma

fundamentação logicista de teorias matemáticas relevantes, no sentido de que elas são suficientes para permitir derivar os axiomas dessas teorias com base em definições logicamente corretas, estabelecendo a analiticidade desses axiomas e dos teoremas que deles se seguem (o argumento, em cada caso, sendo essencialmente o mesmo que o argumento a favor da analiticidade da aritmética elementar esquematizado acima)[3].

Com o intuito de executar esse programa, Hale, Wright e outros adeptos do logicismo neo-fregeano passaram a apresentar e discutir propostas concretas de fundamentações neo-fregeanas de outras duas teorias matemáticas, a saber, a teoria de conjuntos e a análise real: em [41], por exemplo, Wright considera a possilidade de fornecer um fundamentação neo-fregeana baseada em um princípio de abstração proposto por Boolos em [3] como uma alternativa para reparar a Lei Básica V, denominado *New V*[4]; e em [21], Hale propõe uma fundamentação da análise em uma teoria fregeana que estende **AF** mediante a introdução de três princípios de abstração adicionais, destinados a definir implicitamente os conceitos de *número inteiro*, *número racional* e *número real*, respectivamente.

Desde a sua proposição inicial em [40], o programa logicista neo-fregeano de Hale e

[3]Obviamente, nem todos os princípios de abstração podem ser considerados como definições implícitas pelos neo-fregeanos, uma vez que alguns deles são incompatíveis entre si ou resultam teorias inconsistentes quando introduzidos em um sistema dedutivo de *LSO* – a exemplo da seguinte versão simplificada da Lei Básica V:

$$\forall F \forall G(\varepsilon F = \varepsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx))$$

(em que εF denota a *extensão* do conceito F) a partir da qual o paradoxo de Russell é facilmente derivado na lógica de segunda ordem. O expediente de fornecer uma distinção bem motivada dos princípios de abstração adequados para definir implicitamente conceitos matemáticos relevantes dos princípios que não servem para esse próprosito é uma componente importante da agenda filosófica neo-fregeana e está relacionado à chamada *objeção da má companhia*, uma das objeções mais contundentes apresentadas contra o programa (para uma apresentação das diversas versões dessa objeção, ver [41]).

[4]Essa proposta foi, contudo, recusada posteriormente devido ao fato de a teoria fregeana baseada em *New V* ser insuficiente para provar todos os axiomas da teoria de conjuntos de Zermelo-Fraenkel e, principalmente, devido ao fato de esse princípio não satisfazer um dos critérios impostos por Wright para delimitar os princípios de abstração aceitáveis do ponto de vista neo-fregeano (ver nota 3).

Wright passou a fazer parte do debate contemporâneo em filosofia da matemática, envolvendo uma série de discussões acerca do status epistemológico do Princípio de Hume e do status epistemológico dos princípios de abstração e de definições implícitas, em geral – e também acerca da questão de saber se a lógica de segunda ordem deve ou não ser apropriadamente considerada como *lógica*. Esses debates despertaram o interesse pelo Teorema de Frege em alguns autores orientados para os problemas dessa disciplina: muitos trabalhos formais – mas na maior parte dos casos com alguma motivação filosófica – foram publicados sobre o tema no decorrer das décadas de 1980, 1990 e no início dos anos 2000, os quais trouxeram novos e importantes elementos para a avaliação das teses propostas pelos neo-fregeanos. Embora amplamente investigado, entretanto, nenhuma exposição inteiramente detalhada do teorema pode ser até hoje encontrada na literatura; muito provavelmente, porque as atenções daqueles autores estiveram, em grande medida, voltadas para questões históricas e conceituais ou para resultados técnicos mais avançados. Assim, nosso principal objetivo nesta dissertação é apresentar uma exposição autocontida e inteiramente sistemática do Teorema de Frege como um resultado técnico em lógica matemática. Nesse sentido, nos colocamos na posição do leitor mencionado por Wright no início da seção final de [40] (seção xix), na qual ele apresenta, pela primeira vez após Frege, a derivação dos axiomas de Dedekind-Peano a partir do Princípio de Hume:

> O propósito dessa seção final é esboçar como uma lógica de segunda ordem apropriada, suplementada por [*PH*] como um postulado, pode resultar provas dos cinco axiomas de [Dedekind-Peano]. Não tentarei estabelecer aqui a sintaxe ou a teoria da prova do tipo relevante de lógica; nem, *a fortiori*, oferecei derivações rigorosas e plenamente desenvolvidas dos axiomas de [Dedekind-Peano]. Meu propósito é meramente colocar carne suficiente nos ossos delineados por Frege nas seções 70–83 do *Grundlagen* de maneira a possibilitar ao leitor seguir mais facilmente e apreciar o brilhantismo da própria visão de Frege; e a possibilitar que ele resconstrua provas adequadas em detalhe, se desejar.

No Capítulo 1, argumentaremos que a lógica de segunda ordem, tal como usualmente apresentada em livros-texto de lógica matemática, é inapropriada para formalizar a linguagem Aritmética de Frege e, em geral, para formalizar as linguagens de quaisquer teorias fregeanas.

A partir desse diagnóstico, ensaiaremos então uma caracterização geral de linguagens que sejam adequadas para esse próposito, fornecendo, em seguida, uma semântica e um sistema dedutivo para essas linguagens. Também definiremos uma série de conceitos metateóricos e provaremos alguns metateoremas que permitirão relacionar a Aritmética de Frege com outras teorias de segunda ordem, os quais serão utilizados para formular precisamente o Teorema de Frege.

No Capítulo 2, apresentaremos uma prova detalhada do Teorema de Frege como relação entre a Aritmética de Frege e uma teoria de segunda ordem que formaliza os axiomas de Dedekind-Peano, a *Aritmética de Dedekind-Peano*. Com intuito de evidenciar que, mesmo quando desenvolvido em um sistema formal com profundas diferenças técnicas e conceituais em relação aos sistemas fregeanos, o teorema ainda assim pode ser justificadamente creditado a Frege, indicaremos, no decorrer do capítulo, as seções do *Grundlagen* onde as definições e teoremas em questão foram enunciados.

No Capítulo 3, trataremos do tema da consistência da Aritmética de Frege: apresentaremos a prova, devida a Boolos, da consistência de **AF** em relação a (a consistência da) *Aritmética de Peano de Segunda Ordem* e utilizaremos esse resultado e o Teorema de Frege para estabelecer a equiconsistência entre **AF**, a Aritmética de Dedekind-Peano e a *Aritmética de Peano de Segunda Ordem*.

Capítulo 1

Lógica de Segunda Ordem

Na Introdução, enunciamos o Teorema de Frege como o resultado segundo o qual os axiomas de Dedekind-Peano para a aritmética podem ser derivados na teoria formal resultante da introdução do Princípio de Hume a um sistema dedutivo de lógica de segunda ordem, a qual, seguindo Boolos, denominamos Aritmética de Frege (**AF**). Essa formulação do teorema e suas variantes similares são comumente encontradas na literatura sobre o tema. Na entrada da *Stanford Encyclopedia of Philosophy* sobre o Teorema de Frege, por exemplo, Edward Zalta afirma:

> O Teorema de Frege estabelece que os cinco axiomas de Dedekind-Peano para a teoria dos números podem ser derivados a partir do Princípio de Hume na lógica de segunda ordem [43].[1]

Mas o que se entende, nessas formulações, por 'lógica de segunda ordem'? Ao que tudo indica, os autores são unâmimes em considerar que se trata da *lógica de segunda ordem padrão*, tal como caracterizada em livros-texto de lógica matemática. Nessas caracterizações, grosso modo, linguagens de segunda ordem são usualmente descritas como sendo obtidas a partir de linguagens de primeira ordem pela introdução de variáveis para relações e, eventualmente, de variáveis para funções – denominadas *variáveis de segunda ordem* – e permitindo-se a quantificação sobre elas, o que significa que linguagens de segunda ordem são tratadas

[1] Outras caracterizações similares são encontradas em [6], [7], [9], [23], [41] e em [42].

como extensões de linguagens de primeira ordem apenas na medida em que incluem *símbolos lógicos* adicionais em seu vocabulário primitivo.

Apontamos que em virtude desse aspecto, entretanto, a lógica de segunda ordem padrão é inadequada para formalizar a Aritmética de Frege e, em particular, para formalizar o Princípio de Hume:

$$\forall F \forall G(\#F = \#G \leftrightarrow F \approx G)$$

Como descrevemos na Introdução, #, o operador de cardinalidade, é um símbolo de função que se aplica a variáveis de segunda ordem unárias, resultando termos, i.é., resultando expressões da mesma categoria sintática de variáveis e constantes individuais, de modo que, assim como Gx e $x = y$, $G\#F$, $\#F = y$ e $\#F = \#G$ são tratadas como fórmulas atômicas da linguagem da Aritmética de Frege. Mas seria possível descrever precisamente tal comportamento sintático de # com os recursos envolvidos na caracterização de linguagens de segunda ordem padrão? A resposta é não. Para constatá-lo, considere-se a definição usual do conjunto dos termos de uma linguagem de segunda ordem $\mathcal{L}$ sem variáveis para funções[2]:

Definição.

(i) Toda variável individual é um termo de $\mathcal{L}$.

(ii) Toda constante individual de $\mathcal{L}$ é um termo de $\mathcal{L}$.

(iii) Para $n > 0$, se f é um símbolo de função n-ário e $t_1, \ldots, t_n$ são de termos de $\mathcal{L}$, então $ft_1 \ldots t_n$ é um termo de $\mathcal{L}$.

(iv) Nada mais é termo de $\mathcal{L}$ (cláusula de fechamento).

Essa definição não acomoda expressões que, assim como $\#F$, resultam da aplicação de um símbolo de função a variáveis de segunda ordem, pois, de acordo com a cláusula (iii), para que esse fosse o caso, seria necessário que, além das variáveis individuais, as variáveis de

[2]Nesta dissertação, não trabalharemos com linguagens que possuam variáveis (de segunda ordem) para funções, embora o formalismo desenvolvido a seguir possa ser facilmente estendido para incluí-las.

segunda ordem também estivessem incluídas no conjunto dos termos de $\mathcal{L}$. Obviamente, a definição acima pode ser sempre extendida para cada caso particular, adicionando-se uma nova cláusula especificamente formulada para contemplar expressões daquele tipo; no caso da Aritmética de Frege, adicionando-se a cláusula (iii'):

(iii') Se F é uma variável de segunda ordem unária, então $\#F$ é um termo da linguagem de **AF**.

Essa é a maneira como o conjunto dos termos de **AF** é comumente definido, mas, embora tecnicamente correta, existem ao menos três razões pelas quais deveríamos evitá-la e que motivam a busca por uma formulação de uma definição alternativa de *linguagem de segunda ordem* que permita uma caracterização direta da linguagem da Aritmética de Frege:

Em primeiro lugar, notamos que, no decorrer do desenvolvimento da prova do Teorema de Frege em **AF**, introduzem-se símbolos definidos cujo comportamento também não pode ser descrito com os mesmos recursos envolvidos na caracterização de linguagens de segunda ordem padrão. O exemplo mais imediato é o símbolo $\approx$ para a relação de equinumericidade na formulação do Princípio de Hume acima, o qual se aplica a variáveis de segunda ordem unárias F e G, resultando fórmulas atômicas da forma $F \approx G$. Nesse caso, a definição das fórmulas atômicas de uma linguagem de segunda ordem padrão como expressões resultantes da concatenação de uma variável de segunda ordem ou de um símbolo de predicado n-ário a n termos não acomoda $F \approx G$, pois, novamente, F e G não são consideradas como termos dessa linguagem. Assim, para incluir $F \approx G$ entre as fórmulas de **AF** com o símbolo definido $\approx$, seria preciso estabelecer uma cláusula adicional também na definição conjunto das fórmulas da teoria. Exemplos como esse serão recorrentes quando apresentarmos em detalhe a prova do Teorema de Frege no próximo capítulo, o que evidencia que seria inviável, de um ponto de vista prático, adicionar, para cada símbolo definido de comportamento sintático "irregular", uma nova cláusula nas definições de *termo* ou de *fórmula*.

A segunda razão diz respeito à proposta neo-fregeana de fundamentar diferentes teorias matemáticas, além da própria aritmética elementar, em teorias fregeanas, i.é., teorias formais que, como vimos, resultam da introdução de um (ou mais) princípio de abstração

($P@$)

$$\forall\alpha\forall\beta(@\alpha = @\beta \leftrightarrow \alpha \; Equiv \; \beta)$$

a um sistema dedutivo de lógica de segunda ordem. Em $P@$, $@$ representa esquematicamente qualquer operador que se aplique a variáveis α e β, resultando termos, de modo que se α e β são variáveis de segunda ordem (de mesma aridade), expressões como $F@\alpha$, $@\alpha = x$ e $@\alpha = @\beta$ são fórmulas bem formadas da linguagem da teoria em questão. Portanto, as dificuldades apontadas acima não incidem especificamente sobre a Aritmética de Frege, mas também sobre as demais teorias formais desenvolvidas (ou a serem desenvolvidas) pelos neo-fregeanos em seu programa fundacionista em filosofia da matemática. Dessa forma, é certamente mais conveniente formular uma definição uniforme das linguagens dessas teorias do que definí-las, em cada caso particular, introduzindo-se cláusulas adicionais.

A terceira e última razão consiste no fato de que a definição particularizada acima da linguagem da Aritmética de Frege dificulta consideralmente a formulação técnica de teoremas metateóricos que a relacionam com outras teorias. Ao enunciarem-se teoremas desse tipo sobre, digamos, teorias K e K', é desejável que suas linguagens sejam fixadas dentro do mesmo arcabouço e que ambas sejam definidas tendo-se por base um sistema dedutivo comum (i.é., que elas compartilhem os mesmo axiomas e regras lógicos). Esses *desiderata* facilitam, por exemplo, a formulação de definições precisas dos conceitos de *sublinguagem*, *extensão*, *extensão definicional*, *interpretação* etc., utilizados para relacionar teorias distintas, e a confecção das provas de resultados envolvendo esses conceitos.

Em vista desse diagnóstico, pretendemos, neste capítulo, formular uma definição geral de *linguagem de segunda ordem* que possibilite uma caracterização direta da linguagem da Aritmética de Frege e de outras teorias fregeanas. Nossa proposta consiste em apresentar uma definição recorrendo a *tipos*, o que permitirá descrever *uniformemente* o comportamento sintático de #, $\approx$ e de outros símbolos definidos utilizados na prova do Teorema de Frege. A partir dessa definição, forneceremos então uma semântica e um sistema dedutivo apropriados para essas linguagens: o sistema dedutivo apresentado servirá para definirmos o conceito geral de *teoria de segunda ordem*, com o qual caracterizaremos **AF**, e para estabelecermos as regras

de inferência derivadas que serão empregadas na prova do teorema; a semântica, por sua vez, desempenhará um papel importante no Capítulo 3, quando abordarmos o tema da consistência dessa teoria.

Por fim, alertamos o leitor que não intentamos, como nossas definições, reproduzir a linguagem e o sistema formais apresentados por Frege no *Bregriffsshrift* ou no *Grundgesetze der Arithmetik*, utilizando, para isso, a notação lógica moderna. Muito pelo contrário, nossas caracterizações da linguagem de **AF** e do sistema dedutivo subjacente apresentam diferenças fundamentais em relação às linguagens e sistemas formulados em ambos os livros. Uma delas, por exemplo, diz respeito ao fato de tratarmos sentenças e termos como expressões de categorias sintáticas distintas, ao passo que no *Grundgesetze*, Frege considerava as expressões de ambos os tipos como nomes próprios de indivíduos ou objetos – sendo que sentenças se distinguem dos demais termos por denotarem apenas os valores de verdade *o verdadeiro* e *o falso*[3]. Em nossa proposta de formalização, portanto, não estaremos concernidos com os inúmeros problemas interpretativos envolvidos na compreensão das linguagens e sistemas formais elaborados por Frege.

1.1 Sintaxe

Como notamos acima, a dificuldade em caracterizar, com base em uma definição indutiva, expressões formadas pelo operador # como termos de uma linguagem de segunda ordem padrão $\mathcal{L}$ deriva do fato de as variáveis de segunda ordem não estarem inicialmente incluídas no conjunto dos termos de $\mathcal{L}$ (e, pela mesma razão, não é possível caracterizar expressões da forma $F \approx G$ como fórmulas atômicas). Esse diagnóstico sugere que uma possível alternativa para evitarmos aquela dificuldade consistiria em generalizar a definição de *termo* para contemplar essas variáveis. No entanto, não poderíamos simplesmente incluí-las no conjunto dos termos de $\mathcal{L}$, substituindo a cláusula (i) da definição acima por:

(i') Todas as variáveis de primeira e de segunda ordem são termos de $\mathcal{L}$.

[3]Por essa razão, a lógica fregeana no *Grundgesetze* é algumas vezes denominada uma *lógica de termos*.

pois, nesse caso, além de $\#F, \#G, \ldots$, incluiríamos também expressões da forma $\#x, \#y, \ldots, \#\#x$, $\#\#y, \ldots, \#\#F, \#\#G, \ldots, \#R, \#S, \ldots$ entre os termos de $\mathcal{L}$.

Alternativamente, nossa estratégia consiste em generalizar a definição do conceito de *termo*, estratificando-o, porém, em conjuntos de termos que possuam um determinado *tipo*, de modo a identificar os termos da linguagem aos quais o operador de cardinalidade pode ou não ser aplicado: atribuiremos tipos distintos às variáveis de primeira e de segunda ordem – e também às variáveis de segunda ordem de diferentes aridades –, descrevendo # como um símbolo de função que se aplica unicamente a termos que possuam o mesmo tipo que as variáveis de segunda ordem unárias e que resulta termos do mesmo tipo das variáveis individuais. Esse expediente permitirá então incluir $\#F, \#G, \ldots$ entre os termos da linguagem e excluir $\#x, \#y, \ldots, \#R, \#S, \ldots$ como expressões mal formadas; e servirá para descrever de maneira similar o comportamento sintático de símbolos de predicado como $\approx, Her, In$, definidos em **AF** no Capítulo 2.

Diferentemente das definições usuais de linguagens tipadas (como em [11]), que atribuem tipos a todas as expressões bem formadas, incluindo-se as fórmulas, em nossa caracterização das linguagens de segunda ordem a serem consideradas nesta dissertação, restringiremos a atribuição de tipos unicamente aos termos, definindo o conjunto das fórmulas de maneira similar à definição correspondente na lógica de primeira ordem. Para atribuir tipos aos termos de uma linguagem $\mathcal{L}$ *indutivamente*, começaremos por associar um único tipo específico a cada símbolo não lógico de $\mathcal{L}$ e a todas as variáveis. Os tipos associados às variáveis serão mantidos fixos para todas as linguagens (por se tratarem de símbolos lógicos); e os tipos associados aos símbolos não lógicos, bem como os próprios símbolos não lógicos, dependerão das especificidades de cada linguagem particular[4].

[4]O recurso a tipos para formalizar teorias fregeanas na lógica de segunda ordem nos foi pela primeira vez sugerido pela seguinte passagem de Boolos em [3] (p. 172), na qual ele comenta sobre a definição fregeana de *número cardinal* no *Grundlagen* (*Gl*, §68):

> The sophisticated definition of numbers as extensions of certain concepts of concepts and extensive use of binary relations found in the *Grundlagen* are evidence that Frege was there committed to the existence of objects of all finite types: "objects", the items of the lowest type 0, and, for any

Para os nossos propósitos, será suficiente considerar um tipo como 0 ou uma sequência finita não vazia de tipos – $\langle 0\rangle, \langle 0,0\rangle, \langle 0,0,0\rangle, \ldots, \langle\langle 0\rangle\rangle, \langle 0,\langle 0\rangle\rangle, \langle\langle 0\rangle,0\rangle, \langle\langle 0\rangle,\langle 0\rangle\rangle, \langle 0,0,\langle 0\rangle\rangle, \ldots,$ $\langle\langle\langle 0\rangle\rangle\rangle, \langle 0,\langle\langle 0\rangle\rangle\rangle, \langle\langle\langle 0\rangle\rangle,0\rangle, \ldots$ – limitando os tipos que efetivamente utilizaremos a uma determinada *ordem*:

Definição 1.1. O conjunto dos *tipos*, denotado por $\mathcal{T}$, é definido indutivamente por

(i) $0 \in \mathcal{T}$.

(ii) Para $n > 0$, se $\tau_1, \ldots, \tau_n \in \mathcal{T}$, então $\langle \tau_1, \ldots, \tau_n\rangle \in \mathcal{T}$.

Definição 1.2. Definimos a função *ordem* $Ord : \mathcal{T} \to N$ recursivamente por

(i) $Ord(0) = 1$.

(ii) Para $n > 0$, $Ord(\langle \tau_1, \ldots, \tau_n\rangle) = \max\{Ord(\tau_1), \ldots, Ord(\tau_n)\} + 1$.

Exemplo.

1. $Ord(\langle 0\rangle) = \max\{Ord(0)\} + 1 = \max\{1\} + 1 = 2$.
2. $Ord(\langle\langle 0\rangle, 0\rangle) = \max\{Ord(\langle 0\rangle), Ord(0)\} + 1 = \max\{2,1\} + 1 = 3$.
3. $Ord(\langle\langle 0\rangle, \langle 0\rangle\rangle) = \max\{Ord(\langle 0\rangle), Ord(\langle 0\rangle)\} + 1 = \max\{2,2\} + 1 = 3$.

Para cada $n \in N$, seja $\mathcal{T}_n = \{\tau \in \mathcal{T} : Ord(\tau) \leq n\}$ o conjunto de *tipos de no máximo ordem n*: $\mathcal{T}_0 = \varnothing$, $\mathcal{T}_1 = \{0\}$, $\mathcal{T}_2 = \{0, \langle 0\rangle, \langle 0,0\rangle, \langle 0,0,0\rangle, \ldots\}, \ldots$. No nosso caso, não necessitaremos de tipos de ordem maior do que três e, por isso, convencionaremos escrever $\mathcal{T}$ no lugar de $\mathcal{T}_3$, de agora em diante; $\mathcal{T}_3$ sendo parcialmente descrito por

$$\mathcal{T}_2 \cup \{\langle\langle 0\rangle\rangle, \langle 0,\langle 0\rangle\rangle, \langle\langle 0\rangle,0\rangle, \langle\langle 0\rangle,\langle 0\rangle\rangle, \langle 0,0,\langle 0\rangle\rangle, \ldots, \langle\langle 0,0\rangle\rangle, \langle 0,\langle 0,0\rangle\rangle, \ldots, \langle\langle 0,0\rangle,\langle 0,0\rangle\rangle, \ldots\}$$

Ao definir uma teoria de segunda ordem como a Aritmética de Frege, é desejável estabelecer sua linguagem da maneira mais econômica possível. Para isso, fixaremos de antemão os

types $[\tau_1, \ldots, \tau_n]$, relations of type $[\langle \tau_1, \ldots, \tau_n\rangle]$ among items of types $[\tau_1, \ldots, \tau_n]$ [...].

No entanto, em seus trabalhos sobre o tema, Boolos nunca empregou efetivamente uma linguagem tipada para formalizar a Aritmética de Frege.

símbolos lógicos e de pontuação que utilizaremos (Definição 1.5), sendo necessário indicar, em cada caso, os símbolos não lógicos da teoria e associar a cada um deles um tipo específico que determinará o seu comportamento sintático. Todo esse expediente pode ser realizado especificando-se unicamente a assinatura da teoria em questão:

Definição 1.3. Uma *assinatura* (*de segunda ordem*) Σ é um par ordenado $\langle \mathcal{R}, \mathcal{F} \rangle$ tal que

(i) $\mathcal{R} = \langle \mathcal{R}_\tau \rangle_{\tau \in \mathcal{T}}$ é uma família de conjuntos disjuntos.

(ii) $\mathcal{F} = \langle \mathcal{F}_\tau \rangle_{\tau \in \mathcal{T}}$ é uma família de conjuntos disjuntos tal que $\mathcal{F}_0 = \varnothing$ e $\mathcal{F}_{\langle \tau \rangle} = \varnothing$, para todo $\tau \in \mathcal{T}_2$.

(iii) $(\bigcup_{\tau \in \mathcal{T}} \mathcal{R}_\tau) \cap (\bigcup_{\tau \in \mathcal{T}} \mathcal{F}_\tau) = \varnothing$[5].

Os elementos de $\mathcal{R}_0$ serão denominados *constantes individuais*. Para cada $n > 0$, os elementos de $\mathcal{R}_{\langle \tau_1, \ldots, \tau_n \rangle}$ serão denominados *contantes relacionais de tipo* $\langle \tau_1, \ldots, \tau_n \rangle$ ou *constantes relacionais n-árias*, e os elementos de $\mathcal{F}_{\langle \tau_1, \ldots, \tau_n, \tau \rangle}$, denominados *constantes funcionais de tipo* $\langle \tau_1, \ldots, \tau_n, \tau \rangle$ ou *constantes funcionais n-árias*.

Na definição de assinatura, a atribuição de um tipo $\langle \tau_1, \ldots, \tau_n \rangle$ a uma constante relacional R ($R \in \mathcal{R}_{\langle \tau_1, \ldots \tau_n \rangle}$) servirá para especificar quantos e quais termos podem ser concatenados a R para formar uma fórmula atômica: R resulta uma fórmula apenas se concatenado a n termos, cada um dos quais é de tipo $\tau_1, \ldots, \tau_n$, exatamente nessa ordem (Definição 1.8). Similarmente, a atribuição de $\langle \tau_1, \ldots, \tau_n, \tau \rangle$ a uma constante funcional f ($f \in \mathcal{F}_{\langle \tau_1, \ldots, \tau_n, \tau \rangle}$), servirá especificar quantos e quais termos formam um novo termo quando concatenados a f e qual o tipo do termo assim obtido: f resulta um termo de tipo τ, quando concatenado a n termos de tipo $\tau_1, \ldots, \tau_n$, exatamente nessa ordem (Definição 1.7). Como, ao atribuir um tipo a uma constante funcional, é preciso determinar não apenas a quais termos ela é aplicável, mas também o tipo do termo resultante, impede-se, através das condições $\mathcal{F}_0 = \varnothing$ e $\mathcal{F}_{\langle \tau \rangle} = \varnothing$ em (ii), a atribuição de tipos de comprimento menor do que 2 a qualquer constante funcional.

Exemplo.

[5]Como uma cláusula adicional não explicitada, assumimos que nenhum símbolo lógico ou símbolo de pontuação (Definição 1.5) pertence a $\bigcup_{\tau \in \mathcal{T}} \mathcal{R}_\tau$ ou a $\bigcup_{\tau \in \mathcal{T}} \mathcal{F}_\tau$.

1. A assinatura da Aritmética de Frege é o par $\Sigma_{\mathbf{F}} = \langle \mathcal{R}^{\mathbf{F}}, \mathcal{F}^{\mathbf{F}} \rangle$ tal que

(i) Para todo τ, $\mathcal{R}^{\mathbf{F}}_{\tau} = \varnothing$.

(ii) $\mathcal{F}^{\mathbf{F}}_{\langle\langle 0\rangle,0\rangle} = \{\#\}$.

(iii) Para todo $\tau \neq \langle\langle 0\rangle, 0\rangle$, $\mathcal{F}^{\mathbf{F}}_{\tau} = \varnothing$.

2. A assinatura da Aritmética de Frege com o símbolo $\approx$ é o par $\Sigma_{\mathbf{F}^{\approx}} = \langle \mathcal{R}^{\mathbf{F}^{\approx}}, \mathcal{F}^{\mathbf{F}^{\approx}} \rangle$ tal que

(i) $\mathcal{R}^{\mathbf{F}^{\approx}}_{\langle\langle 0\rangle,\langle 0\rangle\rangle} = \{\approx\}$.

(ii) Para todo $\tau \neq \langle\langle 0\rangle, \langle 0\rangle\rangle$, $\mathcal{R}^{\mathbf{F}^{\approx}}_{\tau} = \varnothing$.

(iii) $\mathcal{F}^{\mathbf{F}^{\approx}}_{\langle\langle 0\rangle,0\rangle} = \{\#\}$.

(iv) Para todo $\tau \neq \langle\langle 0\rangle, 0\rangle$, $\mathcal{F}^{\mathbf{F}^{\approx}}_{\tau} = \varnothing$.

$\Sigma_{\mathbf{F}}$ não possui, portanto, constantes individuais ou relacionais, mas apenas a constante funcional unária # de tipo $\langle\langle 0\rangle, 0\rangle$: # se aplica a termos de tipo $\langle 0\rangle$, resultando termos de tipo 0. Além de #, $\Sigma_{\mathbf{F}^{\approx}}$ possui a constante relacional binária $\approx$ de tipo $\langle\langle 0\rangle, \langle 0\rangle\rangle$, a qual resulta uma fórmula atômica, quando concatenada a termos de tipo $\langle 0\rangle$.

A partir dos exemplos acima e da descrição informal do operador # na Introdução, é fácil presumir que o tipo 0 será atribuído às variáveis individuais e que $\langle 0\rangle$ será atribuído às variáveis de segunda ordem unárias. De maneira geral, se $\mathbf{v}$ é uma variável de segunda ordem de n-ária, atribuir-se-á o tipo $\langle \tau_1, \ldots, \tau_n \rangle$ a $\mathbf{v}$, em que $\tau_i = 0$, para $1 \leq i \leq n$:

Definição 1.4. Seja $\mathcal{V} = \langle \mathcal{V}_\tau \rangle_{\tau \in \mathcal{T}_2}$ uma família de conjuntos tal que, para cada $\tau \in \mathcal{T}_2$, $\mathcal{V}_\tau = \{v^\tau_i : i \in N\}$ é o conjunto das *variáveis de tipo* τ; e seja $\mathbf{V} = \bigcup_{\tau \in \mathcal{T}_2} \mathcal{V}_\tau$ o conjunto de todas as *variáveis*.

Note-se que, como os elementos de $\mathcal{T}_2$ são 0 e todas as sequências finitas não vazias de 0s, o tipo de qualquer variável possui no máximo ordem 2 e diremos, em acordo com a nomeclatura usual da *LSO*, que $\mathcal{V}_0$ é o conjunto das *variáveis de primeira ordem* (ou *variáveis individuais*) e que $\mathcal{V}_{\langle \tau_1, \ldots, \tau_n \rangle}$ (i.é., $\mathcal{V}_{\langle 0, \ldots, 0 \rangle}$) é o conjunto das *variáveis de segunda ordem n-árias* (ou *variáveis relacionais n-árias*). Além disso, convencionaremos escrever $v_0, v_1, \ldots$ no lugar de $v^0_0, v^0_1, \ldots$; e $V^n_0, V^n_1, \ldots$ no lugar de $v^{\langle \tau_1, \ldots, \tau_n \rangle}_0, v^{\langle \tau_1, \ldots, \tau_n \rangle}_1, \ldots$, utilizando $\mathbf{v}^\tau, \mathbf{v}^\tau_1, \mathbf{v}^\tau_2, \ldots$ como

metavariáveis de tipo τ (metavariáveis que varrem $\mathcal{V}_\tau$) e $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \ldots$ como metavariáveis em geral (metavariáveis que varrem $\mathbf{V}$).

Fixaremos os *símbolos lógicos* das linguagens das teorias com as quais trabalharemos como as variáveis de primeira e segunda ordem, os *conectivos* $\neg$ (*símbolo de negação*) e $\rightarrow$ (*símbolo de implicação material*), o *símbolo de quantifição universal* $\forall$, e o *símbolo de identidade* $=$. Os *símbolos auxiliares* ou *símbolos de pontuação* serão apenas $($ (*parênteses esquerdo*) e $)$ (*parênteses direito*).

Definição 1.5. Seja Σ uma assinatura de segunda ordem. A *linguagem gerada por* Σ (ou simplesmente a *linguagem de* Σ), denotada por $\mathcal{L}(\Sigma)$, é a upla ordenada

$$\langle \Sigma, \mathcal{V}, \neg, \rightarrow, \forall, =, (,) \rangle$$

Além de atender ao ideal de economia, a definição de assinatura fornece um arcabouço geral no qual as linguagens de teorias de segunda ordem distintas podem ser relacionadas entre si. Em particular, podemos estabelecer precisamente quando a linguagem de uma dada teoria está incluída na linguagem de uma segunda: se $\Sigma = \langle \mathcal{R}, \mathcal{F} \rangle$ e $\Sigma' = \langle \mathcal{R}', \mathcal{F}' \rangle$ são assinaturas tais que $\mathcal{R}_\tau \subseteq \mathcal{R}'_\tau$ e $\mathcal{F}_\tau \subseteq \mathcal{F}'_\tau$, para todo $\tau \in \mathcal{T}$, escreveremos $\Sigma \subseteq \Sigma'$ e diremos que Σ é uma *subassinatura* de Σ' ou que $\mathcal{L}(\Sigma)$ é uma *sublinguagem* de $\mathcal{L}(\Sigma')$. Nos exemplos acima, temos, portanto, que $\Sigma_{\mathbf{F}} \subseteq \Sigma_{\mathbf{F}^\approx}$.

Definição 1.6. Seja $\Sigma = \langle \mathcal{R}, \mathcal{F} \rangle$ uma assinatura. O conjunto de Σ-*parâmetros*, denotado por $[\Sigma]$, é definido por

$$[\Sigma] = \{\forall\} \cup \bigcup_{\tau \in \mathcal{T}} \mathcal{R}_\tau \cup \bigcup_{\tau \in \mathcal{T}} \mathcal{F}_\tau$$

Assim, $[\Sigma_{\mathbf{F}}] = \{\forall, \#\}$ e $[\Sigma_{\mathbf{F}^\approx}] = \{\forall, \approx, \#\}$.

Com a atribuição inicial de tipos aos símbolos não lógicos de uma linguagem, através da especificação de uma assinatura, e também às variáveis de primeira e de segunda ordem, definimos indutivamente o conjunto de *termos de tipo* τ da linguagem, para cada $\tau \in \mathcal{T}$:

Definição 1.7. Seja $\Sigma = \langle \mathcal{R}, \mathcal{F} \rangle$ uma assinatura de segunda ordem. Definimos o conjunto de Σ-*termos de tipo* τ (denotado por $\mathrm{T}(\Sigma)_\tau$) por indução, simultâneamente para todo $\tau \in \mathcal{T}$:

(i) Para todo $\tau \in \mathcal{T}_2$, $\mathcal{V}_\tau \subseteq \mathrm{T}(\Sigma)_\tau$.

(ii) Para todo $\tau \in \mathcal{T}$, $\mathcal{R}_\tau \subseteq \mathrm{T}(\Sigma)_\tau$.

(iii) Para $1 \leq i \leq n$, se $t_i \in \mathrm{T}(\Sigma)_{\tau_i}$ e $f \in \mathcal{F}_{\langle \tau_1, \ldots, \tau_n, \tau \rangle}$, então $f t_1 \ldots t_n \in \mathrm{T}(\Sigma)_\tau$.

O conjunto de todos os Σ-*termos* (denotado por $\mathbf{T}(\Sigma)$) é definido por $\mathbf{T}(\Sigma) = \bigcup_{\tau \in \mathcal{T}} \mathrm{T}(\Sigma)_\tau$.

Na Definição 1.7, a cláusula (i) identifica as variáveis individuais e as variáveis relacionais n-árias como termos de tipo 0 e $\langle 0, \ldots, 0 \rangle$, respectivamente; e de maneira similar, a cláusula (ii) identifica as constantes individuais e as constantes relacionais de tipo $\langle \tau_1, \ldots, \tau_n \rangle$ como termos de tipo 0 e $\langle \tau_1, \ldots, \tau_n \rangle$, respectivamente. A cláusula (iii), a cláusula indutiva, utiliza a atribuição de tipos às constantes funcionais de uma dada linguagem para determinar como termos complexos podem ser formados por essas constantes: se à constante funcional f é atribuído o tipo $\langle \tau_1, \ldots, \tau_n, \tau \rangle$, sabemos, em primero lugar, que um termo é formado por f apenas se ela for concatenada a n outros termos; em segundo lugar, que esses termos têm de ser de tipo $\tau_1, \ldots, \tau_n$, exatamente nessa ordem; e, por fim, que o termo resultante dessa concatenação é de tipo τ, i.é, que ele pertence a $\mathrm{T}(\Sigma)_\tau$. Note-se que, diferentemente das constantes relacionais, as constantes funcionais não são consideradas como termos, mas a atribuição de tipos a esses símbolos na assinatura é fundamental para descrever o seu comportamento sintático. Note-se, ademais, que se $\Sigma \subseteq \Sigma'$, então $\mathrm{T}(\Sigma)_\tau \subseteq \mathrm{T}(\Sigma')_\tau$, para todo $\tau \in \mathcal{T}$, e que, portanto, $\mathbf{T}(\Sigma) \subseteq \mathbf{T}(\Sigma')$, se $\Sigma \subseteq \Sigma'$.

Para recuperar o sentido original de *termo* da lógica de primeira ordem, iremos nos referir aos elementos de $\mathrm{T}(\Sigma)_0$ por *termos de primeira ordem* ou *termos individuais*; e, em geral, aos elementos de $\mathrm{T}(\Sigma)_\tau$ por *termos de Ord*(τ)*-ésima ordem*. Assim, os termos individuais de $\Sigma_{\mathbf{F}}$ são as variáveis de primeira ordem e expressões formadas pela aplicação do operador de cardinalidade a termos de tipo $\langle 0 \rangle$ (termos em $\mathrm{T}(\Sigma_{\mathbf{F}})_{\langle 0 \rangle}$) – *e.g.*, $\#V_0^1, \#V_1^1, \#V_2^1, \ldots$ – e os termos de segunda ordem são apenas as variáveis relacionais. $\Sigma_{\mathbf{F}}$ não possui termos de terceira ordem e $\approx$ é o único termo de terceira ordem de $\Sigma_{\mathbf{F}^\approx}$.

Como anunciamos acima, diferentemente da definição do conjunto dos termos de uma linguagem de segunda ordem, a definição do conjunto das fórmulas será essencialmente a mesma que a definição correspondente na lógica de primeira ordem; a única diferença consiste apenas na maneira como as fórmulas atômicas são caracterizadas:

Definição 1.8. Seja $\Sigma = \langle \mathcal{R}, \mathcal{F} \rangle$ uma assinatura de segunda ordem. Definimos o conjunto de Σ-*fórmulas* (denotado por $\mathrm{F}(\Sigma)$) indutivamente por

(i) Para $1 \leq i \leq n$, se $t_i \in \mathrm{T}(\Sigma)_{\tau_i}$ e $t \in \mathrm{T}(\Sigma)_{\langle \tau_1, \ldots, \tau_n \rangle}$, então $tt_1 \ldots t_n \in \mathrm{F}(\Sigma)$.

(ii) Se $t_1, t_2 \in \mathrm{T}(\Sigma)_0$, então $t_1 = t_2 \in \mathrm{F}(\Sigma)$.

(iii) Se $A \in \mathrm{F}(\Sigma)$, então $(\neg A) \in \mathrm{F}(\Sigma)$.

(iv) Se $A, B \in \mathrm{F}(\Sigma)$, então $(A \to B) \in \mathrm{F}(\Sigma)$.

(v) Se $A \in \mathrm{F}(\Sigma)$ e $\mathbf{v} \in \mathbf{V}$, então $\forall \mathbf{v} A \in \mathrm{F}(\Sigma)$.

A cláusula (i) acima utiliza a atribuição de tipos aos termos da linguagem, mediante a Definição 1.7, para caracterizar as fórmulas atômicas nas quais o símbolo de identidade não ocorra: se $t \in \mathrm{T}(\Sigma)_{\langle \tau_1, \ldots, \tau_n \rangle}$, ou seja, se t é um termo de tipo $\langle \tau_1, \ldots, \tau_n \rangle$, sabemos que ele resultará uma fórmula atômica apenas se concatenado a n termos e se cada um deles for de tipo $\tau_1, \ldots, \tau_n$, exatamente nessa ordem: ao atribuir o tipo $\langle \langle 0 \rangle, \langle 0 \rangle \rangle$ a $\approx$ na definição de $\Sigma_{\mathbf{F}^{\approx}}$, estabelecemos que fórmulas são formadas por essa constante somente se ela for concatenada a exatamente dois termos em $\mathrm{T}(\Sigma_{\mathbf{F}^{\approx}})_{\langle 0 \rangle}$. A cláusula (ii) estabelece que o símbolo de identidade se aplica unicamente a termos individuais para formar fórmulas atômicas e as cláusulas restantes são exatamente as mesmas que na definição correspondente na lógica de primeira ordem com os símbolos $\neg, \to, \forall$ tomados como primitivos – com a diferença de que, em (iv), $\mathbf{v}$ representa qualquer variável de primeira ou de segunda ordem. Note-se que, assim como no caso dos termos, se $\Sigma \subseteq \Sigma$, então $\mathrm{F}(\Sigma) \subseteq \mathrm{F}(\Sigma')$.

De agora em diante, escreveremos $t_1 \neq t_2$ no lugar de $(\neg\, t_1 = t_2)$; e eventualmente escreveremos $t(t_1, \ldots, t_n)$ no lugar de $tt_1 \ldots t_n$ e $t_1 t t_2$ no lugar de $tt_1 t_2$. Como exemplos de $\Sigma_{\mathbf{F}}$-fórmulas temos: $V_0 v_1$, $V_0 v_1 v_2$, $V_1 \# V_0$ (ou $V_1(\# V_0)$), $V_5 v_1 \# V_0$ (ou $V_5(v_1, \# V_0)$), $\# V_5 = v_{10}$,

$(V_0v_1v_1v_1 \to (\neg V_0v_1))$, $(\neg\forall v_1(\neg\ \#V_0 = v_1))$ (ou $(\neg\forall v_1(\#V_0 \neq v_1))$); e, como exemplos de $\Sigma_{\mathbf{F}^{\approx}}$-fórmulas, temos $\approx V_0V_1$ (ou $V_0 \approx V_1$) e $\forall V_5\ V_5 \approx V_5$.

Adotaremos, a seguir, as convenções usuais relativas à omissão de parênteses em uma fórmula e dos índices sobrescritos das variáveis de segunda ordem (sempre que a aridade da variável for clara pelo contexto), escrevendo, por exemplo, $\neg\neg V_0v_1$ no lugar de $(\neg(\neg V_0^1v_1))$ e $V_0v_1v_2 \to (V_5v_1 \to \neg V_5v_2)$ no lugar de $(V_0^2v_1v_2 \to (V_5^1v_1 \to (\neg V_5^1v_2)))$. Os conectivos $\vee$, $\wedge$, $\leftrightarrow$ e o *símbolo de quantificação existencial* $\exists$ serão introduzidos mediante as seguintes abreviações, assumindo que as convenções apropriadas sobre a omissão dos parênteses também se aplicam a esses casos:

$$A \vee B := \neg A \to B$$
$$A \wedge B := \neg(A \to \neg B)$$
$$A \leftrightarrow B := A \to B \wedge B \to A$$
$$\exists \mathbf{v} A := \neg\forall \mathbf{v}\neg A$$

A expressão $\forall\mathbf{v}$ é denominada *quantificador universal* e diremos que $\forall\mathbf{v}$ *liga* a variável $\mathbf{v}$. Na fórmula $\forall\mathbf{v}A$, A é o *escopo do quantificador* $\forall\mathbf{v}$. Uma ocorrência de uma variável em uma fórmula é dita *ligada* se ela ocorre em um quantificador universal ou no escopo de um quantificador universal ligando aquela variável; caso contrário, ela é dita *livre*. Uma variável é dita *livre/ligada* em uma fórmula se ao menos uma ocorrência da variável nessa fórmula for livre/ligada. Assim, V_0 é livre e ligada em $\forall V_0\forall v_1V_0v_1 \to V_0v_1$. Se uma fórmula não contém ocorrências livres de qualquer variável, ela é denominada *fórmula fechada* ou *sentença*.

Se $\mathbf{v}^\tau$ é uma variável, t um termo de tipo τ e A uma fórmula, diremos que uma *ocorrência* de $\mathbf{v}^\tau$ em A é *substituível por* t se ela é ligada ou se ela não está no escopo de um quantificador que liga uma variável em t, e diremos que $\mathbf{v}^\tau$ é *substituível por* t em A se $\mathbf{v}^\tau$ não ocorre em A ou se todas as ocorrências de $\mathbf{v}^\tau$ em A são substituíveis por t. A primeira ocorrência de v_2 em $V_5v_1v_2 \to \forall V_0\neg V_0v_2$ é, por exemplo, substituível por $\#V_0$, embora a segunda não o seja (portanto, v_2 não é substituível por $\#V_0$ em $V_5v_1v_2 \to \forall V_0\neg V_0v_2$)[6].

[6]Nessa definição, consideramos que uma ocorrência ligada de $\mathbf{v}^\tau$ em A é substituível por t para garantir que $\mathbf{v}^\tau$ seja sempre substituível por si própria, mesmo que ela seja ligada em A. Assim, diremos que v_1 é substituível

1.2 Semântica

Uma vez estabelecida a sintaxe da *LSO* de uma maneira apropriada para caracterizar a linguagem da Aritmética de Frege e de outras teorias fregeanas, as definições dos conceitos semânticos de *estrutura*, *denotação*, *satisfação*, *verdade*, *validade*, *modelo* e *consequência lógica* não oferecem dificuldade substancial, sendo, em certa medida, mimetizadas das definições correspondentes para a lógica de segunda ordem padrão. No entanto, diferentemente da lógica (clássica) de primeira ordem, a lógica de segunda ordem padrão possui semânticas alternativas a depender do significado atribuído aos quantificadores de segunda ordem (quantificadores que ligam variáveis de segunda ordem)[7]. Essas diferentes semânticas também poderiam ser replicadas para as linguagens caracterizadas acima, mas em virtude do fato de considerações semânticas desempenharem apenas um papel auxiliar no Capítulo 3, não abordaremos aqui essas distinções, apresentando, como *a* semântica daquelas linguagens, uma adaptação da chamada *semântica padrão*, na qual as variáveis de segunda ordem varrem todas as relações construídas sobre o domínio de uma dada estrutura.

Na definição de uma estrutura para as linguagens de segunda ordem tipadas caracterizadas acima, será preciso considerar um domínio para cada τ em $\mathcal{T}$, no qual os itens de tipo τ da linguagem em questão serão interpretados: a partir de um domínio inicial fixado pela estrutura, definiremos *o domínio de tipo* τ (da estrutura) fazendo uso da seguinte definição preliminar:

Definição 1.9. Seja X um conjunto. Para cada $\tau \in \mathcal{T}$, definimos o conjunto X_τ recursivamente por:

(i) $X_0 = X$.

(ii) $X_{\langle \tau_1, \ldots, \tau_n \rangle} = \mathcal{P}\left(X_{\tau_1} \times \cdots \times X_{\tau_n}\right)$.

Exemplo. Se $X = X_0 = \{a, b\}$, então

por v_1 em $\forall v_1 V_5 v_1 \rightarrow V_{10} v_1$.

[7] Para uma apresentação das diferentes semânticas alternativas para lógica de segunda ordem padrão, ver [34], Capítulo 3.

1. $X_{\langle 0\rangle} = \mathcal{P}(\{a,b\}) = \{\varnothing, \{a\}, \{b\}, \{a,b\}\}$

2. $X_{\langle 0,0\rangle} = \mathcal{P}(\{a,b\} \times \{a,b\}) =$

$$\{\varnothing, \{\langle a,a\rangle\}, \ldots, \{\langle a,a,\rangle\ , \langle a,b\rangle\}, \ldots, \{\langle a,a\rangle, \langle a,b\rangle, \langle b,a\rangle\}, \ldots, \{\langle a,a\rangle, \langle a,b\rangle, \langle b,a\rangle, \langle b,b\rangle\}\}$$

3. $X_{\langle\langle 0\rangle\rangle} = \mathcal{P}\mathcal{P}(\{a,b\}) =$

$$\{\varnothing, \{\varnothing\}, \{\{a\}\}, \{\{b\}\}, \ldots, \{\varnothing, \{a\}\}, \ldots, \{\varnothing, \{a\}, \{a,b\}\}, \ldots, \{\varnothing, \{a\}, \{b\}, \{a,b\}\}\}$$

Uma estrutura para uma linguagem particular será então caracterizada como uma função que especifica um conjunto não vazio como o domínio (inicial) da estrutura e que associa a cada símbolo não lógico da linguagem um item (conjuntístico) construído sobre esse domínio, de acordo com a definição acima:

Definição 1.10. Uma Σ-*estrutura* $\mathfrak{A}$ é uma função no cojunto $[\Sigma]$ de Σ-parâmetros tal que

(i) $\mathfrak{A}(\forall)$ é um conjunto não vazio $|\mathfrak{A}|$ (denominado o *domínio* (*inicial*) de $\mathfrak{A}$).

(ii) Para todo $\tau \in \mathcal{T}$, para toda $R \in \mathcal{R}_\tau$, $\mathfrak{A}(R) \in |\mathfrak{A}|_\tau$.

(iii) Para todo $\tau = \langle \tau_1, \ldots, \tau_n, \tau\rangle \in \mathcal{T}$, para toda $f \in \mathcal{F}_\tau$, $\mathfrak{A}(f)$ é uma função

$$\mathfrak{A}(f) : |\mathfrak{A}|_{\tau_1} \times \cdots \times |\mathfrak{A}|_{\tau_n} \to |\mathfrak{A}|_\tau$$

de $|\mathfrak{A}|_{\tau_1} \times \cdots \times |\mathfrak{A}|_{\tau_n}$ em $|\mathfrak{A}|_\tau$.

(De acordo com a notação mais usual, convencionaremos escrever $R^{\mathfrak{A}}$ no lugar de $\mathfrak{A}(R)$ e $f^{\mathfrak{A}}$ no lugar de $\mathfrak{A}(f)$).

Assim, além de especificar um domínio inicial $|\mathfrak{A}|$, $\mathfrak{A}$ associa a cada constante em $\mathcal{R}_\tau$ um elemento do domínio de tipo τ de $\mathfrak{A}$, ou seja, um elemento de $|\mathfrak{A}|_\tau$: se c é uma constante individual, ela é associada pela estrutura a um elemento de $|\mathfrak{A}|_0 = |\mathfrak{A}|$; e se R é uma constante relacional de tipo $\langle \tau_1, \ldots, \tau_n\rangle$, ela é associada a um elemento de $|\mathfrak{A}|_{\langle \tau_1, \ldots, \tau_n\rangle} = \mathcal{P}(|\mathfrak{A}|_{\tau_1} \times \cdots \times |\mathfrak{A}|_{\tau_n})$. A toda constante relacional n-ária R de tipo $\langle 0, \ldots, 0\rangle$, por exemplo, $\mathfrak{A}$ associa uma

relação n-ária em $|\mathfrak{A}|$; e a toda constante relacional R de tipo $\langle\langle 0\rangle,\langle 0\rangle\rangle$, uma relação binária em $P(|\mathfrak{A}|)$ – i.é., $R^{\mathfrak{A}} \in \mathcal{P}(\mathcal{P}(|\mathfrak{A}|)\times\mathcal{P}(|\mathfrak{A}|))$.

Similarmente, a toda constante funcional n-ária f de tipo $\langle 0,\ldots,0,0\rangle$, $\mathfrak{A}$ associa uma função $f^{\mathfrak{A}}: |\mathfrak{A}|^n \to |\mathfrak{A}|$, e a toda constante funcional f de tipo $\langle\langle 0,0\rangle,\langle 0,0\rangle\rangle$, uma função de $\mathcal{P}(|\mathfrak{A}|\times|\mathfrak{A}|)$ em $\mathcal{P}(|\mathfrak{A}|\times|\mathfrak{A}|)$. Note-se que, assim como na lógica (clássica) de primeira ordem, para toda constante funcional f de tipo $\langle\tau_1,\ldots,\tau_n,\tau\rangle$, $f^{\mathfrak{A}}$ é uma função total em $|\mathfrak{A}|_{\langle\tau_1,\ldots,\tau_n\rangle}$: ela atribui um valor (em $|\mathfrak{A}|_\tau$) a absolutamente todos os elementos de $|\mathfrak{A}|_{\tau_1}\times\cdots\times|\mathfrak{A}|_{\tau_n}$. Em particular, em uma $\Sigma_{\mathbf{F}}$-estrutura $\mathfrak{F}$, $\#^{\mathfrak{F}}$ é uma função total de $\mathcal{P}(|\mathfrak{F}|)$ em $|\mathfrak{F}|$, i.é., uma função que atribui a cada subconjunto do domínio da estrutura $\mathfrak{F}$, um elemento do próprio domínio.

Exemplo. Como um exemplo de uma $\Sigma_{\mathbf{F}^{\approx}}$-estrutura $\mathfrak{F}$ temos

(i) O domínio $|\mathfrak{F}|$ de $\mathfrak{F}$ é o conjunto $\{a,b\}$;

(ii) $\approx^{\mathfrak{A}} = \{\langle\{a\},\{a\}\rangle,\langle\{a\},\{b\}\rangle,\langle\{b\},\{a\}\rangle,\langle\{b\},\{b\}\rangle,\langle\{a,b\},\{a,b\}\rangle\}$; e

(iii) $\#^{\mathfrak{F}} = \{\langle\varnothing,a\rangle,\langle\{a\},a\rangle,\langle\{b\},b\rangle,\langle\{a,b\},b\rangle\}$, ou seja,

$$\#^{\mathfrak{F}}(\varnothing) = \#^{\mathfrak{F}}(\{a\}) = a \text{ e } \#^{\mathfrak{F}}(\{b\}) = \#^{\mathfrak{F}}(\{a,b\}) = b.$$

Da mesma forma que uma estrutura $\mathfrak{A}$ associa um elemento de $|\mathfrak{A}|_\tau$ a cada constante em $\mathcal{R}_\tau$, uma *atribuição de valores a variáveis* em $\mathfrak{A}$ associa um elemento de $|\mathfrak{A}|_\tau$ a cada variável em $\mathcal{V}_\tau$, em que $\tau \in \mathcal{T}_2$:

Definição 1.11. Uma *atribuição* (*de valores a variáveis*) σ *em* $\mathfrak{A}$ é uma função em $\mathbf{V}$ tal que, para todo $\tau \in \mathcal{T}_2$, $\sigma(\mathbf{v}^\tau) \in |\mathfrak{A}|_\tau$.

Portanto, a toda variável individual, uma atribuição σ em $\mathfrak{A}$ associa um elemento de $|\mathfrak{A}|$, e a toda variável relacional n-ária, uma relação n-ária em $|\mathfrak{A}|$, i.é., um elemento de $|\mathfrak{A}|_{\langle 0,\ldots,0\rangle}$. Na $\Sigma_{\mathbf{F}^{\approx}}$-estrutura $\mathfrak{F}$ acima, por exemplo, a função σ em $\mathbf{V}$ que associa o indivíduo a a todas as variáveis individuais; o conjunto $\{a\}$ a V_i^1, se i é um número par, e o conjunto $\{a,b\}$, se i é um número ímpar; e, para $n \geq 2$, a relação $\{\langle a,\ldots,a\rangle\}$ a todas as variáveis relacionais n-árias é uma atribuição em $\mathfrak{F}$.

Dadas uma assinatura Σ, uma Σ-estrutura $\mathfrak{A}$ e uma atribuição σ em $\mathfrak{A}$, definimos a σ-*denotação* de um termo $t \in \mathbf{T}(\Sigma)$ em $\mathfrak{A}$ e a relação entre $\mathfrak{A}$, σ e uma fórmula $A \in \mathrm{F}(\Sigma)$ quando σ *satisfaz* A em $\mathfrak{A}$ da mesma forma que na lógica de primeira ordem:

Definição 1.12. A σ-*denotação* de t em $\mathfrak{A}$ (denotado por $t^{\mathfrak{A}}[\sigma]$) é definida recursivamente por:

(i) Se $t \in \mathbf{V}$, então $t^{\mathfrak{A}}[\sigma] = \sigma(t)$.

(ii) Se $t \in \mathcal{R}_\tau$, então $t^{\mathfrak{A}}[\sigma] = t^{\mathfrak{A}}$.

(iii) Para $1 \leq i \leq n$, se $t_i \in \mathrm{T}(\Sigma)_{\tau_i}$, t é $ft_1 \ldots t_n$ e $f \in \mathcal{F}_{\langle \tau_1, \ldots, \tau_n, \tau \rangle}$, então

$$t^{\mathfrak{A}}[\sigma] = f^{\mathfrak{A}}(t_1^{\mathfrak{A}}[\sigma], \ldots, t_n^{\mathfrak{A}}[\sigma]).$$

Definição 1.13. A atribuição σ *satisfaz* A *em* $\mathfrak{A}$ (denotado por $\vDash_{\mathfrak{A}} A\ [\sigma]$) como se segue:

(i) Se A é $tt_1 \ldots t_n$, então $\vDash_{\mathfrak{A}} A\ [\sigma]$ sse $\langle t_1^{\mathfrak{A}}[\sigma], \ldots, t_n^{\mathfrak{A}}[\sigma] \rangle \in t^{\mathfrak{A}}[\sigma]$.

(ii) Se $t_1, t_2 \in \mathrm{T}(\Sigma)_0$ e A é $t_1 = t_2$, então $\vDash_{\mathfrak{A}} A\ [\sigma]$ sse $t_1^{\mathfrak{A}}[\sigma] = t_2^{\mathfrak{A}}[\sigma]$.

(iii) Se A é $\neg B$, então $\vDash_{\mathfrak{A}} A\ [\sigma]$ sse $\nvDash_{\mathfrak{A}} B\ [\sigma]$.

(iv) Se A é $B \to C$, então $\vDash_{\mathfrak{A}} A\ [\sigma]$ sse $\nvDash_{\mathfrak{A}} B\ [\sigma]$ ou $\vDash_{\mathfrak{A}} C\ [\sigma]$.

(v) Se A é $\forall \mathbf{v}^\tau B$, então $\vDash_{\mathfrak{A}} A\ [\sigma]$ sse $\vDash_{\mathfrak{A}} B\ [\sigma(a/\mathbf{v}^\tau)]$, para todo $a \in |\mathfrak{A}|_\tau$.

Em (v), $\sigma(a/\mathbf{v}^\tau)$ denota a atribuição que concorda com σ em todas variáveis, exceto possivelmente em $\mathbf{v}^\tau$:

$$\sigma(a/\mathbf{v}^\tau)(\mathbf{v}) = \begin{cases} \sigma(\mathbf{v}) & \text{se } \mathbf{v} \neq \mathbf{v}^\tau \\ a & \text{se } \mathbf{v} = \mathbf{v}^\tau \end{cases}$$

Com o conceito de *satisfação*, defininem-se os conceitos semânticos de *verdade*, *validade*, *modelo*, *satisfatibilidade* e *consequência lógica* da maneira usual: uma fórmula A é *verdadeira* em uma estrutura $\mathfrak{A}$ ou $\mathfrak{A}$ é *modelo* de A (denotado por $\vDash_{\mathfrak{A}} A$) se $\vDash_{\mathfrak{A}} A\ [\sigma]$, para

toda atribuição σ em $\mathfrak{A}$; e A é (*logicamente*) *válida* ($\vDash A$) se $\vDash_{\mathfrak{A}} A$, para toda estrutura $\mathfrak{A}$. $\mathfrak{A}$ é *modelo* de um conjunto Γ de fórmulas ($\vDash_{\mathfrak{A}} \Gamma$) se todas as fórmulas em Γ são verdadeiras em $\mathfrak{A}$; e Γ é *satisfatível* se existem uma estrutura $\mathfrak{A}$ e uma atribuição σ em $\mathfrak{A}$ tais que $\vDash_{\mathfrak{A}} \Gamma\,[\sigma]$. Por fim, A é uma *consequência lógica* (*semântica*) de Γ ($\Gamma \vDash A$) se, para toda estrutura $\mathfrak{A}$ e para toda atribuição σ em $\mathfrak{A}$, $\vDash_{\mathfrak{A}} A\,[\sigma]$, sempre que $\vDash_{\mathfrak{A}} \Gamma\,[\sigma]$.

1.3 Um Sistema Dedutivo

Fixaremos agora um sistema axiomático particular que servirá como base para definirmos o conceito de *teoria de segunda ordem*, com o qual caracterizaremos a Aritmética de Frege. Alguns metateoremas serão estabelecidos no sentido de aprimorar a tratabilidade dedutiva de **AF** no desenvolvimento da prova do teorema homônimo. Tanto o sistema apresentado quanto os metateoremas enunciados em seguida são uma extensão do sistema e dos metateoremas formulados em [16] para a lógica de primeira ordem e omitiremos, portanto, as provas dos resultados da seção 1.3.2, assumindo que elas podem ser facilmente adaptadas para o nosso caso. Atenção especial será dedicada ao *operador de abstração* na seção 1.3.3, em razão de não encontrarmos um tratamento rigoroso e plenamente desenvolvido desse operador na literatura.

1.3.1 Axiomas Lógicos

Começaremos por especificar os *axiomas lógicos* de nosso sistema dedutivo, o qual denominaremos $\mathcal{D}_2$: se Σ é uma assinatura de segunda ordem e $A, B \in \mathrm{F}(\Sigma)$, diremos que A é uma *generalização* de B se, para algum $n \geq 1$, existem variáveis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ tais que A é a fórmula $\forall \mathbf{v}_1 \ldots \forall \mathbf{v}_n B$. O conjunto dos axiomas lógicos sobre Σ será então constituído por todas as instâncias dos esquemas $L1$–$L10$ abaixo e por quaisquer generalizações dessas instâncias. Mais precisamente:

Definição 1.14. O conjunto $\mathrm{AL}(\Sigma)$ de *axiomas lógicos sobre* Σ é definido indutivamente por

(i) Todas as fórmulas de $\mathrm{F}(\Sigma)$ que são instâncias dos esquemas $L1$–$L10$ abaixo são elementos de $\mathrm{AL}(\Sigma)$:

$L1$. $A \to (B \to A)$

$L2$. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

$L3$. $(\neg A \to \neg B) \to (B \to A)$

$L4$. $\forall \mathbf{v}^\tau A \to A[t/\mathbf{v}^\tau]$, em que $t \in \mathrm{T}(\Sigma)_\tau$, $\mathbf{v}^\tau$ é substituível por t em A e $A[t/\mathbf{v}^\tau]$ denota o resultado da substituição de todas as ocorrências livres de $\mathbf{v}^\tau$ em A por t.

$L5$. $A \to \forall \mathbf{v} A$, em que $\mathbf{v}$ não é livre em A.

$L6$. $\forall \mathbf{v}(A \to B) \to (\forall \mathbf{v} A \to \forall \mathbf{v} B)$

$L7$. $v_i = v_i$

$L8$. $v_i = v_j \to (A \to A[v_j|v_i])$, em que $A[v_j|v_i]$ denota o resultado da substituição de 0 ou mais ocorrências livres de v_i em A por v_j e todas as ocorrências de v_i substituídas para se obter $A[v_j|v_i]$ são substituíveis por v_j.

$L9$. $\forall v_{i_1} \ldots \forall v_{i_n}(V_j^n v_{i_1} \ldots v_{i_n} \leftrightarrow V_k^n v_{i_1} \ldots v_{i_n}) \to (A \to A[V_k^n|V_j^n])$, em que $A[V_k^n|V_j^n]$ denota o resultado da substituição de 0 ou mais ocorrências livres de V_j^n em A por V_k^n e todas as ocorrências de V_j^n substituídas para se obter $A[V_k^n|V_j^n]$ são substituíveis por V_k^n.

$L10$. $\exists V_j^n \forall v_{i_1} \ldots \forall v_{i_n}(V_j^n v_{i_1} \ldots v_{i_n} \leftrightarrow A)$, em que $v_{i_1} \ldots v_{i_n}$ são livres e V_j^n não é livre em A.

(ii) Se $A \in \mathrm{AL}(\Sigma)$ e $\mathbf{v} \in \mathbf{V}$, então $\forall \mathbf{v} A \in \mathrm{AL}(\Sigma)$.

$L1$–$L3$ são os axiomas proposicionais de $\mathcal{D}_2$ e $L4$–$L6$ regulamentam o comportamento do quantificador universal, quando ligado a variáveis de primeira e de segunda ordem; assim, tanto

$$\forall v_1 \; V_0 v_1 \to V_0 v_5$$

quanto

$$\forall V_0\, V_0 v_1 v_1 \to V_5 v_1 v_1$$

são instâncias de $L4$ (o papel de $L5$ e $L6$ – e também da cláusula (ii) da definição de $\mathrm{AL}(\Sigma)$ – se restringe apenas a possibilitar a prova da *regra de generalização universal* abaixo – Teorema 1.19)[8].

$L7$ e $L8$ são os axiomas de identidade: $L7$ estabelece a reflexividade de $=$ e $L8$ é o *princípio de indiscernibilidade dos idênticos*. $L9$ é uma versão de segunda ordem desse princípio, tratando-se do único axioma de $\mathcal{D}_2$ que não possui um correspondente similar nos sistemas axiomáticos para as formulações padrão da *LSO*: nas linguagens desses sistemas, todas as ocorrências livres de uma variável de segunda ordem V_j^n em uma fórmula são concatenadas a n termos de primeira ordem, o que permite que $L9$ seja facilmente obtido como um corolário do *Teorema da Substituição*[9]. No nosso caso, entretanto, variáveis relacionais e, em geral, termos de segunda ordem podem ocorrer como argumentos de constantes relacionais e constantes funcionais de terceira ordem – como, por exemplo, a variável V_0^1 nas fórmulas $\#V_0^1 = v_1$ e $V_0^1 \approx V_1^1$ – e, para garantir a substitutibilidade desses termos também nesses contextos, foi preciso incluir as instâncias de $L9$ no conjunto $\mathrm{AL}(\Sigma)$.

Note-se que $L9$ é logicamente válido, pois, na presença do axioma da extensionalidade na metateoria, a satisfatição do antecedente $\forall v_{i_1} \ldots \forall v_{i_n}(V_j^n v_{i_1} \ldots v_{i_n} \leftrightarrow V_k^n v_{i_1} \ldots v_{i_n})$ em uma dada estrutura e atribuição implica que as variáveis V_j^n e V_k^n denotam a mesma relação n-ária construída sobre o domínio; e como todos os símbolos não lógicos de uma linguagem são interpretados extensionalmente (conjuntisticamente), isso implica que o consequênte do axioma também é satisfeito naquela estrutura e atribuição. No caso particular da instância (*) $\forall v_1(V_1 v_1 \leftrightarrow V_2 v_1) \to (\#V_1 = \#V_1 \to \#V_1 = \#V_2)$, por exemplo, dadas uma $\Sigma_{\mathbf{F}}$-estrutura $\mathfrak{F}$ e

[8] A rigor, $L1$–$L10$ não são axiomas de $\mathcal{D}_2$, e sim todas as instâncias de $L1$–$L10$. No entanto, iremos nos referir a esses esquemas e a suas instâncias indistintamente por 'axioma'/'axiomas', deixando que o contexto esclareça o sentido dessas expressões em cada ocasião.

[9] O *Teorema da Substituição* afirma que se B é uma subfórmula de A e $A[C|B]$ é o resultado da substituição de 0 ou mais ocorrências de B em A pela fórmula C, então $\vdash \forall \mathbf{v}_1 \ldots \mathbf{v}_n(B \leftrightarrow C) \to (A \leftrightarrow A[B|C])$, em que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ são todas as variáveis livres em B ou C e ligadas em A ([28], Proposição 2.9).

uma atribuição σ em $\mathfrak{F}$, se $\vDash_{\mathfrak{F}} \forall v_1(V_1v_1 \leftrightarrow V_2v_1)\ [\sigma]$, então exatamente os mesmos elementos de $|\mathfrak{F}|$ pertencem aos conjuntos $\sigma(V_1)$ e $\sigma(V_2)$, e $\sigma(V_1) = \sigma(V_2)$. Logo, a função $\#^{\mathfrak{F}}$ associa o mesmo indivíduo de $|\mathfrak{F}|$ a $\sigma(V_1)$ e a $\sigma(V_2)$ e, portanto, $\vDash_{\mathfrak{F}} \#V_1 = \#V_2\ [\sigma]$[10].

Por fim, $L10$ é denominado *axioma da compreensão* e afirma que toda fórmula A em que as variáveis de primeira ordem $v_{i_1} \dots v_{i_n}$ são livres define ou determina uma relação n-ária ou, alternativamente, que para toda A em que $v_{i_1}, \dots, v_{i_n}$ são livres, existe uma relação n-ária de mesma extensão. A restrição de que V_j^n não seja livre em A serve para bloquear instânciais paradoxais como

$$\exists V_j^1 \forall v_i (V_j^1 v_i \leftrightarrow \neg V_j^1 v_i)$$

$$\exists V_j^2 \forall v_{i_1} \forall v_{i_2} (V_j^2 v_{i_1} v_{i_2} \leftrightarrow \neg V_j^2 v_{i_1} v_{i_2})$$

$$\vdots$$

Assumiremos, sem prova, que todas instâncias de $L1$-$L10$ são logicamente válidas: a prova da validade de $L1$-$L8$ é inteiramente análoga à prova da validade de suas versões para a lógica de primeira ordem e, como notamos, $L9$ também é logicamente válido, uma vez que todos os símbolos não lógicos de uma dada linguagem são interpretados extensionalmente por uma estrutura (ver nota 12 acima). $L10$, por sua vez, é irrestritamente válido apenas se, para cada $\tau = \langle 0, \dots, 0 \rangle$, o domínio de tipo τ de uma estrutura $\mathfrak{A}$ é constituído por todas as relações n-árias contruídas sobre $|\mathfrak{A}|$. Em particular, as instâncias da restrição do esquema $L10$ a variáveis de segunda ordem unárias – $\exists V_j \forall v_i (V_J v_i \leftrightarrow A)$ – são válidas somente se $|\mathfrak{A}|_{\langle 0 \rangle}$ compreende todos os subconjuntos de $|\mathfrak{A}|$.

[10]Uma prova plenamente desenvolvida da validade de $L9$ procede por indução sobre as subfórmulas de A e é similar à prova da validade do axioma análogo $L8$ (ver [16], p. 132). Assim como no caso dos outros axiomas do sistema $\mathcal{D}_2$, nesta dissertação, não apresentaremos uma tal prova, mas esperamos que os comentários acima sejam suficientes para convencer o leitor da validade de todas as instâncias de $L9$.

Embora não usual em sistemas axiomáticos para a lógica de segunda ordem padrão, uma versão da instância (*) de $L9$ acima foi adotada por Boolos a título de *axioma lógico* em [7] para investigar as relações entre diferentes formalizações da aritmética na *LSO*, incluindo-se a Aritmética de Frege. A versão de (*) em questão é a fórmula $\forall v_i (V_j v_i \leftrightarrow V_k v_i) \rightarrow \#V_j = \#V_k$, denominada 'FE' (para 'functional extensionality').

1.3.2 Metateoremas, Regras Derivadas e Demonstrações

O sistema $\mathcal{D}_2$ possui *modus ponens* como única regra de inferência:

(*MP*)

$$\frac{A, A \to B}{B}$$

Uma *derivação* em $\mathcal{D}_2$ de uma fórmula A a partir de um conjunto Γ de fórmulas é definida da maneira usual como uma sequência finita de fórmulas $A_1 \ldots A_n$ tal que A é A_n e, para todo $1 \leq i \leq n$,

(i) $A_i \in \mathrm{AL}(\Sigma)$; ou

(ii) $A_i \in \Gamma$; ou

(iii) Existem $j, k < n$ tais que A_k é $A_j \to A_i$.

Se existe uma derivação de A a partir de Γ, diremos que A é uma *consequência lógica sintática* de Γ e escreveremos $\Gamma \vdash A$; e se $\varnothing \vdash A$, diremos que A é um *teorema lógico* e escreveremos $\vdash A$. Também convencionaremos escrever $\Gamma, A \vdash B$ no lugar de $\Gamma \cup \{A\} \vdash B$. Se não existe A tal que $\Gamma \vdash A$ e $\Gamma \vdash \neg A$, Γ será dito um conjunto *consistente* – e *inconsistente*, caso contrário.

A princípio, a exiguidade na quantidade de regras do sistema $\mathcal{D}_2$ representa um empecilho para a manipulação de teorias concretas e um sistema de dedução natural pode parecer mais adequado para o desenvolvimento de uma teoria formalizada particular. No entanto, essa desvantagem é contrabalanceada pelo fato de que provas por indução de resultados metateóricos sobre sistemas axiomáticos são, em geral, consideravelmente mais fáceis do que as provas dos resultados correspondentes sobre sistemas de dedução natural, exigindo a análise de poucos casos no passo indutivo. Além disso, é possível remediar a intratabilidade dedutiva de sistemas axiomáticos, simulando, através de metaregras de inferência, as regras que normalmente estão presentes em um sistema de dedução natural. Nesse sentido, enunciaremos a seguir uma série de teoremas que atendam a esse propósito:

Teorema 1.15. TEOREMA DA DEDUÇÃO (*TD*). Sejam Σ uma assinatura de segunda ordem e $\Gamma \cup \{A, B\} \subseteq \mathrm{F}(\Sigma)$. Se $\Gamma, A \vdash B$, então $\Gamma \vdash A \to B$.

Teorema 1.16. REDUÇÃO AO ABSURDO (*RAA*). Se $\{\Gamma, A\}$ é inconsistente, então $\Gamma \vdash \neg A$; e se $\{\Gamma, \neg A\}$ é inconsistente, então $\Gamma \vdash A$.

Teorema 1.17. REGRA DE INSTANCIAÇÃO UNIVERSAL ($E\forall$). Seja $t \in \mathrm{T}(\Sigma)_\tau$. Se $\Gamma \vdash \forall \mathbf{v}^\tau A$ e $\mathbf{v}^\tau$ é substituível por t em A, então $\Gamma \vdash A[t/\mathbf{v}^\tau]$.

Teorema 1.18. REGRA DE GENERALIZAÇÃO EXISTENCIAL ($I\exists$). Seja $t \in \mathrm{T}(\Sigma)$. Se $\Gamma \vdash A[t/\mathbf{v}^\tau]$ e $\mathbf{v}^\tau$ é substituível por t em A, então $\Gamma \vdash \exists \mathbf{v}^\tau A$.

Teorema 1.19. REGRA DE GENERALIZAÇÃO UNIVERSAL ($I\forall$). Se $\Gamma \vdash A$ e $\mathbf{v}$ não é livre em Γ, então $\Gamma \vdash \forall \mathbf{v} A$.

Teorema 1.20. REGRA DE INSTANCIAÇÃO EXISTENCIAL ($E\exists$). Se $\mathbf{v}$ não é livre em $\Gamma \cup \{B\}$, $\Gamma, A \vdash B$ e $\Gamma \vdash \exists \mathbf{v} A$, então $\Gamma \vdash B$.

Teorema 1.21. TEOREMA DA VARIAÇÃO ALFABÉTICA. Sejam $A \in \mathrm{F}(\Sigma)$, $\mathbf{v}^\tau \in \mathcal{V}_\tau$ e $t \in \mathrm{T}(\Sigma)_\tau$. Existe $A' \in \mathrm{F}(\Sigma)$ que difere de A apenas na escolha das variáveis quantificadas tal que

(i) $\vdash A \leftrightarrow A'$.

(ii) $\mathbf{v}^\tau$ é substituível por t em A'.

Em razão dos metateoremas 1.17–1.21 ou suas versões para a lógica de primeira ordem serem bastante comuns no desenvolvimento de sistemas axiomáticos apresentados em livros-texto de lógica matemática, não exemplificaremos a aplicação das regras que eles codificam em derivações concretas no sistema $\mathcal{D}_2$. Na verdade, o estabelecimento dessas regras servirá, por um lado, para provarmos alguns resultados gerais sobre teorias de segunda ordem na seção 1.4 (especialmente, para provarmos o Teorema da Interpretação no Apêndice A2) e, por outro, para legitimarmos a formulação de *demonstrações* nas teorias formais com as quais trabalharemos: em virtude de sua complexidade e, principalmente, do comprimento das fórmulas envolvidas, as derivações relevantes nessas teorias dificilmente poderiam ser apresentadas de uma maneira plenamente detalhada de modo a tornar explícitos todos os passos

envolvidos; e, por essa razão, formularemos, de agora em diante, provas semi-formais de teoremas, omitindo passos elementares e utilizando, sem menção explícita, as regras enunciadas acima (e também as regras para o operador de abstração provadas abaixo). Nesta dissertação, provas desse tipo serão sistematicamente denominadas *demonstrações*. A partir de uma dada demonstração, acreditamos ser possível reconstruir de maneira mais ou menos imediata a derivação correspondente em $\mathcal{D}_2$, embora essa possa ser uma tarefa extremamente trabalhosa em alguns casos particulares.

Além de recorrermos a demonstrações, também adotaremos convenções adicionais que otimizarão a leitura das fórmulas de uma dada linguagem e que reestabelecem um acordo notacional com a maior parte das exposições do Teorema de Frege presentes na literatura e com as discussões preliminares da Introdução e do início deste capítulo: de agora em diante, escreveremos

(i) $u, x, y, z, w, u_1, \ldots$ no lugar de $v_0, v_1, v_2, v_3, v_4, v_5, \ldots$;

(ii) $F, G, H, F_1, \ldots$ no lugar de $V_0^1, V_1^1, V_2^1, V_3^1, \ldots$; e

(iii) $R, S, T, R_1, \ldots$ no lugar de $V_0^2, V_1^2, V_2^2, \ldots$.

(iv) $R^3, S^3, T^3, R_1^3, \ldots$ no lugar de $V_0^3, V_1^3, V_2^3, \ldots$.

Como exemplos do emprego de (i)–(iii), temos as seguintes instâncias dos axiomas *L*4 e *L*10:

1. $\forall x(Fx \to Gx) \to (Fy \to Gy)$

2. $\forall F \forall x Fx \to \forall x Gx$

3. $\forall R \forall x \exists y Rxy \to \forall x \exists y Sxy$

4. $\exists F \forall u(Fu \leftrightarrow u \neq u)$

5. $\exists R \forall u \forall u_1 (Ruu_1 \leftrightarrow u = u_1)$

1.3.3 O Operador de Abstração

Em grande parte das exposições da prova do Teorema de Frege na literatura, recorre-se a um *operador de abstração* ou *abstrator* $[_\!:_]$ que permite formar um termo de segunda ordem $[v_{i_1}\ldots v_{i_n} : B]$ a partir das variáveis $v_{i_1},\ldots,v_{i_n}$ e de uma fórmula B na qual elas ocorram livres, de modo que se $t,t_1,\ldots,t_n$ são termos individuais, expressões da forma $[v_{i_1}\ldots v_{i_n} : B]t_1\ldots t_n$, $V_j(\#[v_i : B])$, $\#[v_i : B] = t$ são tratadas como fórmulas atômicas da linguagem da Aritmética de Frege. O recurso a esse operador representa um expediente extremamente conveniente para denotar as propriedades e relações que o axioma da compreensão afirma existir, pois ele permite expressá-las localmente através das fórmulas que as definem, facilitando imensamente a confecção de determinardas demonstrações/derivações em teorias de segunda ordem: a instância $\exists F\forall u(Fu \leftrightarrow u \neq u)$ de $L10$, por exemplo, garante a existência da propriedade (vazia) de todos os objetos distintos de si próprios, a qual é expressa por $[u : u \neq u]$; e em geral, instâncias da forma $\exists V_j^n \forall v_{i_1}\ldots\forall v_{i_n}(V_j^n v_{i_1}\ldots v_{i_n} \leftrightarrow B)$ garantem a existência da relação n-ária $[v_{i_1}\ldots v_{i_n} : B]$ definida por B[11].

No próximo capítulo, também utilizaremos o operador de abstração na exposição da prova do Teorema de Frege como um meio para denotar as propriedades e relações definíveis pelas fórmulas de $\Sigma_{\mathbf{F}}$. Contudo, como não incluímos $[_\!:__]$ entre os símbolos lógicos das linguagens caracterizadas acima, iremos introduzí-lo como um símbolo metalinguisticamente ou contextualmente definido: assim como "fórmulas" da forma $A \vee B$ são consideradas como abreviações de fórmulas da forma $(\neg A \rightarrow B)$, "fórmulas" em que $[_\!:__]$ ocorra serão tratadas como abreviações de fórmulas mais complexas nas quais ele não apareça. Para legitimarmos nossa definição, provaremos então alguns teoremas que *simulam* em $\mathcal{D}_2$ os axiomas e regras

[11] Na literatura, não existe uma notação uniforme para o operador de abstração: por exemplo, Boolos [2] e Boolos e Heck [9] (Apêndice 2) utilizam $\{_\!:_\}$, Boolos [3, 4, 5, 7] utiliza $[_\!:__]$ e Zalta [43] utiliza $[\lambda_{\ldots}\ _]$. Outro autores, como Dummett, Wright e, em muitas ocasiões, Heck tratam # como um operador que, quando aplicado a uma variável individual x e a uma fórmula A, resulta um termo de primeira ordem $\#xA$, ou seja, como um *variable binding term operator* ou *VBTO*; esse tratamento dispensa o uso operador abstração na prova do Teorema de Frege, mas impõe algumas dificuldades técnicas na definição das linguagens de segunda ordem apropriadas para formalizar teorias fregeanas.

que são usualmente empregados na manipulação desse operador e que presumivelmente estariam presentes em qualquer sistema dedutivo para uma linguagem em que ele figurasse como um símbolo lógico, da mesma forma que, para legitimar a definição de $A \vee B$ como $(\neg A \to B)$, provam-se teoremas que correspondem às regras de introdução e eliminação de $\vee$[12].

Com o intuito de simplificar a discussão a seguir, consideraremos aqui apenas o caso do *operador de abstração unário*, que permite formar "termos" da forma $[v_i : B]$, uma vez que a definição e os teoremas subsequêntes podem ser facilmente generalizados para operadores de qualquer aridade. No entanto, nos próximos capítulos também utilizaremos o *operador de abstração binário* $[v_i v_j : B]$ para denotar relações binárias definíveis pelas fórmulas das linguagens sob consideração e assumiremos versões análogas dos teoremas e das regras estabelecidos abaixo para "termos" de segunda ordem formados por esse operador. De agora em diante, reservaremos as variáveis $u, u_1, u_2, \ldots$ para ocuparem o lugar de $_{\ldots}$ em $[_{\ldots} : __]$.

Sejam Σ uma assinatura de segunda ordem, $A, B \in \mathrm{F}(\Sigma)$, v_i uma variável individual livre em B e V_j uma variável relacional unária. Introduzimos a notação $A\big[[v_i : B]/V_j\big]$ para denotar o resultado da substituição de todas as ocorrências livres de V_j em A por $[v_i : B]$. Por exemplo, se A é $\forall x \forall y(Fx \to Rxy)$ e B é $u \neq u$, então

$$A\big[[u : B]/F\big] = \forall x \forall y(Fx \to Rxy)\big[[u : u \neq u]/F\big] =$$

$$\forall x \forall y([u : u \neq u]x \to Rxy)$$

Definição 1.22. Formalmente, definimos $A\big[[v_i : B]/V_j\big]$ recursivamente como se segue:

(i) Se A é uma fórmula atômica, então

[12] O método de definição do operador de abstração a ser empregado aqui é essencialmente o mesmo que o da definição do *operador de descrição definida* $\imath$ originalmente proposto por Russell em *On Denoting* [33] e formalmente elaborado no Volume I do *Principia Mathematica* [39]: assim como "termos" (individuais) formados por $\imath$ na análise russeliana, "termos" (de segunda ordem) formados por $[_{\ldots} : __]$ serão tratados *expressões incompletas*, no sentido de que eles não se comportam semanticamente como termos e não possuem significado isoladamente dos contextos em que ocorram; e assim como a definição de Russell, nossa definição consistirá em fornecer um procedimento geral e sistemático para traduzir *contextos sentencias* que contenham ocorrências do operador de abstração em contextos em que ele esteja ausente; em outras palavras, consistirá em uma *definição contextual*.

$$A\big[[v_i:B]/V_j\big] = \begin{cases} A & \text{se } V_j \text{ não ocorre em } A \\ \exists V_k(\forall v_i(V_k v_i \leftrightarrow B) \wedge A[V_k/V_j]) & \text{caso contrário} \end{cases}$$

(ii) Se A é $(\neg C)$, então $A\big[[v_i:B]/V_j\big] = \big(\neg C\big[[v_i:B]/V_j\big]\big)$.

(iii) Se A é $(C \to D)$, então $A\big[[v_i:B]/V_j\big] = \big(C\big[[v_i:B]/V_j\big] \to D\big[[v_i:B]/V_j\big]\big)$.

(iv) Se A é $\forall \mathbf{v} C$, então

$$A\big[[v_i:B]/V_j\big] = \begin{cases} A & \text{se } \mathbf{v} = V_j \\ \forall \mathbf{v} C\big[[v_i:B]/V_j\big] & \text{caso contrário} \end{cases}$$

Em (i), assumimos que V_k não ocorre em A e não é livre em B[13].

De acordo com essa definição, "fórmulas atômicas" em que a expressão $[v_i:B]$ ocorra, como $[v_i:B]t$ e, em se tratando de $\Sigma_{\mathbf{F}}$, $V_l(\#[v_i:B])$ e $\#[v_i:B] = t$, são consideradas como abreviações de fórmulas da forma $\exists V_k(\forall v_i(V_k v_i \leftrightarrow B) \wedge A[V_k/V_j])$:

$$[v_i:B]t \;:=\; \exists V_k(\forall v_i(V_k v_i \leftrightarrow B) \wedge V_k t)$$

$$V_l(\#[v_i:B]) \;:=\; \exists V_k(\forall v_i(V_k v_i \leftrightarrow B) \wedge V_l(\#V_k))$$

$$\#[v_i:B] = t \;:=\; \exists V_k(\forall v_i(V_k v_i \leftrightarrow B) \wedge \#V_k = t)$$

Nesses contextos, V_k desempenha o papel de um representativo para $[v_i:B]$: a conjunta $\forall v_i(V_k v_i \leftrightarrow B)$ a identifica como uma variável que "denota" a propriedade definida por B – cuja existência é garantida pela instância $\exists V_k \forall v_i(V_k v_i \leftrightarrow B)$ de $L10$ – e $A[V_k/V_j]$ expressa, em termos dessa variável, o que inicialmente se enuncia de $[v_i:B]$ através de $A\big[[v_i:B]/V_j\big]$. As "fórmulas moleculares", por sua vez, simplesmente abreviam o resultado da substituição de todas as "subfórmulas atômicas" em que $[v_i:B]$ ocorra pelas fórmulas expandidas correspondentes de acordo com a cláusula (i); no exemplo acima,

$$\forall x \forall y([u:u \neq u]x \to Rxy) \;:=\; \forall x \forall y(\exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge Gx) \to Rxy)$$

[13] A Definição 1.22 foi formulada a partir das indicações de Boolos e Heck no Apêndice 2 de [9].

Exemplo.

1.	$[u : u \neq u]x$	:=	$\exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge Gx)$
2.	$F(\#[u : u \neq u])$	:=	$\exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge F(\#G))$
3.	$\forall x \neg [u : u \neq u]x$	:=	$\forall x \neg \exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge Gx)$
4.	$\exists F\ \#F = \#[u : u \neq u]$	:=	$\exists F \exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge \#F = \#G)$

A partir da Definição 1.22, estabelecemos primeiramente duas regras elementares, denominadas *redução* e *abstração*, que governam o comportamento do operador de abstração em contextos atômicos do forma $[v_i : B]t$ e que são obtidas imediatamente do seguinte teorema:

Teorema 1.23. Se A é $V_j v_i$, então $\vdash \forall v_i\big(A\big[[v_i : B]/V_j\big] \leftrightarrow B\big)$; ou seja, $\vdash \forall v_i([v_i : B]v_i \leftrightarrow B)$.

Prova: Apêndice A1.

■

Exemplo.

1. $\vdash \forall u([u : u \neq u]u \leftrightarrow u \neq u)$

2. $\vdash \forall u([u : Fu \vee Gu]u \leftrightarrow Fu \vee Gu)$

3. $\vdash \forall u([u : \exists x Rxu]u \leftrightarrow \exists x Rxu)$

Corolário 1.24. Seja t um termo individual tal que v_i é substituível por t em B.

(i) REDUÇÃO (*RED*). Se $\Gamma \vdash [v_i : B]t$, então $\Gamma \vdash B[t/v_i]$.

(ii) ABSTRAÇÃO (*ABS*). Se $\Gamma \vdash B[t/v_i]$, então $\Gamma \vdash [v_i : B]t$.

A regra de redução permite eliminar o operador de abstração em $[v_i : B]t$ em favor da fórmula $B[t/v_i]$ e estabelece que se o objeto denotado por t satisfaz a propriedade $[v_i : B]$, então ele satisfaz B. Inversamente, a regra de abstração permite extrair ou abstrair $[v_i : B]$ a partir de $B[t/v_i]$ e estabelece que se o objeto denotado por t satisfaz B, então ele satisfaz a propriedade $[v_i : B]$. Abaixo, encontram-se dois exemplos da aplicação dessas regras em uma demonstração:

Exemplo.

1. Como um exemplo de uma demonstração envolvendo a regra de redução, mostramos que $\vdash \forall x \neg [u : u \neq u]x$:

Demonstração: Suponha $[u : u \neq u]x$. Por *RED*, obtemos $x \neq x$. Absurdo. Logo, $\neg [u : u \neq u]x$ e, portanto, $\forall x \neg [u : u \neq u]x$.

■

2. E como um exemplo de uma demonstração envolvendo a combinação de ambas as regras, mostramos que $\vdash [u : Rxu]y \rightarrow [u : \exists x\, xRxu]y$:

Demonstração: Suponha $[u : Rxu]y$. Por *RED*, obtemos Rxy e, portanto, $\exists x Rxy$. Por *ABS*, inferimos $[u : \exists x\, Rxu]y$. Logo, $[u : Rxu]y \rightarrow [u : \exists x\, xRxu]y$.

■

Um segundo resultado obtido da Definição 1.22 simula no sistema $\mathcal{D}_2$ uma regra de instanciação universal para "termos" de segunda ordem formados pelo operador de abstração (Corolário 1.27), a qual, grosso modo, permite instanciar a variável V_j em uma dada fórmula $\forall V_j A$ para $[v_i : B]$, resultando $A\big[[v_i : B]/V_j\big]$[14]. Essa regra será amplamente utilizada nos próximos capítulos e, para prová-la, recorreremos ao seguinte teorema, que, sob a hipótese $\forall v_i (V_j v_i \leftrightarrow B)$, estabelece a intersubstituibilidade da variável V_j e de $[v_i : B]$ em A[15]:

Teorema 1.25. Sejam $A, B \in \mathrm{F}(\Sigma)$ tais que v_i é livre em B. Se nenhuma das variáveis livres em B é ligada em A, então

$$\forall v_i (V_j v_i \leftrightarrow B) \vdash A \leftrightarrow A\big[[v_i : B]/V_j\big]$$

Prova: A prova do Teorema 1.25 procede por indução sobre as subfórmulas de A e encontra-se desenvolvida em detalhe no Apêndice A1.

■

[14] A notação $A\big[[v_i : B]/V_j\big]$ foi especificamente escolhida para facilitar a formulação dessa regra, tornando-a similar à formulação da regra $E\forall$ acima (Teorema 1.17).

[15] Os resultados seguintes foram formulados a partir das indicações gerais de Boolos em [2], pp. 161–162.

Utilizando o Teorema 1.25, provamos então uma versão do axioma $L4$ para "termos" de segunda ordem formados por $[\ldots\colon __]$, do qual a regra de instanciação universal em questão é uma consequência imediata:

Teorema 1.26. Se nenhuma das variáveis livres em B é ligada em A e V_j não é livre em B, então $\vdash \forall V_j A \to A\big[[v_i : B]/V_j\big]$.

Prova: Pelo Teorema 1.25, temos

$$\forall V_j A, \forall v_i (V_j v_i \leftrightarrow B) \vdash A \leftrightarrow A\big[[v_i : B]/V_j\big]\big]$$

e, por $E\forall$ (aplicada a $\forall V_j A$),

$$\forall V_j A, \forall v_i (V_j v_i \leftrightarrow B) \vdash A$$

Logo, (*) $\forall V_j A, \forall v_i (V_j v_i \leftrightarrow B) \vdash A\big[[v_i : B]/V_j\big]$ (por MP). Como V_j não é livre em B,

$$(**)\ \forall V_j A \vdash \exists V_j \forall v_i (V_j v_i \leftrightarrow B)$$

(por $L10$). Aplicando $E\exists$ a (*) e (**), obtemos $\forall V_j A \vdash A\Big[[v_i : B]/V_j\Big]$, que, por TD, implica

$$\vdash \forall V_j A \to A\big[[v_i : B]/V_j\big]$$

■

Exemplo. $\vdash \forall F\big(\exists x F x \vee \neg \exists x F x\big) \to \exists x[u : u = u]x \vee \neg \exists x[u : u = u]x$

Corolário 1.27. Regra de Instanciação Universal II ($E\forall^{\dagger}$). Se nenhuma das variáveis livres em B é ligada em A, V_j não é livre em B e $\Gamma \vdash \forall V_j A$, então $\Gamma \vdash A\big[[v_i : B]/V_j\big]$.

No Teorema 1.26 e no Corolário 1.27, a restrição de que nenhuma das variáveis livres em B sejam ligadas em A serve para prevenir que essas variáveis sejam capturadas pelos quantificadores em A ao se substituir V_j por $[v_i : B]$ (da mesma forma que a restrição no axioma $L4$ de que a variável $\mathbf{v}^{\tau}$ seja substituível por t em A serve para prevenir que as variáveis em t sejam capturadas pelos quantificadores em A). Como um exemplo no qual a inobservância dessa restrição resulta ser problemática, considere-se o teorema (*) $\forall F \exists G \forall x (Fx \leftrightarrow Gx)$, obtido de

$\forall x(Fx \leftrightarrow Fx)$ pelas regras $I\exists$ e $I\forall$. Se instanciarmos a variável F para $[u : \neg Gu]$ em $(*)$ ao aplicarmos (erroneamente) a regra acima, obtemos $\exists G\forall x([u : \neg Gu]x \leftrightarrow Gx)$, que, pelas regras *RED* e *ABS*, implica imediatamente uma contradição.

Como uma consequência adicional do Teorema 1.26, obtemos uma regra de generalização existencial para "termos" de segunda ordem formados pelo operador de abstração:

Corolário 1.28. REGRA DE GENERALIZAÇÃO EXISTENCIAL II ($I\exists^\dagger$). Se nenhuma das variáveis livres em B é ligada em A, V_j não é livre em B e $\Gamma \vdash A\big[[v_i : B]/V_j\big]$, então $\Gamma \vdash \exists V_j A$.

Exemplo. No exemplo 1 para o Corolário 1.24, mostramos que $\vdash \forall x\neg[u : u \neq u]x$. Aplicando a regra $I\exists^\dagger$, obtemos $\vdash \exists F\forall x\neg Fx$, que afirma a existência de uma propriedade vazia. Utilizando essa mesma regra, é possível mostrar também que existe uma propriedade universal: $\vdash \exists F\forall xFx$.

Demonstração: Pelo axioma *L*7, temos $x = x$, que, por *ABS*, implica $[u : u = u]x$. Portanto, $\forall x[u : u = u]x$. Pela regra $I\exists^\dagger$, inferimos $\exists F\forall xFx$.

■

Ao simular, no sistema $\mathcal{D}_2$, o axioma *L*4 para "termos" de segunda ordem formados pelo operador de abstração, o Teorema 1.26 garante que esses "termos" se comportam dedutivamente como os termos de segunda ordem genuínos de uma dada linguagem, da mesma forma que os teoremas $A \to A \vee B$, $A \to B \vee A$ e $(A \to C) \to ((A \to C) \to ((A \vee B) \to C))$ da lógica proposicional clássica com os conectivos $\neg$ e $\to$ garantem que o símbolo $\vee$, definido contextualmente por $A \vee B := \neg A \to B$, se comporta dedutivamente como um símbolo de disjunção genuíno. É possível mostrar também um resultado similar para linguagens que possuem um operador @ que, assim como #, rebaixa a ordem dos termos aos quais se aplica, ou seja, um operador de tipo $\langle\langle 0\rangle, 0\rangle$ ou $\langle\langle 0,0\rangle, 0\rangle$ ou $\langle\langle 0,0,0\rangle, 0\rangle$ etc. Mais especificamente, é possível provar uma versão de *L*4 para "termos" de primeira ordem que resultam da aplicação de @ a "termos" formados por $[\ldots : __]$; no caso de $\Sigma_{\mathbf{F}}$, "termos" de primeira ordem da forma $\#[v_i : B]$: $\#[u : u = u], \#[u : u \neq u], \#[u : \neg Fu], \#[u : Fu \vee Gu], \#[u : Fu \wedge u \neq x]$.

Para estabelecer esse resultado, utilizaremos a regra $E\forall^\dagger$ acima e introduziremos a notação $A\big[@[v_i : B]/v_j\big]$ para denotar a fórmula $A[@V_j/v_j]\big[[v_i : B]/V_j\big]$, em que @ é um constante

funcional de tipo $\langle\langle 0\rangle,0\rangle$ e V_j, uma variável de segunda ordem unária arbitrária que não ocorre em A ou em B. Assim, $A\big[@[v_i:B]/v_j\big]$ é o resultado da substituição de todas as ocorrências livres da variável individual v_j em A por $@V_j$ e da substituição de V_j por $[v_i:B]$ na fórmula resultante.

Exemplo.

1.	$Fx\big[\#[u:u\neq u]/x\big]$	$=$	$F(\#G)\big[[u:u\neq u]/G\big]$	$=$	$F(\#[u:u\neq u])$
2.	$\#F=x\big[\#[u:u\neq u]/x\big]$	$=$	$\#F=\#G\big[[u:u\neq u]/G\big]$	$=$	$\#F=\#[u:u\neq u]$

Teorema 1.29. Sejam $\Sigma=\langle\mathcal{R},\mathcal{F}\rangle$ uma assinatura tal que $@\in\mathcal{F}_{\langle\langle 0\rangle,0\rangle}$ e $A,B\in \mathrm{F}(\Sigma)$. Se nenhuma das variáveis livres em B é ligada em A, então $\vdash \forall v_jA\rightarrow A\big[@[v_i:B]/v_j\big]$.

Prova: Por $E\forall$, temos $\forall v_jA\vdash A[@V_j/v_j]$ e, como V_j não ocorre em A, por $I\forall$, inferimos $\forall v_jA\vdash \forall V_jA[@V_j/v_j]$. Por $E\forall^{\dagger}$, obtemos então $\forall v_jA\vdash A[@V_j/v_j]\big[@[v_i:B]/V_j\big]$, ou seja, $\forall v_jA\vdash A\big[@[v_i:B]/v_j\big]$. Portanto, $\vdash \forall v_jA\rightarrow A\big[@[v_i:B]/v_j\big]$ (por *TD*).

■

Exemplo. $\vdash \forall x\, x=x\rightarrow \#[u:u\neq u]=\#[u:u\neq u]$

Assim como as regras $E\forall^{\dagger}$ e $I\forall^{\dagger}$ são corolários imediatos do Teorema 1.26, regras de instanciação universal e generalização existencial para "termos" de primeira ordem da forma $@[v_i:B]$ podem ser obtidas imediatamente do Teorema 1.29:

Corolário 1.30. REGRA DE INSTANCIAÇÃO UNIVERSAL III ($E\forall^{\dagger\dagger}$). Se nenhuma das variáveis livres em B é ligada em A e $\Gamma\vdash \forall v_jA$, então $\Gamma\vdash A\big[@[v_i:B]/v_j\big]$.

Corolário 1.31. REGRA DE GENERALIZAÇÃO EXISTENCIAL III ($I\exists^{\dagger\dagger}$). Se nenhuma das variáveis livres em B é ligada em A e $\Gamma\vdash A\big[@[v_i:B]/v_j\big]$, então $\Gamma\vdash \exists v_jA$.

Exemplo. Como um exemplo simples da aplicação de $E\forall^{\dagger\dagger}$ e $I\exists^{\dagger\dagger}$ em uma demonstração, mostramos que $\vdash \exists x\, x=\#[u:u\neq u]$.

Demonstração: Pelo axioma *L*7, temos $\forall x\, x=x$. Aplicando a regra $E\forall^{\dagger\dagger}$, inferimos $\#[u:u\neq u]=\#[u:u\neq u]$ e, por $I\exists^{\dagger\dagger}$, obtemos então $\exists x\, x=\#[u:u\neq u]$.

■

Por fim, como consequência do Teorema 1.23 e da regra $E\forall^{\dagger\dagger}$, temos versões das regras de abstração e redução para "termos" formados por @ e $[_{\ldots}: __]$:

Corolário 1.32. Se nenhuma das variáveis livres em B é ligada em A, então:

(i) REDUÇÃO II. ($RED^{\dagger}$). Se $\Gamma \vdash [v_j : A](@[v_i : B])$, então $\Gamma \vdash A\big[@[v_i : B]/v_j\big]$.

(ii) ABSTRAÇÃO II. ($ABS^{\dagger}$). Se $\Gamma \vdash A\big[@[v_i : B]/v_j\big]$, então $\Gamma \vdash [v_j : A](@[v_i : B])$.

Com as regras $RED, ABS, RED^{\dagger}, ABS^{\dagger}, E\forall^{\dagger}, I\exists^{\dagger}, E\forall^{\dagger\dagger}$ e $I\exists^{\dagger\dagger}$, não será preciso empregar explicitamente o axioma da compreensão $L10$ em qualquer demonstração formulada nos capítulos subsequêntes, embora ele tenha sido necessário para prová-las – como evidencia uma inspeção cuidadosa das provas dos teoremas 1.23 e 1.26, a partir dos quais as regras foram obtidas. Contudo, juntamente com a prova do Teorema de Frege no Capítulo 2, a definicão contextual do operador de abstração em $\mathcal{D}_2$ estabelece a sua dispensabilidade no teorema: como $[_{\ldots}: __]$ não constitui um símbolo lógico genuíno da linguagem da Aritmética de Frege e, em geral, de quaisquer linguagens de segunda ordem caracterizadas na seção 1.1, e como as regras para o operador são derivadas a partir dos axiomas em $\mathrm{AL}(\Sigma_{\mathbf{F}})$, o Teorema de Frege poderia, em princípio, ser provado sem fazer qualquer uso de $[_{\ldots}: __]$, de maneira que $L10$ seria extensivamente utilizado em praticamente todas as demonstrações envolvidas. Entretanto, uma tal prova seria consideravelmente mais trabalhosa e menos intuitiva do que aquela a ser apresentada a seguir, na qual operador é utilizado.

É importante enfatizar que as regras $RED, ABS, E\forall^{\dagger}, I\exists^{\dagger}$ são válidas em quaisquer teorias cujo sistema dedutivo de base é $\mathcal{D}_2$; e que $RED^{\dagger}, ABS^{\dagger}, E\forall^{\dagger\dagger}$ e $I\exists^{\dagger\dagger}$ são válidas em quaisquer teorias que, além disso, possuem um operador @ que rebaixa a ordem dos termos aos quais se aplica a 1. Isso significa que essas regras podem ser utilizadas em toda e qualquer teoria fregeana e não se restringem apenas à Aritmética de Frege.

1.4 Teorias, Definições e Interpretações

Com o sistema dedutivo $\mathcal{D}_2$, definiremos uma teoria (de segunda ordem) da maneira usual como um conjunto de fórmulas fechado sob a relação de consequência lógica sintática e enunciaremos algumas definições de conceitos gerais que permitirão relacionar entre si as teorias a serem tratadas nos próximos capítulos. A maior parte das definições e teoremas desta seção são inteiramente rotineiros e não diferem das definições e teoremas correspondentes para lógica de primeira ordem, mas optamos por incluí-los neste capítulo com o único intuito de *fixar* os resultados metateóricos sobre teorias formais a serem empregados posteriormente. Contudo, no caso particular do conceito de *interpretação*, que desempenhará um papel importante no Capítulo 3, as definições e teoremas relevantes apresentam algumas diferenças significativas em relação à lógica de primeira ordem e, por essa razão, será oportuno fornecer um tratamento mais detalhado desse conceito a seguir.

Definição 1.33. Seja Σ uma assinatura de segunda ordem. Uma Σ-*teoria* (*de segunda ordem*) é um conjunto $K \subseteq \mathrm{F}(\Sigma)$ tal que, para toda $A \in \mathrm{F}(\Sigma)$, se $K \vdash A$, então $A \in K$.

Devido ao fato de $\vdash$ possuir a propriedade de *reflexividade* (se $A \in \Gamma$, então $\Gamma \vdash A$), uma teoria de segunda ordem é um conjunto de fórmulas para o qual vale a seguinte identidade:

$$K = \{A \in \mathrm{F}(\Sigma) : K \vdash A\}$$

Se $A \in K$ ou, equivalentemente, se $K \vdash A$, convencionaremos escrever $\vdash_K A$ e diremos que A é um *teorema* de K. Também utilizaremos a notação $\Gamma \vdash_K A$ para expressar que $\Gamma \cup K \vdash A$. Note-se que, pelo teorema $A \to (\neg A \to B)$ da lógica proposicional, se K é uma teoria inconsistente, então $K = \mathrm{F}(\Sigma)$, ou seja, todas as fórmulas da linguagem são teoremas de K.

Definição 1.34. Seja $\Gamma \subseteq \mathrm{F}(\Sigma)$. A *teoria gerada por* Γ (ou simplesmente a *teoria de* Γ) é o conjunto $\{A \in \mathrm{F}(\Sigma) : \Gamma \vdash A\}$.

É imediato provar que a teoria gerada por um conjunto Γ de fórmulas é, de fato, uma teoria de segunda ordem (a prova procede por indução sobre o comprimento de uma derivação a partir de $\{A \in \mathrm{F}(\Sigma) : \Gamma \vdash A\}$). Se existe um conjunto decidível que gera uma teoria, ela é

dita ser *axiomatizável*. Note-se que, como os axiomas em $\mathcal{D}_2$ são logicamente válidos e *MP* preserva verdade, se uma estrutura $\mathfrak{A}$ é modelo de Γ, então $\mathfrak{A}$ é modelo da teoria K gerada por Γ, ou seja, $\mathfrak{A}$ é modelo de todos os teoremas de K.

Nos próximos capítulos, ao definir uma Σ-teoria particular K, fixaremos apenas um conjunto que gera K (todas as teorias com as quais trabalharemos são, de fato, axiomatizáveis), nos referindo aos seus elementos por *axiomas próprios de K*, em oposição aos *axiomas lógicos de K* em $\mathrm{AL}(\Sigma)$; e, para preservar a aplicabilidade da regra de generalização universal $I\forall$ (Teorema 1.19) a todos os teoremas de uma teoria, tomaremos o cuidado de fixar seus axiomas próprios como fórmulas fechadas, de modo que se $K = \{A \in \mathrm{F}(\Sigma) : \Gamma \vdash A\}$, poderemos sempre inferir $\vdash_K \forall \mathbf{v} B$ ($\Gamma \vdash \forall \mathbf{v} B$) a partir de $\vdash_K B$ ($\Gamma \vdash B$). Todavia, convencionaremos omitir os quantificadores universais prenexos ao enunciarmos axiomas e teoremas de uma teoria formal: assumiremos que se $\mathbf{v}_1, \ldots, \mathbf{v}_n$ são, na ordem de ocorrência, todas as variáveis livres em uma fórmula A enunciada como um axioma ou teorema, então o axioma ou teorema em questão é, na verdade, a fórmula $\forall \mathbf{v}_1 \ldots \forall \mathbf{v}_n A$.

Os conceitos de *extensão* e *extensão conservativa* são particularmente úteis para relacionar diferentes teorias:

Definição 1.35. Sejam Σ e Σ' assinaturas de segunda ordem tais que $\Sigma \subseteq \Sigma'$, K uma Σ-teoria e K' uma Σ'-teoria.

(i) K' é uma *extensão* de K (ou K é uma *subteoria* de K') se e somente se $K \subseteq K'$. Equivalentemente, se e somente se, para toda $A \in \mathrm{F}(\Sigma)$, se $\vdash_K A$, então $\vdash_{K'} A$.

(ii) K' é uma *extensão conservativa* de K se e somente se K' é uma extensão de K e $\mathrm{F}(\Sigma) \cap K' \subseteq K$. Equivalentemente, se e somente se K' é uma extensão de K e, para toda $A \in \mathrm{F}(\Sigma)$, se $\vdash_{K'} A$, então $\vdash_K A$.

Portanto, K' é uma extensão de K se todos os teoremas de K são teoremas de K'; e K' é uma extensão conservativa de K se K' é uma extensão de K e se todas as fórmulas *da linguagem de K* que são teoremas de K' também são teoremas de K – o que significa que, embora inclua K, K' não possui ou não permite provar novos teoremas expressos na linguagem de K, mas apenas, possivelmente, novos teoremas cuja formulação envolva ao menos um

símbolo próprio de K' (note-se que se duas teorias compartilham a mesma linguagem e se uma delas é uma extensão conservativa da outra, então elas são, na verdade, idênticas).

O seguinte teorema afirma que, dadas duas teorias K e K', se todos os axiomas próprios de K são teoremas de K', então K' é uma extensão de K:

Teorema 1.36. Sejam Σ e Σ' assinaturas de segunda ordem tais que $\Sigma \subseteq \Sigma'$ e K' uma Σ'-teoria. Sejam $\Gamma \subseteq \mathrm{F}(\Sigma)$ e K a Σ-teoria de Γ. Se $\Gamma \subseteq K'$, então K' é uma extensão de K.

Prova: Suponha que $\vdash_K A$. Logo, $\Gamma \vdash A$ e, como $\Gamma \subseteq K'$, pela *monotonicidade* de $\vdash$ (se $\Delta \vdash B$ e $\Delta \subseteq \Delta'$, então $\Delta' \vdash B$), obtemos $K' \vdash A$, ou seja, $\vdash_{K'} A$.

■

Assim, para mostrar que os teoremas de K são teoremas de K' é suficiente provar os axiomas próprios de K em K' (contanto que K seja, de fato, uma teoria axiomatizável). Esse resultado será utilizado na prova do Teorema de Frege para mostrarmos que absolutamente todos os teoremas de uma formalização particular da aritmética elementar podem ser provados em **AF** com o auxílio de definições suplementares. O procedimento de enriquecer a linguagem de uma dada teoria com definições também pode, por sua vez, ser tecnicamente caracterizado em termos dos conceitos de extensão e extensão conservativa.

De um ponto de vista estritamente formal, uma definição em uma teoria K será tratada aqui simplesmente como um axioma adicional D introduzido ao conjunto de axiomas próprios de K no qual ocorra um novo símbolo não lógico α, de modo que D será denominada uma *definição de α em K*. A rigor, o procedimento de definir um símbolo em K mediante D consistirá, portanto, em considerar uma extensão K' de K cuja linguagem inclui o símbolo α. No entanto, embora se comporte inferencialmente como um axioma, uma definição deve atender a determinadas condições que a distinguem de um axioma no sentido usual: (i) ela não deve permitir provar novos teoremas na linguagem da teoria original em que é introduzida; e (ii) deve garantir a eliminabilidade do símbolo definido em todos os contextos sentenciais em que ele ocorra ou possa ocorrer. A definição abaixo caracteriza, em termos de relações entre teorias, essas condições, comumente denominadas *critério de conservatividade* (ou *não criatividade*) e *critério de eliminabilidade*:

Definição 1.37. Sejam Σ e Σ' assinaturas de segunda ordem tais que $\Sigma \subseteq \Sigma'$, K uma Σ-teoria e K' uma Σ'-teoria. K' é uma *extensão definicional* de K se e somente se

(i) K' é uma extensão conservativa de K (*critério de conservatividade*); e

(ii) Para toda $A \in \mathrm{F}(\Sigma')$, existe $B \in \mathrm{F}(\Sigma)$ tal que $A \leftrightarrow B \in K'$, ou seja, tal que $\vdash_{K'} A \leftrightarrow B$ (*critério de eliminabilidade*).

De acordo com a cláusula (i), o critério de conservatividade, quando introduzida em uma teoria K, uma definição D de um símbolo α deve resultar uma extensão conservativa K' de K; esse critério expressa a exigência de que, diferentemente de axiomas genuínos, definições devem ser neutras em relação às teorias em que são introduzidas e não podem permitir provar novos teoremas nas linguagens dessas teorias, mas apenas teoremas que, por assim dizer, dizem respeito unicamente ao comportamento do símbolo definido. A cláusula (ii), o critério de eliminabilidade, exige que, para cada fórmula A da linguagem de K' em que α ocorra, deve existir um fórmula B da linguagem da teoria original K que seja equivalente a A em K'; em outras palavras, (ii) exige que *todos* os contextos sentenciais em que α ocorra possam ser "reescritos" ou "traduzidos" em K' em contextos sentenciais nos quais apenas símbolos da linguagem de K estejam presentes ou, alternativamente, que α seja sempre dedutivamente eliminável em favor dos símbolos primitivos de K[16].

Um aspecto importante do conceito de extensão definicional consiste no fato de o critério de conservatividade implicar que qualquer extensão definicional de uma teoria consistente K preserva a consistência de K, o que significa que definições não podem, por si sós, resultar contradições:

Teorema 1.38. Sejam Σ e Σ' assinaturas de segunda ordem tais que $\Sigma \subseteq \Sigma'$, K uma Σ-teoria e K' uma Σ'-teoria. Se K' é uma extensão conservativa de K, então K é consistente se e somente se K' é consistente.

Prova: Suponha que K' é inconsistente. Logo, $\vdash_{K'} B$, para toda $B \in \mathrm{F}(\Sigma')$. Em particular, $\vdash_{K'} B$, para toda $B \in \mathrm{F}(\Sigma)$ e, portanto, $\vdash_K B$, para toda $B \in \mathrm{F}(\Sigma)$. Logo, existe $C \in \mathrm{F}(\Sigma)$ tal

[16]Para uma explicação mais detalhada da teoria de definições, ver [36], Capítulo 8; e para uma discussão acerca dos pressupostos conceituais subjacentes aos critérios de conservatividade e eliminabilidade, ver [1].

que $\vdash_K C$ e $\vdash_K \neg C$: K é inconsistente. Suponha, por outro lado, que K é inconsistente. Logo, existe $B \in \mathrm{F}(\Sigma)$ tal que $\vdash_K B$ e $\vdash_K \neg B$. Como K' é uma extensão de K, $\vdash_{K'} B$ e $\vdash_{K'} \neg B$ e, portanto, K' é inconsistente.

■

Corolário 1.39. Sejam Σ e Σ' assinaturas de segunda ordem tais que $\Sigma \subseteq \Sigma'$, K uma Σ-teoria e K' uma Σ'-teoria. Se K é um subteoria de uma extensão definicional de K' e K' é consistente, então K é consistente.

Embora a Definição 1.33 caracterize o conceito de extensão definicional como uma relação estática entre teorias e não descreva efetivamente o *processo* de definir novos símbolos em uma teoria particular, nos próximos capítulos, iremos aderir à prática matemática corrente de apresentar definições sequencialmente: a partir de uma dada teoria inicial K_0, introduziremos uma série de definições $D_0, D_1, \ldots, D_n$ por etapas, cada uma das quais definirá um único novo símbolo α e resultará uma extensão definicional da teoria precedente. Desse modo, ao fim do processo, teremos uma sequência de teorias $K_0, K_1, \ldots, K_n, K_{n+1}$, na qual, para todo $1 \leq i \leq n$, K_{i+1} é uma extensão definicional de K_i que resulta da introdução de D_i ao conjunto de axiomas próprios de K_i. Assim, o procedimento de definir novos símbolos em K_0 será tratado como uma maneira de enriquecer a linguagem de K_0 gradualmente e possibilitará que símbolos previamente definidos sejam utilizados para definir novos símbolos em etapas subsequêntes. O seguinte teorema legitima esse expediente ao estabelecer que não apenas K_{n+1} será uma extensão definicional de K_n, ...; K_2, uma extensão definicional de K_1 e K_1, uma extensão definicional de K_0, mas também que K_{n+1} será uma extensão definicional da teoria inicial K_0:

Teorema 1.40. Sejam Σ_1, Σ_2 e Σ_3 assinaturas tais que $\Sigma_1 \subseteq \Sigma_2 \subseteq \Sigma_3$ e, para $1 \leq i \leq 3$, K_i uma Σ_i-teoria de segunda ordem. Se K_3 é uma extensão definicional de K_2 e K_2 é uma extensão definicional de K_1, então K_3 é uma extensão definicional de K_1.

Prova: Basta mostrar que (i) K_3 é um extensão conservativa de K_1 e que (ii) para toda $A \in \mathrm{F}(\Sigma_3)$, existe $C \in \mathrm{F}(\Sigma_1)$ tal que $\vdash_{K_3} A \leftrightarrow C$:

(i) Como $K_1 \subseteq K_2 \subseteq K_3$, K_3 é uma extensão de K_1. Para mostrar que K_3 não prova teoremas

adicionais na lingugem de K_1, suponha que $\vdash_{K_3} A$, em que $A \in \mathrm{F}(\Sigma_1)$. Como $\mathrm{F}(\Sigma_1) \subseteq \mathrm{F}(\Sigma_2)$, $A \in \mathrm{F}(\Sigma_2)$. Logo, $\vdash_{K_2} A$ (pois K_3 é uma extensão conservativa de K_2) e, portanto, $\vdash_{K_1} A$ (pois K_2 é uma extensão conservativa de K_1).

(ii) Seja $A \in \mathrm{F}(\Sigma_3)$. Logo, existe $B \in \mathrm{F}(\Sigma_2)$ tal que $\vdash_{K_3} A \leftrightarrow B$ e existe $C \in \mathrm{F}(\Sigma_1)$ tal que $\vdash_{K_2} B \leftrightarrow C$ (pela cláusula (ii) da Definição 1.33). Como $K_2 \subseteq K_3$, $\vdash_{K_3} B \leftrightarrow C$ e, portanto, $\vdash_{K_3} A \leftrightarrow C$.

■

No Capítulo 3, faremos uso de outro conceito importante para relacionar diferentes teorias de segunda ordem: o conceito de *interpretação*. Nesta dissertação, o papel de interpretações se restringirá apenas a possibilitar provar resultados de consistência relativa acerca de **AF** e de duas outras formalizações alternativas da aritmética e, por essa razão, trataremos desse conceito a seguir com o único objetivo de formular esses resultados precisamente[17].

Grosso modo, dadas duas teorias K e K', uma interpretação de K em K', no sentido estrito mencionado acima, consiste em uma maneira de mimetizar uma estrutura que é modelo de K em K' ou, em outras palavras, em uma maneira de replicar sintaticamente uma tal estrutura no interior da teoria K'. Em geral, estruturas são descritas utilizando-se um fragmento de uma linguagem natural específica, como o inglês ou, no nosso caso, o português, enriquecida por uma certa quantidade de simbolismo matemático; em grande medida, mas não exclusivamente, proveniente da teoria de conjuntos. Além disso, ao definir uma estrutura para uma linguagem formal, recorre-se frequentemente a uma teoria matemática não formalizada para caracterizar e manipular a estrutura em questão. Se a teoria matemática utilizada nessa caracterização for passível de formalização em uma teoria K', a discussão acerca da estrutura poderá ser inteiramente traduzida para a linguagem utilizada para formalizá-la – desde que a linguagem de K' seja, em certo sentido, suficientemente rica. O conceito de interpretação a seguir pretende regimentar e tornar precisa a noção de tradução envolvida nesse processo.

Dada duas assinaturas Σ e Σ', definimos, primeiramente, uma interpretação da linguagem

[17] As definições e teoremas abaixo foram adaptados para o sistema dedutivo $\mathcal{D}_2$ a partir das definições e teoremas apresentados em [35], §4.7 para a lógica de primeira ordem.

de Σ na linguagem de Σ' como função injetiva do conjunto de Σ-parâmetros no conjunto de Σ'-parâmetros que associa uma constante relacional distinguida de tipo $\langle 0 \rangle$ a $\forall$ e que, a cada símbolo não lógico de Σ, associa um símbolo de Σ' de categoria sintática e tipo apropriados:

Definição 1.41. Sejam $\Sigma = \langle \mathcal{R}, \mathcal{F} \rangle$ e $\Sigma' = \langle \mathcal{R}', \mathcal{F}' \rangle$ assinaturas de segunda ordem. Uma *interpretação de* $\mathcal{L}(\Sigma)$ *em* $\mathcal{L}(\Sigma')$ é uma função injetiva $I : [\Sigma] \to [\Sigma']$ tal que

(i) $\forall^I$ é uma constante relacional U em $\mathcal{R}'_{\langle 0 \rangle}$ (denominada o *domínio* (*inicial*) de I).

(ii) Para todo $\tau \in \mathcal{T}$, para toda $R \in \mathcal{R}_\tau$, $R^I \in \mathcal{R}'_\tau$.

(iii) Para todo $\tau \in \mathcal{T}$, para toda $f \in \mathcal{F}_\tau$, $f^I \in \mathcal{F}'_\tau$.

U é denominada o domínio de I pois, ao definirmos o correlato sintático da noção de estrutura envolvendo interpretações, ela desempenhará um papel análogo ao domínio $|\mathfrak{A}|$ de uma Σ-estrutura $\mathfrak{A}$. Note-se que a definição exige que a linguagem $\mathcal{L}(\Sigma')$ seja suficientemente rica para conter ao menos uma constante em $\mathcal{R}'_\tau$, para cada elemento de $\mathcal{R}_\tau$; uma constante em $\mathcal{F}'_\tau$, para cada elemento de $\mathcal{F}_\tau$, além de conter a constante relacional unária U em $\mathcal{R}'_{\langle 0 \rangle}$. Como veremos, nem todas as linguagens sob consideração atendem a essas exigências, mas, em muitas ocasiões, elas podem ser enriquecidas mediante definições de maneira a possibilitar formular a interpretação relevante.

Assim como utilizamos a notação X_τ para denotar o domínio $|\mathfrak{A}|_\tau$ de tipo τ de uma estrutura $\mathfrak{A}$ (Definição 1.9), dada uma interpretação I de $\mathcal{L}(\Sigma)$ em $\mathcal{L}(\Sigma)$, utilizaremos a notação U_τ para denotar o *domínio de tipo* τ *de* I:

Definição 1.42. Sejam $\Sigma = \langle \mathcal{R}, \mathcal{F} \rangle$ uma assinatura de segunda ordem, $U \in \mathcal{R}_{\langle 0 \rangle}$ e $t \in \mathrm{T}(\Sigma)_\tau$. Definimos a notação $U_\tau(t)$ recursivamente por

(i) $U_0(t) := U(t)$

(ii) Se $\tau = \langle \tau_1, \ldots, \tau_n \rangle$ e, para $1 \leq i \leq n$, $\mathbf{v}_i \in \mathcal{V}_{\tau_i}$, então

$$U_\tau(t) := \forall \mathbf{v}_1 \ldots \forall \mathbf{v}_n (t\mathbf{v}_1 \ldots \mathbf{v}_n \to U_{\tau_1}(\mathbf{v}_1) \wedge \cdots \wedge U_{\tau_n}(\mathbf{v}_n))$$

em que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ não ocorrem em t e $\mathbf{v}_i \neq \mathbf{v}_j$, para todo $1 \leq i, j \leq n$.

Exemplo.

1. $U_0(x) := Ux$

2. $U_0(\#F) := U(\#F)$

3. $U_{\langle 0\rangle}(F) := \forall x(Fx \to Ux)$

4. $U_{\langle 0,0\rangle}(R) := \forall x \forall y(Rxy \to Ux \wedge Uy)$

5. $U_{\langle\langle 0\rangle,\langle 0\rangle\rangle}(\approx) := \forall F \forall G(F \approx G \to \forall x(Fx \to Ux) \wedge \forall x(Gx \to Ux))$

Com a notação acima, definimos então quando uma interpretação de $\mathcal{L}(\Sigma)$ em $\mathcal{L}(\Sigma')$ é *interpretação de* $\mathcal{L}(\Sigma)$ *em* uma Σ'-teoria K', que corresponde à relação entre a linguagem $\mathcal{L}(\Sigma)$ e uma Σ-estrutura $\mathfrak{A}$:

Definição 1.43. Seja K' uma Σ'-teoria. Uma *interpretação de* $\mathcal{L}(\Sigma)$ *em* K' é uma interpretação I de $\mathcal{L}(\Sigma)$ em $\mathcal{L}(\Sigma')$ tal que

(i) $\vdash_{K'} \exists x Ux$.

(ii) Para todo $\tau \in \mathcal{T}$, para toda $R \in \mathcal{R}_\tau$, $\vdash_{K'} U_\tau(R^I)$.

(iii) Para todo $\tau \in \mathcal{T}$, para toda $f \in \mathcal{F}_{\langle \tau_1,\ldots,\tau_n,\tau\rangle}$,

$$\vdash_{K'} \forall \mathbf{v}_1 \ldots \forall \mathbf{v}_n(U_{\tau_1}(\mathbf{v}_1) \to \cdots \to U_{\tau_n}(\mathbf{v}_n) \to U_\tau(f^I \mathbf{v}_1 \ldots \mathbf{v}_n))$$

em que $\mathbf{v}_i \in \mathcal{V}_{\tau_i}$, para cada $1 \leq i \leq n$.

De maneira análoga à definição de uma Σ-estrutura na seção 1.2 (Definição 1.10), a cláusula (i) exige que o domínio inicial U de I seja não vazio; a cláusula (ii), que todas as constantes em $\mathcal{R}'_\tau$ nos quais os elementos de $\mathcal{R}_\tau$ são interpretados por I satisfaçam $U_\tau(\ldots)$ em K'; e a cláusula (iii), que U seja fechada sob todas as constantes funcionais de $\mathcal{F}'_\tau$ nas quais os elementos de $\mathcal{F}_\tau$ são interpretadas por I.

Dados uma intepretação I de $\mathcal{L}(\Sigma)$ em K', um termo $t \in \mathbf{T}(\Sigma)$ e uma fórmula $A \in \mathrm{F}(\Sigma)$, a *interpretação de* t *por* I, denotada por t^I, é o termo em $\mathbf{T}(\Sigma')$ obtido pela substituição de todo símbolo $\alpha \in [\Sigma]$ que ocorre em t por α^I; e a *intepretação de* A *por* I, denotada por $A^{(I)}$, a fórmula em $\mathrm{F}(\Sigma')$ obtida pela substituição de todo termo t em A por t^I, pela relativização

de todos os quantificadores universais $\forall \mathbf{v}^\tau$ em A a $U_\tau(\mathbf{v}^\tau)$ e pela relativização da fórmula resultante a $U_\tau(\mathbf{v}^\tau)$, para toda variável $\mathbf{v}^\tau$ livre em A:

Definição 1.44. Sejam I uma interpretação de $\mathcal{L}(\Sigma)$ em K' e $A \in \mathrm{F}(\Sigma)$. Definimos a notação A^I recursivamente por

(i) Se A é $tt_1 \dots t_n$, então A^I é $t^I t_1^I \dots t_n^I$.

(ii) Se A é $t_1 = t_2$, então A^I é $t_1^I = t_2^I$.

(iii) Se A é $\neg B$, então A^I é $\neg B^I$.

(iv) Se A é $B \to C$, então A^I é $B^I \to C^I$.

(v) Se A é $\forall \mathbf{v}^\tau B$, então A^I é $\forall \mathbf{v}^\tau (U_\tau(\mathbf{v}^\tau) \to B^I)$

Sejam $\mathbf{v}_1, \dots, \mathbf{v}_n$ todas as variáveis livres em A (na ordem de ocorrência), em que $\mathbf{v}_i \in \mathcal{V}_{\tau_i}$, para todo $1 \leq i \leq n$. A *interpretação $A^{(I)}$ de A por I* é a fórmula

$$U_{\tau_1}(\mathbf{v}_1) \to \dots \to U_{\tau_n}(\mathbf{v}_n) \to A^I$$

Por fim, definimos quando uma interpretação de $\mathcal{L}(\Sigma)$ em K' é uma interpretação de uma Σ-teoria K em K', que corresponde à relação entre K e uma Σ-estrutura $\mathfrak{A}$, quando $\mathfrak{A}$ é modelo de K:

Definição 1.45. Sejam K uma Σ-teoria e K' uma Σ'-teoria. Uma *interpretação* de K *em* K' é uma interpretação I de $\mathcal{L}(\Sigma)$ em K' tal que tal que $\vdash_{K'} A^{(I)}$, para todo teorema A de K. Se existe uma interpretação de K em K', K é dita ser *interpretável* em K'.

Portanto, uma interpretação de K em K' é uma interpretação de $\mathcal{L}(\Sigma)$ que preserva todos os teoremas de K em K'. É fácil notar que se K é interpretável em K', então a consistência de K' implica na consistência de K:

Teorema 1.46. Se K é interpretável em K' e K' é consistente, então K é consistente.

Prova: Suponha que K é inconsistente e seja I uma interpretação de K em K'. Logo, existe uma fórmula fechada $A \in \mathrm{F}(\Sigma)$ tal que $\vdash_K A$ e $\vdash_K \neg A$ (pois $K = \mathrm{F}(\Sigma)$). Pela definição acima, $\vdash_{K'} A^{(I)}$ e $\vdash_{K'} (\neg A)^{(I)}$. Mas, como nenhuma variável é livre em A, $A^{(I)}$ é A^I e $(\neg A)^{(I)}$ é $\neg A^I$. Portanto, $\vdash_{K'} A^I$ e $\vdash_{K'} \neg A^I$: K' é inconsistente.

■

Ao empregar o resultado acima no Capítulo 3, faremos uso do seguinte teorema, que estabelece que se uma interpretação I de $\mathcal{L}(\Sigma)$ em K' preserva todos os axiomas próprios de uma Σ-teoria axiomatizável K, então I preserva todos os teoremas de K:

Teorema 1.47. TEOREMA DA INTERPRETAÇÃO. Sejam K uma Σ-teoria axiomatizável, K' uma Σ'-teoria e I uma interpretação de $\mathcal{L}(\Sigma)$ em K'. Se $\vdash_{K'} A^{(I)}$, para todo axioma próprio A de K, então I é uma interpretação de K em K'.

Prova: Apêndice A2.

■

Capítulo 2

O Teorema de Frege

> I do not know whether Frege realized that Hume's principle plus the logic of the *Begriffsschrift* was all he used or needed. Perhaps not; *he* would have had no reason to value the observation. In any event, it is a pity that the derivability of arithmetic from Hume's principle isn't known as *Frege's Theorem*.
>
> George Boolos

Agora que definimos a sintaxe e a semântica da lógica de segunda ordem e caracterizamos o conceito de *teoria* de uma maneira apropriada para formalizar a Aritmética de Frege, podemos definí-la precisamente e formular o teorema homônimo tecnicamente como uma relação entre **AF** e a *Aritmética de Dedekind-Peano* (**ADP**), uma teoria de segunda ordem que formaliza a aritmética elementar. Do modo como será provado neste capítulo, o teorema estabele que o vocabulário primitivo de **ADP** pode ser definido em **AF** e seus axiomas próprios, demonstrados a partir das definições fornecidas e do Princípio de Hume. A prova será desenvolvida de uma maneira que concorda quase inteiramente com o esboço apresentado por Frege em *Gl*, §§70-83, e, no decorrer do capítulo, nos referiremos às seções do livro onde

as definições e teoremas em questão foram enunciados. Ao indicarmos o percurso adotado no *Grundlagen*, pretendemos fazer jus à epígrafe acima, na qual Boolos cunha o nome 'Teorema de Frege', e evidenciar que, mesmo quando desenvolvido em um arcabouço formal com profundas diferenças técnicas e conceituais em relação aos sistemas do *Begriffsschrift* e do *Grundgesetze*, o resultado ainda assim pode ser justificadamente creditado a Frege.

Começaremos por definir as teorias **AF** e **ADP**, apresentando, em seguida, uma prova detalhada do Teorema de Frege. Ao definirmos o vocabulário primitivo de **ADP** em **AF**, consideraremos algumas extensões definicionais de **AF**, assumindo, sem prova, que todas as definições a seguir são logicamente corretas – i.é., que sua introdução ao conjunto de axiomas próprios de **AF** resulta, de fato, uma extensão definicional da teoria. Como a categoria sintática e o tipo de um símbolo definido podem ser facilmente apreendidos de sua definição, não especificaremos as assinaturas resultantes de cada extensão, nos referindo a elas e às extensões correspondentes indistintamente por '$\Sigma_{\mathbf{F}}$' e '**AF**', respectivamente.

2.1 Aritmética de Frege

Repetimos aqui a definição de $\Sigma_{\mathbf{F}}$, a assinatura da Aritmética de Frege, apresentada no Capítulo 1 a título de exemplo, remetendo o leitor aos comentários explicativos da seção 1.1:

Definição 2.1. $\Sigma_{\mathbf{F}}$ é o par $\langle \mathcal{R}^{\mathbf{F}}, \mathcal{F}^{\mathbf{F}} \rangle$ tal que

(i) Para todo τ, $\mathcal{R}^{\mathbf{F}}_{\tau} = \varnothing$.

(ii) $\mathcal{F}^{\mathbf{F}}_{\langle\langle 0\rangle,0\rangle} = \{\#\}$.

(iii) Para todo $\tau \neq \langle\langle 0\rangle,0\rangle$, $\mathcal{F}^{\mathbf{F}}_{\tau} = \varnothing$.

Termos individuas formados pelo operador de cardinalidade #, como $\#F$, serão lidos como 'o número de Fs' ou 'o número (cardinal) do conceito F'; também estenderemos essa terminologia para "termos" individuais que resultam da aplicação do operador de cardinalidade a "termos" de segunda ordem formados pelo abstrator $[\ldots\colon __]$: $\#[u : u \neq u]$, por exemplo, será lido como 'o número do conceito 'ser diferente de si próprio".

Em muitos ocasiões, utilizamos previamente as expressões 'Aritmética de Frege' ou '**AF**' para nos referirmos à teoria semi-formalmente descrita na literatura como o resultado da introdução do Princípio de Hume a um sistema dedutivo de lógica de segunda ordem padrão. Com a definição de *teoria de segunda ordem* da seção 1.4, podemos agora definí-la precisamente:

Definição 2.2. A *Aritmética de Frege* (**AF**) é a $\Sigma_{\mathbf{F}}$-teoria cujo único axioma próprio é o *Princípio de Hume*:

(*PH*)

$$\#F = \#G \leftrightarrow \exists R(\forall x(Fx \rightarrow \exists y(Gy \wedge Rxy)) \wedge \forall y(Gy \rightarrow \exists x(Fx \wedge Rxy)) \wedge$$

$$\forall x \forall y \forall z \forall w(Rxy \wedge Rzw \rightarrow (x = z \leftrightarrow y = w)))$$

A formulação de *PH* acima difere consideravalmente daquela apresentada na Introdução: $\#F = \#G \leftrightarrow F \approx G$ (em que, relembremos, $F \approx G$ é lido como 'F é *equinumérico* a G'). Na ocasião, mencionamos que o símbolo $\approx$ poderia ser definido em termos do vocabulário lógico da *LSO* e, para recuperar aquela formulação, propomos então as seguintes definições em **AF**:

($D_{\approx_R}$)

$$F \approx_R G \leftrightarrow \forall x(Fx \rightarrow \exists y(Gy \wedge Rxy)) \wedge \forall y(Gy \rightarrow \exists x(Fx \wedge Rxy)) \wedge$$

$$\forall x \forall y \forall z \forall w(Rxy \wedge Rzw \rightarrow (x = z \leftrightarrow y = w))$$

($D_{\approx}$)

$$F \approx G \leftrightarrow \exists R\, F \approx_R G$$

$F \approx G_R$ afirma que R correlaciona bijetivamente os objetos que são F aos objetos que são G: a primeira conjunta do lado direito de $\leftrightarrow$ expressa que, a cada F, R relaciona ao menos um G, e a segunda, que cada G é R-relacionado a ao menos um F. A terceira conjunta expressa que R e sua conversa são relações funcionais e condensa em uma única fórmula as fórmulas

$$\forall x \forall y \forall z(Rxy \wedge Rxz \rightarrow y = z)$$

e

$$\forall x \forall y \forall z (Ryx \wedge Rzx \rightarrow y = z)$$

Na presença da terceira conjunta, a primeira e a segunda expressam, portanto, que, a cada F, R relaciona exatamente um G, e que cada G é R-relacionado a exatamente um F, respectivamente[1].

Ao definir **AF**, optamos pela formulação expandida de *PH* para fazer jus ao fato de que $\Sigma_{\mathbf{F}}$ possui apenas um único símbolo não lógico (#); e, do ponto de vista metodológico, não poderíamos propor as definições $D_{\approx_R}$ e $D_{\approx}$ e formular *PH* em sua forma abreviada na definição de **AF** sem antes estabelecer a teoria em que o símbolo $\approx$ deveria ser definido; no nosso caso, a própria **AF**. De agora em diante, no entanto, designaremos por 'Princípio Hume' ou '*PH*' a fórmula $\#F = \#G \leftrightarrow F \approx G$ (ou mais precisamente, a fórmula $\forall F \forall G(\#F = \#G \leftrightarrow F \approx G)$):

(*PH*)

$$\#F = \#G \leftrightarrow F \approx G$$

Os seguintes fatos sobre $\approx$ são facilmente provados em **AF** a partir de $D_{\approx_R}$ e $D_{\approx}$, os quais, conjuntamente, estabelecem que *equinumericidade* é uma relação de equivalência:

($\approx_1$) $F \approx F$

($\approx_2$) $F \approx G \rightarrow G \approx F$

($\approx_3$) $F \approx G \wedge G \approx H \rightarrow F \approx H$[2]

Definindo-se $F \equiv G$ (F é *coextensivo* a G) por

[1] Uma definição alternativa de $\approx_R$ seria $F \approx_R G \leftrightarrow \forall x(Fx \rightarrow \exists! y(Gy \wedge Rxy)) \wedge \forall y(Gy \rightarrow \exists! x(Fx \wedge Rxy))$, em que $\exists! \mathbf{v}_1 A$ é definido da maneira usual como $\exists \mathbf{v}(A \wedge \forall \mathbf{v}_2(A[\mathbf{v}_2/\mathbf{v}_1] \rightarrow \mathbf{v}_1 = \mathbf{v}_2))$.

[2] Para provar $\approx_1$, basta utilizar o axioma *L*9 e o teorema lógico $\forall x(Fx \leftrightarrow Fx)$ – $\approx_1$ também pode ser provado em **AF** sem o recurso a *L*9, mostrando-se $F \approx_I F$, em que $I := [uu_1 : u = u_1]$ é a *relação identidade*. Para provar $\approx_2$, é suficiente mostrar $G \approx_{R^{-1}} F$, em que $F \approx_R G$ e $R^{-1} := [uu_1 : Ru_1u]$ é a *relação inversa* de R; e para provar $\approx_3$, é suficiente mostrar $F \approx_{S \circ R} H$, em que $F \approx_R G$, $G \approx_S H$ e $S \circ R := [uu_1 : \exists z(Ruz \wedge Szu_1)]$ é a *relação composta* de S e R.

$(D_{\equiv})$

$$F \equiv G \leftrightarrow \forall x(Fx \leftrightarrow Gx)$$

prova-se, ademais, que se dois conceitos são coextensivos, eles são também equinuméricos, um fato que será algumas vezes utilizado na prova do Teorema de Frege:

$(\approx_4)$ $F \equiv G \rightarrow F \approx G$

2.2 Aritmética de Dedekind-Peano

Na Introdução, mencionamos que a tese logicista originalmente formulada por Frege no *Grundlagen* consistia na tese de que todas as verdades aritméticas são analíticas, o que, segundo a caracterização fregeana do conceito de analiticidade em *Gl*, §3, significa que elas poderiam ser provadas unicamente a partir de leis lógicas gerais e definições apropriadas. Como as proposições verdadeiras da aritmética são infinitas em número, para demonstrar essa tese, seria impossível, entretanto, demonstrar a analicidade de cada uma delas em particular. Era preciso, portanto, selecionar um número finito e reduzido de leis aritméticas a partir das quais, presumivelmente, todas as outras proposições aritméticas verdadeiras poderiam ser provadas, mostrando-se apenas a analicidade dessas leis[3]. Mas quais seriam, segundo Frege, as leis aritméticas a partir das quais todas as outras proposições poderiam ser provadas? Embora ele não ofereça uma resposta explícita a essa pergunta, é possível extrair, a partir dos princípais teoremas provados no *Grundlagen* e no *Grundgesetze*, uma axiomatização para aritmética relativamente similar à axiomatização formulada por Dedekind em [13] e elaborada formalmente por Peano em [31]; e cujas formalizações contemportâneas são algumas vezes denominadas *Aritmética de Dedekind-Peano*[4].

[3]Em virtude do primeiro teorema de incompletude de Gödel, hoje sabemos que nenhuma axiomatização recusiva e, portanto, finita da aritmética é suficiente para provar absolutamente todas as verdades da aritmética.

[4]Na literatura, não existe consenso terminológico em relação ao nome dado à teoria aritmética a ser apresentada a seguir: Boolos [7], por exemplo, se refere a seus axiomas por 'Peano postulates' e Heck [25] denomina-a 'second-order Peano arithmetic', utilizando a sigla **PA2**. O nome 'Aritmética de Peano de Segunda Ordem', en-

A Aritmética de Dedekind-Peano possui diversas formulações alternativas que se distinguem pelo número de axiomas ou pelas noções e símbolos primitivos envolvidos e, para provar o Teorema de Frege como uma relação entre e **AF** e essa teoria, escolheremos uma axiomatização que facilitará particularmente esse trabalho: adotaremos o conjunto de axiomas apresentado por Heck em [25], p. 589, com uma ligeira modificação, porém, no axioma de indução matemática. Tal como será formulada aqui, a teoria possui apenas três símbolos não lógicos: a constante individual **0**, para o número 0, e as constantes relacionais *Nat*, para *número natural*, e *Pred*, para a relação *predecessor* (*imediato*)[5]:

Definição 2.3. A assinatura da Aritmética de Dedekind-Peano é o par $\Sigma_{\mathbf{DP}} = \langle \mathcal{R}^{\mathbf{DP}}, \mathcal{F}^{\mathbf{DP}} \rangle$ tal que

(i) $\mathcal{R}_0^{\mathbf{DP}} = \{\mathbf{0}\}$.

(ii) $\mathcal{R}_{\langle 0 \rangle}^{\mathbf{DP}} = \{Nat\}$.

(iii) $\mathcal{R}_{\langle 0,0 \rangle}^{\mathbf{DP}} = \{Pred\}$.

(iv) Para todo $\tau \neq 0, \langle 0 \rangle, \langle 0,0 \rangle$, $\mathcal{R}_\tau^{\mathbf{DP}} = \varnothing$.

(v) Para todo τ, $\mathcal{F}_\tau^{\mathbf{DP}} = \varnothing$.

Assim, os termos individuais de $\Sigma_{\mathbf{DP}}$ são **0** e as variáveis de primeira ordem; e *Pred* e *Nat* são as únicas constantes relacionais (de aridade 1 e 2, respectivamente) da linguagem, as quais se aplicam apenas a termos individuais para formar fórmulas atômicas. $\Sigma_{\mathbf{DP}}$ não possui constantes funcionais e, assim como $\Sigma_{\mathbf{F}}$, não possui também termos de terceira ordem. Como

tretanto, se consolidou, em lógica-matemática, para designar uma teoria de segunda ordem para aritmética que apresenta algumas diferenças em relação à axiomatização adotada aqui e, por essa razão, iremos aderir aqui ao nome 'Aritmética de Dedekind-Peano', reservando 'Aritmética de Peano de Segunda Ordem' para aquela teoria, a ser introduzida no Capítulo 3.

[5]Embora *Nat* e *Pred* sejam sequências de, respectivamente, três e quatro letras no alfabeto romano, considerá-las-emos como expressões indecomponíveis em nosso formalismo (i.é., como símbolos). Ao adotar essa notação específica, pretendemos remeter o leitor às interpretações pretendidas desses símbolos, facilitando a leitura de fórmulas complexas em que eles ocorram. Uma alternativa mais usual seria escrever *N* no lugar de *Nat* e *P* no lugar de *Pred*.

exemplos de $\Sigma_{\mathbf{DP}}$-fórmulas temos Fx, Rxy, $x = y$, $F\mathbf{0}$, $Rx\mathbf{0}$, $x = \mathbf{0}$, $Nat(\mathbf{0})$, $Pred(\mathbf{0},\mathbf{0})$, $\neg Fx$, $\neg R\mathbf{00}$, $\forall x(\neg Pred(x,\mathbf{0}))$, $\forall x(Pred(\mathbf{0},x) \to Nat(x))$, $\exists F \forall x(Nat(x) \to \neg Fx)$.

Definição 2.4. A *Aritmética de Dedekind-Peano* (**ADP**) é a $\Sigma_{\mathbf{DP}}$-teoria cujos axiomas próprios são:

(**ADP**1) $Nat(\mathbf{0})$

(**ADP**2) $Nat(x) \wedge Pred(x,y) \to Nat(y)$

(**ADP**3) $Nat(x) \wedge Pred(x,y) \wedge Pred(x,z) \to y = z$

(**ADP**4) $Nat(y) \wedge Nat(z) \wedge Pred(y,x) \wedge Pred(z,x) \to y = z$

(**ADP**5) $\neg \exists x(Nat(x) \wedge Pred(x,\mathbf{0}))$

(**ADP**6) $Nat(x) \to \exists y Pred(x,y)$

(**ADP**7) $F\mathbf{0} \wedge \forall x \forall y(Nat(x) \wedge Pred(x,y) \to (Fx \to Fy)) \to \forall x(Nat(x) \to Fx)$

Na interpretação pretendida de **ADP** – na qual as variáveis individuais varrem números naturais, as variáveis relacionais varrem relações entre números naturais, *Nat* é o universo, $\mathbf{0}$ denota o número 0 e *Pred* denota a relação entre um número natural e o seu sucessor – os axiomas **ADP**1-7 expressam que

1. 0 é um número natural.
2. Se m é um número natural e precede n, então n é um número natural.
3. Se m é um número natural e precede n e n', então $n = n'$.
4. Se m e m' são números naturais e precedem n, então $m = m'$.
5. Nenhum número natural precede 0.
6. Se m é um número natural, então para algum n, m precede n.
7. Para todo conjunto X, se

   (i) 0 pertence a X; e

(ii) sempre que um número natural m precede n e pertence a X, n também pertence a X, então todo o número natural pertencem a X.

ADP2 estabelece, portanto, que *número natural* é fechado sob a relação *predecessor*; **ADP**3, **ADP**4 e **ADP**6, que *predecessor* é, respectivamente, funcional, injetiva e serial nos números naturais; e **ADP**7 é o *princípio de indução matemática*. Note-se que **ADP**3, **ADP**2 e **ADP**6 caracterizam a relação *predecessor* como uma função nos números naturais: ao interpretarmos $Pred(x,y)$ como 'y é sucessor de x', eles estabelecem que todo número natural possui um único sucessor natural. Note-se também que **ADP**3 e **ADP**4 poderiam ser condensados em $Nat(x) \wedge Nat(y) \wedge Nat(z) \wedge Nat(w) \rightarrow (x = z \leftrightarrow y = w)$; e **ADP**2 e **ADP**6, condensados em $Nat(x) \rightarrow \exists y(Nat(x) \wedge Pred(x,y))$, de modo que seria possível formular a teoria **ADP** com apenas cinco axiomas próprios. No entanto, optamos por considerá-los separadamente para facilitar a exposição da prova do Teorema de Frege a seguir.

2.3 A Prova do Teorema

No primeiro parágrafo deste capítulo, afirmamos que a prova do Teorema de Frege a ser apresentada aqui seria desenvolvida de uma maneira que concorda *quase* inteiramente com o esboço presente em *Gl*, §§70-83. Essa qualificação é motivada pela tese sustentada por Boolos e Heck em [9] de que Frege não possuía uma prova completa do teorema em 1884, ano da publicação do *Grundlagen*: a partir da análise de *Gl*, §§82-83, que contêm a indicação de como a prova da proposição de que todo número natural possui um sucessor (**ADP**6) deveria ser desenvolvida, e da seção 114 da Parte II do *Grundgesetze*, onde a derivação de uma versão formalizada dessa mesma proposição no sistema fregeano é explicada a título de introdução ao seu desenvolvimento técnico subsequênte, eles concluem que Frege não somente não apresentou a prova (ou o esboço da prova) daquela proposição no *Grundlagen* como também não sabia, à época, como fazê-lo; alegando, dentre outras razões, que, no segundo livro, ele considerou estar "corrigindo" a prova ensaiada no primeiro.

Não iremos expor o argumento completo elaborado por ambos os autores e muito menos

nos posicionar em relação a essa questão interpretativa, nos restringindo apenas a mencionar, na ocasião da demonstração de **ADP**6 em **AF**, os indícios textuais presentes no *Grundlagen* apontados em [9]. Porém, em vista da interpretação proposta por Boolos e Heck, apresentaremos uma prova de **ADP**6 que, embora não seja uma reprodução inteiramente fidedigna daquela presente no *Grundgesetze* ou da esboçada (talvez erroneamente) no *Grundlagen*, preserva as ideias centrais das partes correspondentes de ambos os livros. Nesse sentido, iremos expor o que eles denominam uma 'prova *conjectural* do Teorema de Frege' ([9], p. 327). Por fim, notamos ainda que a prova abaixo não foi baseada em nenhuma fonte preponderante, mas em diversas exposições distintas; além do próprio *Grundlagen*, citamos como principais referências [4], [9], [24] e [43].

Feitas essas considerações iniciais, passemos à prova propriamente dita, enunciando o teorema como a seguinte relação entre **ADP** e **AF**:

Teorema 2.5. TEOREMA DE FREGE. **ADP** é uma subteoria de uma extensão definicional de **AF**.

De acordo com o Teorema 1.36 do Capítulo 1, para provar o Teorema de Frege nessa formulação técnica particular, é suficiente provar que existe uma extensão definicional de **AF** da qual os axiomas **ADP**1-7 são teoremas.

2.3.1 Relação Ancestral e os Princípios de Indução Generalizada

Como primeira etapa no desenvolvimento da prova do Teorema de Frege, será preciso replicar em **AF** a famosa definição fregeana do conceito de *ancestral*, formulada formalmente na Parte III do *Begriffsshrift* (Definição 76) e retomada em linguagem vernacular em *Gl*, §79. Esse conceito desempenha um papel central no teorema, sendo utilizado na definição de *número natural* e em vários passos fundamentais da prova.

Intuitivamente, a relação ancestral de uma dada relação binária R pode ser entendida como a relação que ocorre entre x e y caso exista uma cadeia que ligue x a y cujos elos são determinados por R: $xRx_1, x_1Rx_2, \ldots, x_{n-1}Rx_n, x_nRy$. Assim, a relação ancestral de R nada mais é do que a própria série ou sequência gerada por R. Para tomarmos um exemplo sugerido pela

terminologia, considere-se a relação *ser pai de* e suponha-se que a é pai de b e c, que b é pai de d e que c é pai de e. Nesse caso, diremos que a é um ancestral, na relação de paternidade, de b, c, d e e, que b é um ancestral de d e que c é um ancestral de e – porém, nem b é um ancestral de e e nem c é um ancestral de d.

No *Begriffsshrift*, Frege pretendeu capturar ou explicar essa noção intuitva utilizando o formalismo da conceitografia. Sua proposta consistia em formular uma definição lógica do conceito de *ancestral* ou, mais precisamente, do conceito de *série* ou *sequência* (o termo 'ancestral' foi cunhado posteriormente por Russell e Whitehead em [39]), de maneira a exemplificar como noções matemáticas elementares e fatos gerais acerca delas, tidos como de natureza intuitiva, poderiam ser fundamentados unicamente na lógica pura. Essa proposta representa, talvez, a primeira expressão do logicismo fregeano[6].

A definição de Frege do conceito de ancestral é baseada sobre a ideia de uma propriedade ser hereditária em uma dada relação: grosso modo, x é dito um *R-ancestral* de y – na terminologia fregeana, *x precede y na série R* ou *y segue após x na série R* – se y possui todas as propriedades hereditárias em R que os descentes imediatos de x também possuem, sendo que as propriedades hereditárias são aquelas que se transmitem através da relação: se x é F e x e y são R-relacionados, então y também é F. Por exemplo, tomemos agora a relação *ser mãe* (*biológica*) *de* e consideremos que Maria é mãe de Luciana e Pedro, que Luciana é mãe de Lucas e que Pedro é pai de Ana. Assim, diremos que Maria é uma ancestral materna de Lucas, pois ele possui todas as propriedades hereditárias na relação de maternidade que foram herdadas pelos descendentes imediatos de Maria, ou seja, Luciana e Pedro. Por outro lado, Maria não é uma ancestral materna de Ana, pois Ana falha em possuir ao menos uma dessas propriedades: ela falha em possuir, digamos, o mesmo DNA mitocondrial de Maria, Luciana, Pedro e Lucas[7].

Para replicar a definição de *ancestral* em **AF**, introduziremos a constante funcional * de tipo $\langle\langle 0,0\rangle,\langle 0,0\rangle\rangle$ em $\Sigma_{\mathbf{F}}$, de modo que R^* denotará a *a relação ancestral de R*; e para replicar a definição de *ancestral fraco* (a seguir), introduziremos a constante funcional $^{*=}$ também de

[6]Para uma discussão acerca da presença de um "protologicismo" no *Begriffsschrift*, ver [2].

[7]O DNA mitocondrial é transmitido da mãe para os filhos e nunca do pai para os filhos.

tipo $\langle\langle 0,0\rangle,\langle 0,0\rangle\rangle$. * e $^{*=}$ se aplicam, portanto, a termos em $\mathrm{T}(\Sigma_{\mathbf{F}})_{\langle 0,0\rangle}$ – variáveis de segunda ordem binárias, por exemplo – e resultam termos em $\mathrm{T}(\Sigma_{\mathbf{F}})_{\langle 0,0\rangle}$. A definição de * recorre a outras duas definições auxiliares:

(D_{Her})

$$Her(F,R) \leftrightarrow \forall x_1 \forall x_2 (Rx_1x_2 \rightarrow (Fx_1 \rightarrow Fx_2))$$

(D_{In})

$$In(R,F,x) \leftrightarrow \forall x_1 (Rxx_1 \rightarrow Fx_1)$$

(D_*)

$$R^*xy \leftrightarrow \forall F(Her(F,R) \wedge In(R,F,x) \rightarrow Fy)$$

$Her(F,R)$, lido como 'F é *hereditário* em R' ou 'F é *R-heraditário*', expressa que F é fechado sob ou transmissível através de R; e $In(R,F,x)$, expressa que todos os objetos com os quais x mantém a relação R, os *R-descendentes imediatos* de x, são F[8]. De acordo com D_*, portanto, x é um *R-ancestral* de y se e somente se y é F, para todo F R-hereditário tal que todo R-descendente imediato de x é F. Notamos que $Rxy \rightarrow R^*xy$ e $R^*xy \wedge R^*yz \rightarrow R^*xz$ (fatos $Anc1$ e $Anc5$ abaixo) são obtidos diretamente de D_* e que esses fatos garantem que a definição fregeana captura aspectos essenciais da noção intuitiva de ancestral descrita acima, a saber, que (i) se x e y são R-relacionados, então y segue após x na série ou sequência gerada por R e que (ii) essa série é transitiva.

A relação *ancestral fraca de R*, definida por Frege no *Begriffsshrift* (Definição 99) e retomada em *Gl*, §81, é reproduzida em **AF** por

($D_{*=}$)

$$R^{*=}xy \leftrightarrow R^*xy \vee x = y$$

Ou seja, x é um *R-ancestral fraco* de y – na terminologia fregeana, *x pertence à série R que*

[8] D_{Her} corresponde à Definição 69 do *Begriffsshrift* e, em D_{In}, que não possui correspondente fregeana, o símbolo '*In*' deriva da palavra de língua inglesa 'inherit'.

termina com y/y pertence à série R que começa com x – se e somente se x é um R-ancestral de y ou $x = y$.

Abaixo encontram-se enunciados todos os fatos sobre * e $^{*=}$ que serão utilizados na prova do Teorema de Frege e cujas demonstrações não oferecem dificuldade[9]:

($Anc1$) $Rxy \to R^*xy$

($Anc2$) $Her(F,R) \wedge In(R,F,x) \to \forall y(R^*xy \to Fy)$

($Anc3$) $Fx \wedge Her(F,R) \to \forall y(R^*xy \to Fy)$

($Anc4$) $R^*xy \to \exists z Rzy$

($Anc5$) $R^*xy \wedge R^*yz \to R^*xz$

($Anc6$) $R^*xy \wedge Ryz \to R^*xz$

($Anc7$) $R^{*=}xx$

($Anc8$) $R^*xy \to R^{*=}xy$

($Anc9$) $Fx \wedge Her(F,R) \to \forall y(R^{*=}xy \to Fy)$

($Anc10$) $R^{*=}xy \wedge Ryz \to R^*xz$

($Anc11$) $R^{*=}xy \wedge Ryz \to R^{*=}xz$

($Anc12$) $R^*xy \to \exists z(R^{*=}xz \wedge Rzy)$

Dentre esses fatos, destacamos que $Anc3$ e $Anc9$ possuem um significado particularmente importante: restituindo o quantificador universal $\forall F$ a $Anc3$, temos a fórmula

$$\forall F(Fx \wedge Her(F,R) \to \forall y(R^*xy \to Fy))$$

que, em terminologia fregeana, expressa que todos os objetos que seguem após x na série R caem sob todo conceito F tal que x cai sob F e F é R-hereditário. Na presença do axioma da compreensão $L10$, ela representa um método de inferência que podemos considerar como

[9] Os fatos $Anc1$, $Anc3$ e $Anc5$ são provados na Parte III do *Begriffsshrift* como as proposições 91, 84 e 98, respectivamente; e os fatos $Anc2$, $Anc4$, $Anc9$, $Anc10$ e $Anc12$ são provados como os teoremas 123, 124, 144 e 141 da Parte I do *Grundgesetze*, respectivamente. As demonstrações de $Anc1$-12 encontram-se no Apêndice B1.

uma forma generalizada de indução matemática: para provar que todos os R-ancestrais de x possuem uma dada propriedade P, definida, digamos, pela fórmula $A \in \Sigma_{\mathbf{F}}$, é suficiente mostrar que (i) x possui P (passo base); e que (ii) P é R-hereditária (passo indutivo): se Rx_1x_2 e x_1 possui P (hipótese indutiva), então x_2 também possui P. Como consequência de *Anc*3, temos ainda uma segunda versão de indução generalizada, codificada em *Anc*9: para provar que todos os R-ancestrais fracos de x possuem P é também suficiente mostrar que (i) x possui P e que (ii) P é R-herditária. Denominemos *Anc*3 e *Anc*9 *princípio de indução generalizada 1* e *princípio de indução generalizada 2*, respectivamente.

Assim como todos os outros fatos enunciados acima, os princípios de indução 1 e 2 são obtidos unicamente das definições D_* e $D_{*=}$, não havendo nenhum recurso a *PH* em suas demonstrações. O mesmo se pode dizer de uma terceira versão de indução generalizada, obtida de *Anc*9 e da qual o axioma de indução **ADP**7 será um caso particular:

(*Anc*13) PRINCÍPIO DE INDUÇÃO GENERALIZADA 3

$$Fx \wedge \forall x_1 \forall x_2 (R^{*=}xx_1 \wedge Rx_1x_2 \rightarrow (Fx_1 \rightarrow Fx_2)) \rightarrow \forall y (R^{*=}xy \rightarrow Fy)$$

Demonstração: Suponha $Fx \wedge \forall x_1 \forall x_2 (R^{*=}xx_1 \wedge Rx_1x_2 \rightarrow (Fx_1 \rightarrow Fx_2))$. Instanciando G para $[u : R^{*=}xu \wedge Fu]$ na seguinte variação alfabética de *Anc*9:

$$Gx \wedge Her(G,R) \rightarrow \forall y (R^{*=}xy \rightarrow Gy)$$

obtemos

$$[u : R^{*=}xu \wedge Fu]x \wedge Her([u : R^{*=}xu \wedge Fu], R) \rightarrow \forall y (R^{*=}xy \rightarrow [u : R^{*=}xu \wedge Fu]y)$$

(i) Para provar $[u : R^{*=}xu \wedge Fu]x$, basta mostrar $R^{*=}xx \wedge Fx$, que é uma consequência imediata de *Anc*7 e da hipótese inicial.

(ii) Para provar $Her([u : R^{*=}xu \wedge Fu], R)$, suponha Rx_1x_2 e $[u : R^{*=}xu \wedge Fu]x_1$. Logo, $R^{*=}xx_1 \wedge Fx_1$. De $R^{*=}xx_1$ e Rx_1x_2, obtemos $R^{*=}xx_2$ (por *Anc*11); e de $R^{*=}xx_1$, Rx_1x_2 e Fx_1, obtemos Fx_2 (pela hipótese inicial). Logo, $R^{*=}xx_2 \wedge Fx_2 \Rightarrow [u : R^{*=}xu \wedge Fu]x_2$.

De (i) e (ii), inferimos $\forall y(R^{*=}xy \to [u : R^{*=}xu \wedge Fu]y) \;\Rightarrow\; \forall y(R^{*=}xy \to Fy)$[10].

■

2.3.2 Demonstração de ADP1, ADP2, ADP7 e ADP5

Agora que definimos *relação ancestral* e *relação ancestral fraca à la* Frege e estabelecemos os fatos sobre * e $^{*=}$ necessários na prova do Teorema de Frege, daremos início à demonstração dos axiomas de **ADP**, começando por aqueles que não oferem dificuldade alguma: **ADP**1, **ADP**2, **ADP**7 e **ADP**5. Antes, porém, replicamos em **AF** as definições fregeanas das noções aritméticas primitivas: 0, *predecessor* e *número natural*, ou, no nosso caso, **0**, *Pred* e *Nat*, os símbolos não lógicos de $\Sigma_{\mathbf{DP}}$.

Em *Gl*, §74, Frege define o número 0 como o número cardinal do conceito 'diferente de si próprio'. Em **AF**, temos

($D_{\mathbf{0}}$)

$$\mathbf{0} = \#[u : u \neq u]$$[11]

A relação *predecessor* é definida em *Gl*, §76; a definição correspondente em **AF** é

(D_{Pred})

$$Pred(x,y) \leftrightarrow \exists F \exists z(Fz \wedge \#[u : Fu \wedge u \neq z] = x \wedge \#F = y)$$[12]

[10]Nessa demonstração, utiliza-se um expediente algumas vezes denominado 'fortalecer a hipótese indutiva': instancia-se G para $[u : R^{*=}xu \wedge Fu]$ em *Anc*9, pois a presença da fórmula $R^{*=}xx_1$ na hipótese indutiva $R^{*=}xx_1 \wedge Fx_1$ em (ii) é necessária para se obter Fx_2 a partir da hipótese inicial. Com isso, obtêm-se um fato mais forte do que aquele almejado, $\forall y(R^{*=}xy \to [u : R^{*=}xu \wedge Fu]y)$, mas do qual o consequente do teorema pode ser obtido de maneira imediata. Se tentássemos recorrer à instância $Fx \wedge Her(F,R) \to \forall y(R^{*=}xy \to Fy)$ de *Anc*9 para obter $\forall y(R^{*=}xy \to Fy)$ diretamente, não conseguiríamos provar $Her(F,R)$, uma vez que, sem $R^{*=}xx_1$, não seria possível inferir Fx_2 de Rx_1x_2, Fx_1 e da hipótese inicial.

[11]Explicitamente, $\exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge \mathbf{0} = \#G)$. $D_{\mathbf{0}}$ poderia ser formulada em forma normalizada como $\mathbf{0} = x \leftrightarrow \exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge \#G = x)$. Nesse caso, prova-se que $\vdash_{\mathbf{AF}} \exists! x \exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge \#G = x)$ a partir dos axiomas *L*9 e *L*10, o que garante a legitimidade da definição.

[12]Explicitamente, $Pred(x,y) \leftrightarrow \exists F \exists z(Fz \wedge \exists G(\forall u(Gu \leftrightarrow (Fu \wedge u \neq z)) \wedge \#G = x) \wedge \#F = y)$.

i.é., x precede y (na terminologia fregeana, *y segue imediatamente após x na série natural dos números*) se e somente se existem F e z tais z é F, o número do conceito 'F e diferente de z' é x, e o número do conceito F é y – para abreviar as demonstrações em que D_{Pred} seja aplicada, convencionaremos escrever F^{-z} no lugar de $[u : Fu \wedge u \neq z]$, de agora em diante.

Em §77, Frege define o número 1 como o número do conceito 'igual a 0' e prova que 0 precede 1:

$(D_{\mathbf{1}})$

$$\mathbf{1} = \#[u : u = \mathbf{0}]^{13}$$

Pelas definições de $\mathbf{0}$, $\mathbf{1}$ e *Pred* é imediato provar $Pred(\mathbf{0},\mathbf{1})$ em **AF**, bastando, para isso, provar $[u : u = \mathbf{0}]\mathbf{0} \wedge \#[u : u = \mathbf{0}]^{-\mathbf{0}} = \mathbf{0} \wedge \#[u : u = \mathbf{0}] = \mathbf{1}$. Embora Frege não defina os números individuais 2, 3, ... em *Gl*, as definições $D_{\mathbf{0}}$, $D_{\mathbf{1}}$ e D_{Pred} sugerem o método: $\mathbf{2} = \#[u : u = \mathbf{0} \vee u = \mathbf{1}]$, $\mathbf{3} = \#[u : u = \mathbf{0} \vee u = \mathbf{1} \vee u = \mathbf{2}]$, ... $Pred(\mathbf{1},\mathbf{2})$, $Pred(\mathbf{2},\mathbf{3})$, ... são então facilmente provados em **AF**[14].

Frege define *número natural*, em sua terminologia, *número finito*, apenas no final do esboço da prova do Teorema de Frege, em *Gl*, §83: x é um número finito se e somente se ele pertence à série natural dos números que começa com 0, a formulação fregeana de 'x é um *Pred*-ancestral fraco de $\mathbf{0}$':

(D_{Nat})

$$Nat(x) \leftrightarrow Pred^{*=}(\mathbf{0},x)$$

Além das definições fregeanas acima, propomos ainda duas definições adicionais que facilitarão a exposição da prova de **ADP**6:

[13] Explicitamente, $\exists G(\forall u(Gu \leftrightarrow u = \mathbf{0}) \wedge \mathbf{1} = \#G)$. Assim como $D_{\mathbf{0}}$, $D_{\mathbf{1}}$, também poderia ser formulada em forma normalizada como $\mathbf{1} = x \leftrightarrow \exists H(\forall u(Hu \leftrightarrow u = \mathbf{0}) \wedge \#H = x)$ ou, eliminando-se a ocorrência do numeral $\mathbf{0}$, como $\mathbf{1} = x \leftrightarrow \exists H(\forall u_1(Hu_1 \leftrightarrow \exists G(\forall u(Gu \leftrightarrow u \neq u) \wedge \#G = u_1)) \wedge \#H = x)$.

[14] Note-se a similaridade das definições fregeanas dos números individuais com as definições correspondentes na teoria de conjuntos de Zermelo-Fraenkel (**ZF**): $\mathbf{0} = \varnothing, \mathbf{1} = \{\mathbf{0}\}, \mathbf{2} = \{\mathbf{0},\mathbf{1}\}, \mathbf{3} = \{\mathbf{0},\mathbf{1},\mathbf{2}\}$; ou seja, $\mathbf{0} = \{x : x \neq x\}, \mathbf{1} = \{x : x = \mathbf{0}\}, \mathbf{2} = \{x : x = \mathbf{0} \vee x = \mathbf{1}\}, \mathbf{3} = \{x : x = \mathbf{0} \vee x = \mathbf{1} \vee x = \mathbf{2}\}$.

($D_<$)

$$x < y \leftrightarrow Pred^*(x,y)$$

($D_\leq$)

$$x \leq y \leftrightarrow Pred^{*=}(x,y)$$

Ou seja, x é *menor* que y se e somente se x é um *Pred*-ancestral de y, e x é *menor ou igual a* y se e somente se x é um *Pred*-ancestral fraco de y. Nos referiremos a < por 'série dos números naturais' (denominada por Frege 'série natural dos números' no *Grundlagen*).

É fácil notar que **ADP**1 é uma consequência trivial da definição de *Nat* e de *Anc*7; **ADP**2, por sua vez, é obtido dessa mesma definição e de *Anc*11:

(**ADP**2) $Nat(x) \wedge Pred(x,y) \rightarrow Nat(y)$

Demonstração: Instanciando R para *Pred*, x_1 para $\mathbf{0}$, x_2 para x e x_3 para y na variação alfabética $R^{*=}x_1x_2 \wedge Rx_2x_3 \rightarrow R^{*=}x_1x_2$ de *Anc*11, obtemos

$$Pred^{*=}(\mathbf{0},x) \wedge Pred(x,y) \rightarrow Pred^{*=}(\mathbf{0},y)$$

i.é.,

$$Nat(x) \wedge Pred(x,y) \rightarrow Nat(y)$$

■

Como anunciado anteriormente, **ADP**7, o princípio de indução matemática, é um caso particular do princípio de indução generalizada 3 (*Anc*13):

(**ADP**7) $F\mathbf{0} \wedge \forall x_1 \forall x_2 (Nat(x_1) \wedge Pred(x_1,x_2) \rightarrow (Fx_1 \rightarrow Fx_2)) \rightarrow \forall x(Nat(x) \rightarrow Fx)$

Demonstração: Instanciando R para *Pred* e y para $\mathbf{0}$ na seguinte variação alfabética de *Anc*13

$$Fy \wedge \forall x_1 \forall x_2 (R^{*=}yx_1 \wedge Rx_1x_2 \rightarrow (Fx_1 \rightarrow Fx_2)) \rightarrow \forall x(R^{*=}yx \rightarrow Fx)$$

obtemos

$$F\mathbf{0} \wedge \forall x_1 \forall x_2 (Pred^{*=}(\mathbf{0},x_1) \wedge Pred(x_1,x_2) \rightarrow (Fx_1 \rightarrow Fx_2)) \rightarrow \forall x(Pred^{*=}(\mathbf{0},x) \rightarrow Fx)$$

i.é.,

$$F\mathbf{0} \wedge \forall x_1 \forall x_2 (Nat(x_1) \wedge Pred(x_1, x_2) \rightarrow (Fx_1 \rightarrow Fx_2)) \rightarrow \forall x (Nat(x) \rightarrow Fx)$$

■

Em **AF**, é possível provar também uma versão mais fraca de indução matemática a partir do princípio de indução generalizada 2 (*Anc*9):

(**ADP**7′) $F\mathbf{0} \wedge \forall x_1 \forall x_1 (Pred(x_1, x_2) \rightarrow (Fx_1 \rightarrow Fx_2)) \rightarrow \forall x (Nat(x) \rightarrow Fx)$

ADP7′ é a fórmula elegida por Heck em [25] como o axioma de indução de **ADP** e representa a única diferença entre a sua formulação da Aritmética de Dedekind-Peano e aquela apresentada acima. No entanto, não é difícil verificar que ambas as formulações da teoria são equivalentes: **ADP**7′ é facilmente obtido de **ADP**7 e **ADP**7 é provado a partir de **ADP**7′ da mesma maneira que o princípio de indução generalizada 3 foi provado a partir do princípio de indução generalizada 2: instanciando F para $[u : Nat(u) \wedge Fu]$ e utilizando **ADP**1 no lugar de *Anc*7 e **ADP**2 no lugar de *Anc*11.

Há que se notar que os axiomas **ADP**1, **ADP**2 e **ADP**7 (ou **ADP**7′) foram provados unicamente a partir de D_{Nat} e dos fatos sobre * e $^{*=}$ acima, cujas demonstrações não envolvem qualquer recurso ao Princípio de Hume. Isso signifca que a definição fregeana de *Nat* como $Pred^{*=}$ é uma forma de, por assim dizer, incorporar as propriedades dos números naturais expressas por aqueles axiomas, incluindo-se o princípio de indução matemática, na própria definição do conceito de número natural. Por outro lado, como veremos, as demonstrações de **ADP**3-6 em **AF** dependem de *PH*.

No *Grundlagen*, Frege não prova explicitamente a proposição de que nenhum número natural precede 0 (**ADP**5), mas, em *Gl*, §75, logo após definir o número 0, ele enuncia um lema com o qual demonstra-se facilmente aquela proposição:

> Deve ser possível provar agora, por meio das suposições anteriores, que todo conceito sob o qual nada cai é equinumérico a todo conceito sob o qual nada cai, e somente a um conceito dessa espécie, de onde se segue que 0 é o número que convém a um tal conceito, e que nenhum objeto cai sob um conceito se o número que lhe convém é 0.

O conteúdo dessa passagem pode ser expresso em **AF** pelas fórmulas (i) $\neg\exists xFx \rightarrow (\neg\exists xGx \leftrightarrow F \approx G)$, (ii) $\neg\exists xFx \rightarrow \#F = \mathbf{0}$ e (iii) $\#F = \mathbf{0} \rightarrow \neg\exists xFx$, de modo que (ii) e (iii) são, segundo Frege, obtidos de (i). No nosso caso, provaremos apenas (iii), sem recorrer a (i), na medida em (ii) não é necessária nos desenvolvimentos subsequêntes.

(**ADP**5.1.1) $\#F = \mathbf{0} \rightarrow \neg\exists xFx$

Demonstração: Suponha $\#F = \mathbf{0}$ e $\exists xFx \Rightarrow Fx$. Logo, $\#F = \#[u : u \neq u]$. Por *PH*, obtemos $F \approx [u : u \neq u]$ e, portanto, $\exists R\, F \approx_R [u : u \neq u] \Rightarrow F \approx_R [u : u \neq u]$, de onde inferimos

$$\forall x(Fx \rightarrow \exists y(Rxy \wedge [u : u \neq u]y))$$

De Fx, obtemos então $\exists y(Rxy \wedge [u : u \neq u]y) \Rightarrow Rxy \wedge [u : u \neq u]y \Rightarrow y \neq y$. Absurdo.

■

Na presença de **ADP**5.1.1, provamos o seguinte lema, do qual o axioma **ADP**5 é uma consequência imediata:

(**ADP**5.1) $\neg\exists xPred(x, \mathbf{0})$

Demonstração: Suponha $\exists xPred(x, \mathbf{0}) \Rightarrow Pred(x, \mathbf{0})$. Logo,

$$\exists F\exists z(Fz \wedge \#F^{-z} = x \wedge \#F = \mathbf{0}) \Rightarrow Fz \wedge \#F^{-z} = x \wedge \#F = \mathbf{0}$$

De $\#F = \mathbf{0}$, inferimos $\neg\exists xFx$ (por **ADP**5.1.1); e de Fz, inferimos $\exists xFx$. Absurdo.

■

2.3.3 Demonstração de ADP3 e ADP4

Em *Gl*, §78, Frege enumera seis proposições sobre as definições de 0, 1 e *predecessor*, afirmando, no ínicio da seção, que o leitor perceberá facilmente como prová-las. Dentre elas, encontra-se a proposição que expressa a bijetividade da relação *predecessor* (Proposição 5), enunciada por Frege como "A relação de *m* e *n* estabelecida pela proposição: '*n* segue imediatamente após *m* na série natural dos números' é uma relação biunívoca". Em **AF**,

podemos expressá-la através da fórmula $Pred(x,y) \wedge Pred(z,w) \rightarrow (x = z \leftrightarrow y = w)$ ou ainda subdividí-la em

(**ADP**3.1) $Pred(x,y) \wedge Pred(x,z) \rightarrow y = z$

(**ADP**4.1) $Pred(y,x) \wedge Pred(z,x) \rightarrow y = z$

ADP3 e **ADP**4 são consequências imediatas de **ADP**3.1 e **ADP**4.1, cujas demonstrações pressupõem dois lemas sobre $\approx$: **ADP**3.1.1 e **ADP**4.1.1 abaixo. Ainda que não ofereçam nenhuma dificuldade substancial, as provas desses lemas são demasiadamente longas para serem apresentadas em detalhe aqui e, por essa razão, forneceremos uma demonstração abreviada de **ADP**3.1.1 e apenas enunciaremos **ADP**4.1.1, inserindo sua demonstração no Apêndice B2.

ADP3.1.1 expressa que se o conceito 'F e diferente de x' é equinumérico ao conceito 'G e diferente de y', x é F e y é G, então F é equinumérico a G, e, intuitivamente, estabelece que o processo de "acrescentar" exatamente um objeto a cada conceito de um par de conceitos equinuméricos preserva a relação de equinumericidade:

(**ADP**3.1.1) $F^{-x} \approx G^{-y} \wedge Fx \wedge Gy \rightarrow F \approx G$

Demonstração: Suponha $F^{-x} \approx G^{-y} \wedge Fx \wedge Gy$. Logo, $\exists R\, F^{-x} \approx_R G^{-y} \;\Rightarrow\; F^{-x} \approx_R G^{-y}$. Considere a relação $[uu_1 : Ruu_1 \vee (u = x \wedge u_1 = y)]$. É imediato mostrar $F \approx_{[uu_1 : Ruu_1 \vee (u = x \wedge u_1 = y)]} G$. Portanto, $F \approx G$.

■

(**ADP**3.1) $Pred(x,y) \wedge Pred(x,z) \rightarrow y = z$

Demonstração: Suponha $Pred(x,y) \wedge Pred(x,z)$. Logo,

$$\exists F \exists x_1 (Fx_1 \wedge \#F^{-x_1} = x \wedge \#F = y) \;\Rightarrow\; Fx_1 \wedge \#F^{-x_1} = x \wedge \#F = y$$

e

$$\exists G \exists y_1 (Gy_1 \wedge \#G^{-y_1} = x \wedge \#G = z) \;\Rightarrow\; Gy_1 \wedge \#G^{-y_1} = x \wedge \#G = z$$

De $\#F^{-x_1} = \#G^{-y_1}$, obtemos $F^{-x_1} \approx G^{-y_1}$ (por PH). De **ADP**3.1.1, Fx_1 e Gy_1, inferimos então $F \approx G$. Por PH novamente, $\#F = \#G$ e, portanto, $y = z$.

■

ADP4.1.1 é a inversa de **ADP**3.1.1 e expressa que se F é equinumérico a G, x é F e y é G, então o conceito 'F mas é diferente de x' é equinumérico a 'G mas é diferente y'. Intuitivamente, **ADP**4.1.1 estabelece que o processo de "retirar" exatamente um objeto de cada conceito de um par de conceitos equinuméricos também preserva a relação de equinumericidade:

(**ADP**4.1.1) $F \approx G \wedge Fx \wedge Gy \rightarrow F^{-x} \approx G^{-y}$

Demonstração: Apêndice B2.

■

(**ADP**4.1) $Pred(y,x) \wedge Pred(y,z) \rightarrow y = z$

Demonstração: Suponha $Pred(y,z) \wedge Pred(z,x)$. Logo,

$$\exists F \exists x_1 (Fx_1 \wedge \#F^{-x_1} = y \wedge \#F = x) \Rightarrow Fx_1 \wedge \#F^{-x_1} = y \wedge \#F = x$$

e

$$\exists G \exists y_1 (Gy_1 \wedge \#G^{-y_1} = z \wedge \#G = x) \Rightarrow Gy_1 \wedge \#G^{-y_1} = z \wedge \#G = x$$

De $\#F = \#G$, obtemos $F \approx G$ (por PH). De **ADP**4.1.1, Fx_1 e Gy_1, inferimos então $F^{-x_1} \approx G^{-y_1}$. Por PH novamente, $\#F^{-x_1} = \#G^{-y_1}$ e, portanto, $y = z$.

■

2.3.4 Demonstração de ADP6

Tendo provado **ADP**1-5 e **ADP**7 em **AF**, chegamos agora à etapa mais desafiadora da prova do Teorema de Frege: a demonstração de **ADP**6. Frege esboça a prova da proposição correspondente de que todo número natural possui um sucessor em *Gl*, §82-83 e comenta, no início da seção 82, que ao provar-se esse fato, ter-se-á estabelecido que a série dos números naturais não possui um último membro. Se combinado com a proposição de que nenhum número natural segue a si próprio nessa série, enunciada por Frege na seção 83 e replicada

em **AF** como **ADP**6.1.2.2 abaixo, ela estabelece ainda que cada número natural possui um sucessor distinto de todos os números que o precedem, o que equivale a dizer que existem infinitos números naturais. Em vista dessa consequência, a relevância filosófica da prova de **ADP** em **AF** é significativamente maior do que a dos demais axiomas de **ADP**, pois ela representa o fato notável de que a existência de uma infinidade de objetos distintos pode ser estabelecida unicamente a partir do Princípio de Hume.

A estratégia fregeana em *Gl*, §82-83 para provar que todo número natural possui um sucessor consiste em mostrar que se n é um número natural, então ele precede o número (cardinal) de números menores ou iguais a n ou, alternativamente, que n precede o número do conceito 'menor ou igual a n'. Em **AF**, essa proposição é expressa por

(**ADP**6.1)

$$Nat(x) \rightarrow Pred(x, \#[u : u \leq x])$$

De acordo com as indicações gerais de Frege em §82, a prova de **ADP**6.1 requer aplicar alguma versão de indução sobre o conceito $\left[u_1 : Pred(u_1, \#[u : u_1 \leq u])\right]$, uma vez que ele se refere ao teorema como resultando da aplicação de um método de inferência obtido da definição de *ancestral* e identifica claramente, entre as etapas principais da prova, um passo base e um passo indutivo. No entanto, nenhuma menção é feita à forma específica do princípio de indução a ser utilizado e, como não nos propomos formular uma prova de **ADP**6 inteiramente fidedigna àquela esboçada no *Grundlagen*, podemos supor que o princípio de indução relevante é expresso em **AF** por **ADP**7.

ADP6.1 é uma variante do consequente da seguinte instância de **ADP**7:

$$\left[u_1 : Pred(u, \#[u : u_1 \leq u])\right]\mathbf{0} \wedge$$

$$\forall x \forall y(Nat(x) \wedge Pred(x,y) \rightarrow (\left[u_1 : Pred(u, \#[u : u_1 \leq u])\right]x \rightarrow \left[u_1 : Pred(u, \#[u : u_1 \leq u])\right]y)) \rightarrow$$

$$\forall x(Nat(x) \rightarrow \left[u_1 : Pred(u, \#[u : u_1 \leq u])\right]x)$$

Para prová-lo, portanto, basta provar

(**ADP**6.1.1) PASSO BASE

$$\big[u_1 : Pred(u_1, \#[u : u \leq u_1])\big]\mathbf{0}$$

(**ADP**6.1.2) PASSO INDUTIVO

$$\forall x \forall y(Nat(x) \wedge Pred(x,y) \rightarrow (\big[u_1 : Pred(u_1, \#[u : u \leq u_1])\big]x \rightarrow \big[u_1 : Pred(u_1, \#[u : u \leq u_1)])\big]y))$$

Para provar **ADP**6.1.1, que afirma que 0 precede o número de números menores ou iguais a 0, é preciso mostrar $Pred(\mathbf{0}, \#[u : u \leq \mathbf{0}])$ e como $\vdash_{\mathbf{AF}} Pred(\mathbf{0}, \mathbf{1})$, é, para isso, suficiente mostrar $\#[u : u \leq \mathbf{0}] = \mathbf{1}$, o que requer o seguinte lema:

(**ADP**6.1.1.1) $\neg\, x < \mathbf{0}$

Demonstração: Suponha $x < \mathbf{0}$. Instanciando R para $Pred$ e y para $\mathbf{0}$ em $Anc4$, obtemos

$$Pred^*(x, \mathbf{0}) \rightarrow \exists z Pred(z, \mathbf{0})$$

i.é.,

$$x < \mathbf{0} \rightarrow \exists z Pred(z, \mathbf{0})$$

Logo, $\exists z Pred(z, \mathbf{0}) \;\Rightarrow\; \exists x Pred(x, \mathbf{0})$, o que contradiz **ADP**5.1. Absurdo.

■

A partir de **ADP**6.1.1.1, demonstra-se **ADP**6.1.1 facilmente:

Demonstração: Suponha $[u : u \leq \mathbf{0}]x \;\Rightarrow\; x \leq \mathbf{0}$. Logo, $x < \mathbf{0} \vee x = \mathbf{0}$. Por **ADP**6.1.1.1, obtemos $x = \mathbf{0} \;\Rightarrow\; [u : u = \mathbf{0}]x$.

Suponha, por outro lado, $[u : u = \mathbf{0}]x \;\Rightarrow\; x = \mathbf{0}$. Logo, $x \leq \mathbf{0} \;\Rightarrow [u : u \leq \mathbf{0}]x$.

Portanto, $[u : u \leq \mathbf{0}] \equiv [u : u = \mathbf{0}]$. Por $\approx_4$ e PH, inferimos $\#[u : u \leq \mathbf{0}] = \#[u : u = \mathbf{0}]$, i.é., $\#[u : u \leq \mathbf{0}] = \mathbf{1}$. De $Pred(\mathbf{0}, \mathbf{1})$, obtemos então $Pred(\mathbf{0}, \#[u : u \leq \mathbf{0}])$.

■

A demonstração de **ADP**6.1.2, o passo indutivo da prova de **ADP**6.1, é consideralmente mais complexa do que a demonstração de **ADP**6.1.1. Para prová-lo, será preciso mostrar $Pred(y, \#[u : u \leq y])$ a partir das hipóteses $Nat(x) \wedge Pred(x,y)$ e $Pred(x, \#[u : u \leq x])$. Pela

definição de *Pred*, isso requer mostrar que existem F e z tais que z é F, que o número do conceito 'F e diferente de z' é y e que o número do conceito F é $\#[u : u \leq y]$:

$$\exists F \exists z (Fz \wedge \#F^{-z} = y \wedge \#F = \#[u : u \leq y])$$

Mas quais seriam as testemunhas para F e z a serem utilizadas na demonstração? Os candidatos mais naturais são $[u : u \leq y]$ e o próprio y, respectivamente: como $y \leq y$ e $\#[u : u \leq y] = \#[u : u \leq y]$, o único desafio da prova seria então mostrar $\#[u : u \leq y]^{-y} = y$ a partir das hipóteses do teorema. Frege aponta a necessidade de se provar esse fato no início de §83, ao afirmar "devemos mostrar que [y] é o número que convém ao conceito 'pertence à série natural dos números que termina em [y], mas não é idêntico a [y]' ". Em seguida, ele acrescenta: "E, para isso, deve-se provar que este conceito possui a mesma extensão que o conceito 'pertence à série natural dos números que termina em [x]' " (em **AF**, **ADP**1.2.3 e **ADP**1.2.4 abaixo)[15].

Ainda em *Gl*, §83, Frege afirma ser necessário, para se obter a proposição acima, mostrar que nenhum número natural segue após si próprio na série $<$, o que significa que $<$ não possui *loops*, ou seja, que situações como $\langle 0,1,2,\ldots,8,9,10,8,9,10,8,9,10,\ldots\rangle$ ou como $\langle 0,1,\ldots,1848,100,101,\ldots,1848,100,101,\ldots\rangle$ não ocorrem. Em **AF**, a fórmula correspondente a esse fato é **ADP**6.1.2.2 abaixo, cuja prova pressupõe

(**ADP**6.1.2.1) $x < y \wedge Pred(z,y) \rightarrow x \leq z$

Demonstração: Suponha $x < y \wedge Pred(z,y)$. Instanciando R para *Pred* na variação alfabética $R^*xy \rightarrow \exists w(R^{*=}xw \wedge Rwy)$ de *Anc*12, obtemos

$$Pred^*(x,y) \rightarrow \exists w(Pred^{*=}(x,w) \wedge Pred(w,y))$$

i.é.,

$$x < y \rightarrow \exists w(x \leq w \wedge Pred(w,y))$$

De $x < y$, inferimos $\exists w(x \leq w \wedge Pred(w,y)) \Rightarrow x \leq w \wedge Pred(w,y)$; e de **ADP**4.1, $Pred(z,y)$ e $Pred(w,y)$, inferimos $z = w$. Logo, $x \leq z$.

■

[15] Em *Gl*, §§82-83, Frege utiliza as variáveis d e a no lugar de x e y, respectivamente.

(**ADP**6.1.2.2) $Nat(x) \rightarrow \neg\, x < x$

Demonstração: Instanciando F para $[u : \neg\, u < u]$ em **ADP**7′, obtemos

$$[u : \neg\, u < u]\mathbf{0} \wedge Her([u : \neg\, u < u], Pred) \rightarrow \forall x(Nat(x) \rightarrow [u : \neg\, u < u]x)$$

(i) De **ADP**6.1.1.1, obtemos $[u : \neg\, u < u]\mathbf{0}$.

(ii) Para provar $Her([u : \neg\, u < u], Pred)$, suponha $Pred(x_1, x_2)$ e $[u : \neg\, u < u]x_1 \;\Rightarrow\; \neg\, x_1 < x_1$. Suponha também $x_2 < x_2$ (para proceder por absurdo). De $x_2 < x_2$ e $Pred(x_1, x_2)$, inferimos $x_2 \leq x_1$ (por **ADP**6.1.2.1). Por $Anc11$ e $Pred(x_1, x_2)$, obtemos então $x_1 < x_1$. Absurdo.

De (i) e (ii), inferimos $\forall x(Nat(x) \rightarrow [u : \neg\, u < u]x) \;\Rightarrow\; \forall x(Nat(x) \rightarrow \neg\, x < x)$.

■

Como indica Frege, com **ADP**6.1.2.2, demonstra-se que exatamente os mesmos objetos caem sob $[u : u \leq y]^{-y}$ e $[u : u \leq x]$:

(**ADP**6.1.2.3) $Nat(x) \wedge Pred(x, y) \rightarrow \forall z([u : u \leq x]z \rightarrow [u : u \leq y]^{-y}z)$

Demonstração: Suponha $Nat(x) \wedge Pred(x, y)$ e $[u : u \leq x]z \;\Rightarrow\; z \leq x$. Precisamos mostrar $z \leq y \wedge z \neq y$.

(i) Para provar $z \leq y$, considere a variação alfabética $R^{*=}x_1x_2 \wedge Rx_2x_3 \rightarrow R^*x_1x_3$ de $Anc10$. Instanciando R para $Pred$, x_1 para z, x_2 para x e x_3 para y, obtemos

$$Pred^{*=}(z, x) \wedge Pred(x, y) \rightarrow Pred^*(z, y)$$

i.é.,

$$z \leq x \wedge Pred(x, y) \rightarrow z < y$$

Logo, $z < y$ e, portanto, $z \leq y$ (por $Anc7$).

(ii) Para provar $z \neq y$, suponha $z = y$. Por (i), temos $z < y \;\Rightarrow\; y < y$. Mas, $Nat(y)$ (por **ADP**2 e $Nat(x) \wedge Pred(x, y)$) e, portanto, $\neg\, y < y$ (por **ADP**6.1.2.2). Absurdo.

De (i) e (ii), inferimos $z \leq y \wedge z \neq y \;\Rightarrow\; [u : u \leq y]^{-y} z$.

■

(**ADP**6.1.2.4) $Pred(x,y) \rightarrow \forall z([u : u \leq y]^{-y} z \rightarrow [u : u \leq x]z)$

Demonstração: Suponha $Pred(x,y)$ e $[u : u \leq y]^{-y} z \;\Rightarrow\; z \leq y \wedge z \neq y$. Precisamos mostrar $z \leq x$. Considere a variação alfabética $x_1 < x_3 \wedge Pred(x_2, x_3) \rightarrow x_1 \leq x_2$ de **ADP**6.1.2.1. Instanciando x_1 para z, x_2 para x e x_3 para y, obtemos

$$z < y \wedge Pred(x,y) \rightarrow z \leq x$$

De $z \leq y$ e $z \neq y$, inferimos $z < y$; e $Pred(x,y)$ está entre as hipóteses iniciais. Portanto, $z \leq x \;\Rightarrow\; [u : u \leq x]z$.

■

ADP6.1.2.3 e **ADP**6.1.2.4 poderiam ser condensados em

$$Nat(x) \wedge Pred(x,y) \rightarrow [u : u \leq x] \equiv [u : u \leq y]^{-y}$$

mas optamos por prová-los separadademente para explicitar que, enquanto **ADP**6.1.2.2 é necessário para inferir $[u : u \leq y]^{-y} z$ de $[u : u \leq x]z$ (na demonstração de **ADP**6.1.2.3), ele não é utilizado para se obter a direção inversa (na demonstração de **ADP**6.1.2.4). É justamente para possibilitar a aplicação de **ADP**6.1.2.2 que a fórmula $Nat(x)$ ocorre no antecedente de **ADP**6.1.2.3, embora esteja ausente em **ADP**6.1.2.4.

Em *Gl*, §82, Frege enuncia o passo indutivo da prova de que todo número natural *n* precede o número de números menores ou iguais a *n* (**ADP**6.1) como

> Se [x] segue imediatamente após [y] na série natural dos números [$Pred(x,y)$], e se vale para [x] que: o número que convém ao conceito 'pertence à série natural dos números que termina em [x]' segue imediatamente após [x] na série natural dos números [$Pred(x, \#[u : u \leq x])$], então também vale para [y] que: o número que convém ao onceito 'pertence à série natural dos números que termina em [y]' segue imediatamente após [y] na série natural dos números [$Pred(y, \#[u : u \leq y])$].

Quando comparada a **ADP**6.1.2, notamos nessa passagem a ausência da hipótese de que x é um número natural. Ao indicar as linhas gerais da prova dessa proposição em *Gl*, §83, Frege também omite essa hipótese quando enuncia o lema correspondente à fórmula $Nat(x) \wedge Pred(x,y) \rightarrow [u : u \leq x] \equiv [u : u \leq y]^{-y}$ acima, embora ele seja enfático ao afirmar que ela é necessária na proposição de que nenhum número segue a si próprio na série dos números naturais (**ADP**6.1.2.2). Como vimos, no entanto, a presença de $Nat(x)$ no antecedente de **ADP**6.1.2.3 é fundamental para que seja possível utilizar **ADP**6.1.2.2 em sua demonstração, o que significa que não poderíamos simplesmente provar $Pred(x,y) \rightarrow [u : u \leq x] \equiv [u : u \leq y]^{-y}$ a partir de **ADP**6.1.2.2, como Frege parece indicar[16]. Caso isso fosse possível, a presença de $Nat(x)$ no antecedente de **ADP**6.1.2 também seria desnecessária (cf. a demonstração de **ADP**6.1.2 abaixo), e não precisaríamos utilizar **ADP**7 para provar **ADP**6.1, bastando, para isso, o princípio de indução **ADP**7′. É devido ao fato de ser necessário utilizar **ADP**7 na prova de **ADP**6.1 que optamos por incluí-lo no lugar de **ADP**7′ na axiomatização de Heck de **ADP** em [25].

Esses indícios compõem a tese interpretativa de Boolos e Heck de que Frege não possuía uma prova completa do teorema homônimo em 1884. De acordo com eles, Frege realmente pensava ser possível provar $Pred(x,y) \rightarrow [u : u \leq x] \equiv [u : u \leq y]^{-y}$ a partir de **ADP**6.1.2.2 e, assim, provar $Pred(x,y) \rightarrow (\left[u_1 : Pred(u_1, \#[u :\leq u_1])\right]x \rightarrow \left[u_1 : Pred(u_1, \#[u :\leq u_1])\right]y)$ no lugar de **ADP**6.1.2. Esse fato explicaria então porque ele afirma em *Gl*, §82 que a proposição correspondente a **AD**6.1 é uma consequência de um método de inferência obtido da definição de *ancestral*, já que, diferentemente de **ADP**7, **ADP**7′ é uma consequência imediata dessa definição. Em suma, Boolos e Heck concluem que Frege cometeu um erro em *Gl*, §§82-83.

Correta ou não, a conclusão de Boolos e Heck não invalida o fato de que Frege apresentou uma prova correta de que todo número natural possui um sucessor no *Grungesetze*. Além disso, não se pode negar que, mesmo que correta, a tese de que Frege não tenha efetivamente provado aquela proposição no *Grundlagen* e que não sabia, à época, como fazê-lo, ele estava bem próximo de alcançar o seu objetivo e, talvez, as imprecisões presentes em *Gl*, §§82-83 fossem devidas ao fato de os resultados *esboçados* no livro não terem sido ainda

[16]Esse ponto foi apontado pela primeira vez por Boolos em [4], p. 194.

plenamente formalizados em um sistema dedutivo, de modo a explicitar todos os detalhes a serem preenchidos. Indubitavelmente, entretanto, a apresentação da prova conjectural do Teorema de Frege deste capítulo evidencia que, em 1884, Frege já dispunha de todos os recursos necessários para formular uma prova sem lacunas ou imprecisões.

Para retomar o fio da exposição da prova do Teorema de Frege e finalizá-la, notamos que o seguinte teorema é uma consequência de **ADP**6.1.2.3, **ADP**6.1.2.4, $\approx_4$ e *PH*:

(**ADP**6.1.2.5) $Nat(x) \wedge Pred(x,y) \rightarrow \#[u : u \leq y]^{-y} = \#[u : u \leq x]$

e que partir de **ADP**6.1.2.5, demonstra-se finalmente **ADP**6.1.2, do qual **ADP**6.1 e, portanto, **ADP**6 são consequências imediatas:

Demonstração: Suponha $Nat(x) \wedge Pred(x,y)$ e

$$\big[u_1 : Pred(u_1, \#[u : u \leq u_1])\big]x \;\Rightarrow\; Pred(x, \#[u : u \leq x])$$

De **ADP**6.1.2.5 e $Nat(x) \wedge Pred(x,y)$, obtemos

$$\#[u : u \leq y]^{-y} = \#[u : u \leq x]$$

De **ADP**3, $Pred(x, \#[u : u \leq x])$ e $Pred(x,y)$, inferimos $\#[u : u \leq x] = y$. Portanto,

$$\#[u : u \leq y]^{-y} = y$$

Logo,

$$[u : u \leq y]y \wedge \#[u : u \leq y]^{-y} = y \wedge \#[u : u \leq y] = \#[u : u \leq y]$$

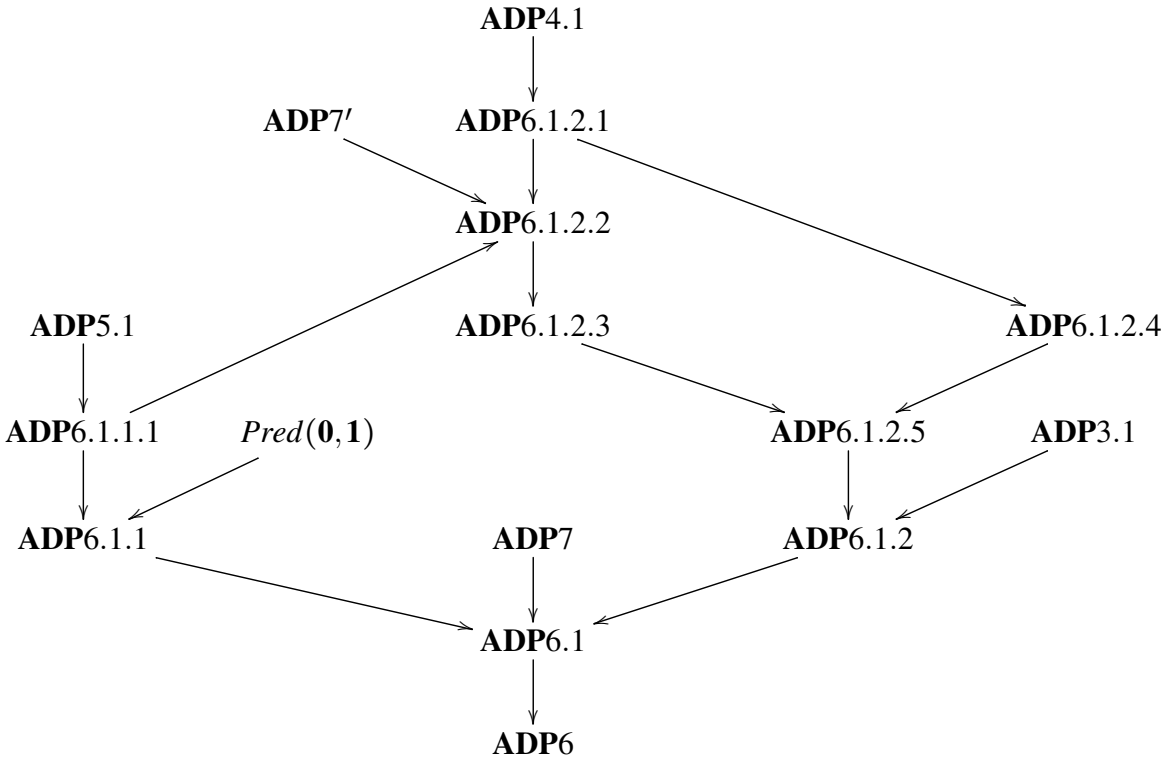
Quantificando sobre y e $[u : u \leq y]$, obtemos

$$\exists F \exists z (Fz \wedge \#F^{-z} = y \wedge \#F = [u : u \leq y])$$

i.é.,

$$Pred(y, \#[u : u \leq y]) \;\Rightarrow\; \big[u_1 : Pred(u_1, \#[u : u \leq u_1])\big]y$$

■



Estrutura da Demonstração de **ADP**6

Como consequência do Teorema de Frege e do Corolário 1.39, obtemos o seguinte resultado de consistência relativa:

Corolário 2.6. Se **AF** é consistente, então **ADP** é consistente.

Capítulo 3

Consistência

> It's a pity that Russell's paradox has obscured Frege's accomplishment in the *Grundlagen* (...). And as Frege's work shows, Hume's principle is much more powerful than we might have supposed it to be, implying, with the aid of second-order logic, the whole of second order arithmetic (while failing to imply $\perp$).
>
> George Boolos

Tendo provado o Teorema de Frege, abordaremos, neste capítulo final, o tema da consistência de **AF**.

Como notamos, a descoberta do paradoxo de Russell foi decisiva para que Frege abandonasse, de uma vez por todas, a tese logicista, por volta de 1906; e, ademais, consistiu na principal razão pela qual o Teorema de Frege permaneceu desconhecido pela tradição filosófica durante grande parte do século XX. Naturalmente, a redescoberta do teorema por Wright, em 1983, suscitou, imediatamente, uma suspeita generalizada acerca de **AF** e motivou a busca por uma prova de consistência da teoria.

A primeira prova da consistência de **AF** foi apresentada por John Burgess [10] em 1984

e consistia simplesmente em considerar a $\Sigma_{\mathbf{F}}$-estrutura $\mathfrak{A}$ cujo domínio $|\mathfrak{A}|$ é o conjunto $\{0,1,2,\ldots,\aleph_0\}$ e na qual $\#^{\mathfrak{A}}$ é a função que, a cada subconjunto X de $|\mathfrak{A}|$, associa o número cardinal de X. De fato, é imediato verificar que *PH* é uma fórmula verdadeira nessa estrutura: $\vDash_{\mathfrak{A}} \#F = \#G\ [\sigma]$ se e somente se a cardinalidade $\sigma(F)$ é igual à cardinalidade de $\sigma(G)$ e $\vDash_{\mathfrak{A}} F \approx G$ se e somente se existe uma correlação um a um entre $\sigma(F)$ e $\sigma(G)$; ou seja, se e somente $\sigma(F)$ e $\sigma(G)$ são conjuntos equipotentes. Como a cardinalidade de $\sigma(F)$ é igual a cardinalidade de $\sigma(G)$ se e somente se $\sigma(F)$ é equipotente a $\sigma(G)$, $\vDash_{\mathfrak{A}} \#F = \#G\ [\sigma]$ se e somente se $\vDash_{\mathfrak{A}} F \approx G$ e, assim, $\vDash_{\mathfrak{A}} \#F = \#G \leftrightarrow F \approx G\ [\sigma]$, para toda atribuição σ em $\mathfrak{A}$.

A prova modelo-teórica de Burgess, no entanto, recorre a um princípio que consiste em uma formulação conjuntística informal do próprio Princípio de Hume; i.é., ao fato de que dois conjuntos possuem o mesmo número cardinal se e somente se eles são equipotentes. Esse aspecto foi apontado por Boolos como um indício da presença de uma circularidade, possivelmente viciosa, envolvida na prova. Logo após reproduzí-la em [4], ele afirma:

> Em um passo crucial da prova da consistência (...) da sentença formal denominada Princípio de Hume, fizemos um apelo a um princípio informal que conecta cardinalidade e correspondência um a um que pode ser simbolizado como – o Princípio de Hume. (Fizemos esse apelo quando dissemos que a cardinalidade de $[\sigma(F)]$ é igual à cardinalidade de $[\sigma(G)]$ se e somente se $[\sigma(F)]$ pode ser colocado em correspondência um a um com $[\sigma(G)]$). Essa argumento deveria contar como uma *prova* da consistência do Princípio de Hume? Que garantias qualquer argumento pode nos dar de que uma certa sentença é consistente, se o argumento apela a um princípio uma de cujas formalizações é a própria sentença que estamos tentando provar ser consistente?

A despeito das dúvidas levantadas por Boolos na passagem acima, o fato de que de quaisquer dois subconjuntos de $\{0,1,2,\ldots,\aleph_0\}$ possuem o mesmo número cardinal se e somente se eles são equipotentes é um teorema da teoria de conjuntos de Zermelo-Fraenkel, o que significa que a prova modelo-teórica de Burgess pode ser facilmente convertida em um prova da consistência de **AF** em relação a **ZF**.

Embora a possibilidade de que uma contradição seja algum dia derivada em **ZF** seja altamente improvável, entretanto, podemos nos perguntar se seria possível provar um resultado

acerca da consistência de **AF** que seja mais forte do que aquele representado pela prova de Burgess. A resposta é sim. Um tal resultado fora provado por Boolos em [3], onde ele mostra, em linhas gerais, a consistência de **AF** em relação *Aritmética de Peano de Segunda Ordem* ($\mathbf{AP_2}$), uma segunda teoria para a aritmética elementar cuja consistência está acima de toda e qualquer dúvida razoável[1].

Neste capítulo, iremos apresentar uma elaboração detalhada desse resultado e utilizá-lo para provar a equiconsistência entre **AF**, **ADP** e $\mathbf{AP}_2$. Na seção 3.1, introduziremos $\mathbf{AP_2}$ e provaremos que essa teoria é consistente em relação a **ADP**. Na seção 3.2, apresentaremos a prova de Boolos da consistência de **AF** em relação a $\mathbf{AP_2}$ e mostraremos como ambos os resultados e o Corolorário 2.6 do Teorema de Frege são suficientes para estabelecer a equiconsistência entre **AF**, **ADP** e $\mathbf{AP_2}$.

3.1 Aritmética de Peano de Segunda Ordem

A Aritmética de Peano de Segunda Ordem é uma versão de segunda ordem da *Aritmética de Peano* (**AP**) cujos símbolos não lógicos são apenas a constante individual **0**, para o número 0, e a constante funcional s, para a função *sucessor*[2]:

Definição 3.1. A assinatura da Aritmética de Peano de Segunda Ordem é o par $\Sigma_{\mathbf{P}} = \langle \mathcal{R}^{\mathbf{P}}, \mathcal{F}^{\mathbf{P}} \rangle$ tal que

(i) $\mathcal{R}^{\mathbf{P}}_0 = \{\mathbf{0}\}$.

(ii) Para todo $\tau \neq 0$, $\mathcal{P}^{\mathbf{R}}_{\tau} = \varnothing$.

(iii) $\mathcal{F}^{\mathbf{P}}_{\langle 0,0\rangle} = \{s\}$.

(iii) Para todo $\tau \neq \langle 0,0\rangle$, $\mathcal{F}^{\mathbf{P}}_{\tau} = \varnothing$.

[1] Mais precisamente, a prova de Boolos foi originalmente formulada para uma teoria de segunda ordem ligeiramente diferente de, mas esquivalente a, **AF**. Em [4], nas página seguintes à passagem citada acima, Boolos reapresenta essa prova para a mesma teoria; e, em [9] (Apêndice 2), encontramos uma versão da prova original adaptada para **AF**.

[2] Para uma apresentação da Aritmética de Peano, ver [17], capítulos 20, 21 e 22.

Assim, $\Sigma_{\mathbf{P}}$ não inclui as constantes funcionais + e × presentes na linguagem de **AP**. Como veremos, esses símbolos podem ser definidos em $\mathbf{AP_2}$ e os axiomas de **AP** para as operações de *soma* e *multiplicação* nos números naturais, facilmente provados a partir da definições fornecidas e dos axiomas próprios de $\mathbf{AP_2}$.

Definição 3.2. A *Aritmética de Peano de Segunda Ordem* ($\mathbf{AP_2}$) é a $\Sigma_{\mathbf{P}}$-teoria cujos axiomas próprios são

($\mathbf{AP_2}1$) $sx \neq \mathbf{0}$

($\mathbf{AP_2}2$) $sx = sy \rightarrow x = y$

($\mathbf{AP_2}3$) $F\mathbf{0} \wedge \forall x(Fx \rightarrow Fsx) \rightarrow \forall xFx$

Na interpretação pretendida de $\mathbf{AP_2}$, $\mathbf{AP_2}1$ e $\mathbf{AP_2}2$ expressam, respectivamente, que 0 não é o sucessor de nenhuma número natural e que a função sucessor é injetiva; e $\mathbf{AP_2}3$ é o princípio de indução matemática.

Note-se que, diferentemente de **ADP**, $\mathbf{AP_2}$ não possui uma constante relacional que formaliza o conceito dos números naturais e que, portanto, nenhum de seus axiomas é relativizado a uma tal constante. Isso significa que $\mathbf{AP_2}$ é uma teoria estritamente mais forte do que **ADP**, no sentido de que não existe uma extensão definicional de **ADP** da qual $\mathbf{AP_2}$1-3 resultam ser teoremas. No entanto, é possível provar as relativizações desses axiomas a *Nat* em uma extensão definicional apropriada de **ADP**:

Teorema 3.3. $\mathbf{AP_2}$ é interpretável em uma extensão definicional de **ADP**.

Prova: Primeiramente, estendemos a linguagem de **ADP**, introduzindo a constante funcional *s* mediante a seguinte definição:

(D_s)

$$sx = y \leftrightarrow \big((Nat(x) \rightarrow Pred(x,y)) \wedge (\neg Nat(x) \rightarrow y = \mathbf{0})\big)$$

cuja legitimidade é garantida pelos axiomas **ADP**3 e **ADP**6[3]. Sejam $\Sigma_{\mathbf{DP}^{df}}$ a assinatura que resulta de $\Sigma_{\mathbf{D}}$ pela introdução de *s* a $\mathcal{F}^{\mathbf{D}}_{\langle 0,0\rangle}$ e $\mathbf{ADP}^{df}$ a extensional definicional correspondente

[3] D_s define o sucessor de um número natural *x* como o único *y* tal que *x* precede *y* e o sucessor de qualquer outro objeto que não seja um número natural como 0.

(os axiomas próprios de $\mathbf{ADP}^{df}$ são, portanto, **ADP**1 – **ADP**7 e D_s). Definimos, em segundo lugar, uma interpretação I de $\mathcal{L}(\Sigma_{\mathbf{P}})$ em $\mathcal{L}(\Sigma_{\mathbf{DP}^{df}})$ (ver Definição 1.41):

(i) $\forall^I = Nat$.

(ii) $\mathbf{0}^I = \mathbf{0}$.

(iii) $s^I = s$.

É imediato mostrar que I é uma interpretação de $\mathcal{L}(\Sigma_{\mathbf{P}})$ em $\mathbf{ADP}^{df}$ (ver Definição 1.43): por **ADP**1, $\vdash_{\mathbf{ADP}^{df}} Nat(\mathbf{0})$ e, portanto, $\vdash_{\mathbf{ADP}^{df}} \exists x Nat(x)$. Para provar que Nat é fechada sob s, suponha $Nat(x)$. De D_s e $sx = sx$, inferimos $Pred(x,sx)$ e, por **ADP**2, $Nat(sx)$. Logo, $\vdash_{\mathbf{ADP}^{df}} Nat(x) \to Nat(sx)$.

Por fim, provamos que I é, de fato, uma interpretação de $\mathbf{AP_2}$ em $\mathbf{AD}_2^{df}$ (Definição 1.45), o que, pelo Teorema 1.47, requer mostrar apenas que as relativizações de $\mathbf{AP_2}$1-3 à constante relacional Nat são teoremas de $\mathbf{AD}_2^{df}$ (ver Definição 1.44):

(i) $\vdash_{\mathbf{ADP}^{df}} Nat(x) \to sx \neq \mathbf{0}$

Demonstração: Suponha $Nat(x)$ e $sx = \mathbf{0}$. Por D_s, obtemos $Nat(x) \to Pred(x,\mathbf{0})$ e, portanto, $Pred(x,\mathbf{0})$. Logo, $\exists x(Nat(x) \wedge Pred(x,\mathbf{0}))$, o que contradiz **ADP**5.

■

(ii) $\vdash_{\mathbf{ADP}^{df}} Nat(x) \to \forall y(Nat(y) \to (sx = sy \to x = y))$

Demonstração: Suponha $Nat(x)$, $Nat(y)$ e $sx = sy$. Por D_s, temos $Pred(x,sx)$ e $Pred(y,sy)$ (pois $sx = sx$ e $sy = sy$). Portanto, $Pred(y,sx)$. Por **ADP**4, inferimos $x = y$.

■

(iii) $\vdash_{\mathbf{ADP}^{df}} \forall x(Fx \to Nat(x)) \to (F\mathbf{0} \wedge \forall x(Nat(x) \to (Fx \to Fsx)) \to \forall x(Nat(x) \to Fx)$

Demonstração: Suponha $\forall x(Fx \to Nat(x))$ e $F\mathbf{0} \wedge \forall x(Nat(x) \to (Fx \to Fsx))$. Suponha também $Nat(x_1) \wedge Pred(x_1,x_2)$ e Fx_1. Por D_s, obtemos $sx_1 = x_2$. Intanciando x para x_1 em $\forall x(Nat(x) \to (Fx \to Fsx))$, obtemos $Nat(x_1) \to (Fx_1 \to Fsx_1)$. De $Nat(x_1)$ e Fx_1, inferimos

Fsx_1 e, como $sx_1 = x_2$, Fx_2. Logo,

$$(*)\ \forall x_1 \forall x_2 (Nat(x_1) \wedge Pred(x_1, x_2) \rightarrow (Fx_1 \rightarrow Fx_2))$$

De **ADP**7, $F\mathbf{0}$ e $(*)$, inferimos $\forall x(Nat(x) \rightarrow Fx)$.

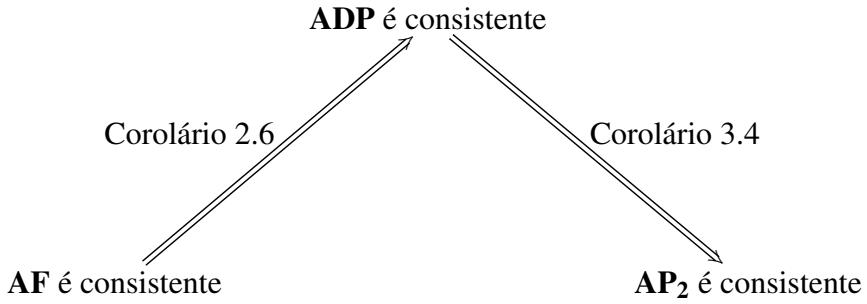
■

■

Como consequência do teorema acima e dos teoremas 1.38 e 1.46, obtemos:

Corolário 3.4. Se **ADP** é consistente, então $\mathbf{AP_2}$ é consistente.

O qual, juntamente com o Corolário 2.6, estabelece a seguinte relação entre **AF**, **ADP** e $\mathbf{AP_2}$:



Assim, para fechar o ciclo representado pelo diagrama e estabelecer a equiconsistência entre **AF**, **ADP** e $\mathbf{AP_2}$, é preciso mostrar apenas que se $\mathbf{AP_2}$ é consistente, então **AF** é consistente.

3.2 Uma Prova de Consistência Relativa de AF

Em [3], a estratégia de Boolos para provar a consistência de **AF** em relação à consistência de $\mathbf{AP_2}$ consiste em considerar, primeiramente, uma $\Sigma_{\mathbf{F}}$-estrutura definida em termos do conjunto dos números naturais na qual PH é uma fórmula verdadeira, ou seja, em considerar uma prova modelo-teórica da consistência da Aritmética de Frege; e, em segundo lugar, em mostrar como essa estrutura pode ser convertida em uma interpretação de **AF** em (uma extensão definicional de) $\mathbf{AP_2}$.

3.2.1 Uma Prova Modelo-Teórica da Consistência de AF

Teorema 3.5. **AF** possui modelo.

Prova: Cosidere $\Sigma_{\mathbf{F}}$-estrutura $\mathfrak{F}$ definida por

(i) $|\mathfrak{F}| = N$

(ii) Para todo $X \in |\mathfrak{A}|_{\langle 0\rangle} = \mathcal{P}(N)$,

$$\#^{\mathfrak{F}}(X) = \begin{cases} n+1 & \text{se } X \text{ é um conjunto finito de números naturais e possui } n \text{ elementos} \\ 0 & \text{se } X \text{ é infinito} \end{cases}$$

Seja σ uma atribuição arbitrária em $\mathfrak{F}$. Mostraremos que $\models_{\mathfrak{F}} \#F = \#G \leftrightarrow F \approx G\ [\sigma]$.

1. ($\Rightarrow$) Suponha que $\models_{\mathfrak{F}} \#F = \#G\ [\sigma]$. Logo, $\#^{\mathfrak{F}}(\sigma(F)) = \#^{\mathfrak{F}}(\sigma(G))$. Temos dois casos:

   (a) $\sigma(F)$ é um conjunto finito de números naturais que possui n elementos. Logo, $\#^{\mathfrak{F}}(\sigma(F)) = \#^{\mathfrak{F}}(\sigma(G)) = n+1$ e $\sigma(G)$ também possui n elementos. Portanto, $\sigma(F)$ é equipotente a $\sigma(G)$ e, assim, $\models_{\mathfrak{F}} F \approx G\ [\sigma]$[4].

   (b) $\sigma(F)$ é um conjunto infinito de números naturais. Logo, $\#^{\mathfrak{F}}(\sigma(F)) = \#^{\mathfrak{F}}(\sigma(G)) = 0$ e $\sigma(G)$ também é infinito. Como quaisquer dois conjuntos infinitos de números naturais são equipotentes, $\sigma(F)$ é equipotente a $\sigma(G)$ e, assim, $\models_{\mathfrak{F}} F \approx G\ [\sigma]$.

2. ($\Leftarrow$) Suponha que $\models_{\mathfrak{F}} F \approx G\ [\sigma]$. Logo, $\sigma(F)$ e $\sigma(G)$ são conjuntos equipotentes. Temos, novamentes, dois casos:

   (a) $\sigma(F)$ é um conjunto finito de números naturais e possui n elementos. Logo, $\sigma(G)$ possui n elementos (pois $\sigma(F)$ é equipotente a $\sigma(G)$) e $\#^{\mathfrak{F}}(\sigma(F)) = \#^{\mathfrak{F}}(\sigma(G)) = n+1$. Portanto, $\models_{\mathfrak{F}} \#F = \#G\ [\sigma]$.

   (b) $\sigma(F)$ é um conjunto inifinito de números naturais. Logo, $\sigma(G)$ é infinito e $\#^{\mathfrak{F}}(\sigma(F)) = \#^{\mathfrak{F}}(\sigma(G)) = 0$. Portanto, $\models_{\mathfrak{F}} \#F = \#G\ [\sigma]$.

[4] É imediato verificar que $\models_{\mathfrak{F}} F \approx G\ [\sigma]$ se e somente se existe uma função bijetiva do conjunto $\sigma(F)$ em $\sigma(G)$; i.é., se, e somente se $\sigma(F)$ e $\sigma(G)$ são conjuntos equipotentes de números naturais.

Concluímos que para toda atribuição σ em $\mathfrak{F}$, $\models_{\mathfrak{F}} \#F = \#G\,[\sigma]$ se e somente se $\models_{\mathfrak{F}} F \approx G\,[\sigma]$ e, portanto, que $\models_{\mathfrak{F}} \forall F \forall G(\#F = \#G \leftrightarrow F \approx G)$, i.é., que PH é verdadeiro em $\mathfrak{F}$.

■

3.2.2 O Operador η

O procedimento para converter a $\Sigma_{\mathbf{F}}$-estrutura $\mathfrak{F}$ acima em uma interpretação de **AF** em (uma extensão definicional de) $\mathbf{AP_2}$ consiste em replicar formalmente, no interior de $\mathbf{AP_2}$, a função $\#^{\mathfrak{F}}$ e provar uma formalização correspondente do correlato modelo-teórico do Princípio de Hume na estrutura $\mathfrak{F}$; i.é., provar uma formalização do fato de que $\#^{\mathfrak{F}}$ associa o mesmo número natural a dois subconjuntos de N se e somente se eles são equipotentes.

Para formalizar a função $\#^{\mathfrak{F}}$ em $\mathbf{AP_2}$, será preciso formular primeiramente uma série de definições preliminares na teoria, a começar por algumas operações e relações aritméticas elementares utilizadas na caracterização da estrutura $\mathfrak{F}$:

(D_+)

$$x+y = x_1 \leftrightarrow \forall R(R(x,\mathbf{0},x) \wedge \forall z \forall w(R(x,z,w) \rightarrow R(x,sz,sw)) \rightarrow R(x,y,x_1))^5$$

$(D_<)$

$$x < y \leftrightarrow \exists z\; sx+z = y$$

$(D_\leq)$

$$x \leq y \leftrightarrow x < y \vee x = y$$

(D_{Nat})

$$Nat(x) \leftrightarrow x = x$$

$(D_\mathbf{1})$

$$\mathbf{1} = s\mathbf{0}$$

[5] A rigor, deveríamos escrever $+xy$ no lugar de $x+y$. No entanto, iremos adotar a notação infixa para facilitar a leitura das fórmulas, inserindo, sempre que necessário, os parênteses externos para evitar ambiguidades.

D_+ define a operação de soma nos números naturais como a menor relação S para a qual valem as condições $S(x,\mathbf{0},x)$ e $\forall z\forall w(S(x,z,w)\to S(x,sz,sw))$, que correspondem aos axiomas

$x+\mathbf{0}=x$

$x+sz=s(x+z)$

de **AP** – os quais são facilmente obtidos a partir da definição. Embora não utilizaremos a operação de multiplicação nos desenvolvimentos subsequentes, a constante funcional $\times$ também pode ser definida em $\mathbf{AP_2}$ de maneira similar:

$$x\times y=x_1\leftrightarrow\forall R(R(x,\mathbf{0},\mathbf{0})\wedge\forall z\forall w(R(x,z,w)\to R(z,sz,w+x)\to R(x,y,x_1))$$

e os axiomas $x\times\mathbf{0}=\mathbf{0}$ e $x\times sz=(x\times z)+x$ de **AP**, provados a partir dessa definição.

Além de $D_+,D_<,D_\leq,D_{Nat}$ e $D_\mathbf{1}$, também introduzimos em $\mathbf{AP_2}$ as definições $D_{\approx_R}$ e $D_\approx$ para a relação de equinumericidade (seção 2.1), e assumimos, sem prova, alguns teoremas elementares envolvendo os símbolos definidos. Dentre eles, destacamos o fato de que $<$ é um relação de ordem nos números naturais e o *princípio do menor número*:

$$\exists xFx\to\exists x(Fx\wedge\forall y(y<x\to\neg Fy))$$

obtido de $\mathbf{AP_2}3$ e que estabelece que, para todo conceito F não vazio de números naturais, existe um menor número que cai sob F.

Para replicar a função $\#^\mathfrak{F}$ em $\mathbf{AP}_2$, definimos primeiramente as constantes relacionais Fin e Eta de tipo $\langle\langle 0\rangle,0\rangle$:

(D_{Fin})

$$Fin(F,x)\leftrightarrow[u:u<x]\approx F$$

(D_{Eta})

$$Eta(F,x)\leftrightarrow(\exists y(Fin(F,y)\wedge x=y+\mathbf{1}))\vee(\neg\exists yFin(F,y)\wedge x=\mathbf{0}))$$

$Fin(F,x)$ expressa que o conceito dos números naturais menores do que x é equinumérico a F e $\exists xFin(F,x)$ expressa, portanto, que F é um conceito finito. $Eta(F,x)$, por sua vez, expressa que ou existe um número y tal que $Fin(F,y)$ e $x = y + \mathbf{1}$ ou F é infinito e $x = \mathbf{0}$.

Fazendo uso de Eta, formalizamos $\#^{\mathfrak{F}}$ através da constante funcional η de tipo $\langle\langle 0\rangle, 0\rangle$, introduzida em $\mathbf{AP_2}$ mediante a seguinte definição:

(D_η)

$$\eta F = x \leftrightarrow Eta(F,x)$$

D_η é a correspodente formal da definição da função $\#^{\mathfrak{F}}$ na $\Sigma_{\mathbf{F}}$-estrutura $\mathfrak{F}$: ela reproduz, no interior de $\mathbf{AP_2}$, a caracterização dessa função envolvida na descrição de $\mathfrak{F}$. Em se tratando de uma definição de uma constante funcional, para legitimar D_η, é preciso mostrar que ela atende às condições de existência e unicidade, provando os seguintes teoremas:

(Eta_1) $\exists xEta(F,x)$

(Eta_2) $Eta(F,x) \wedge Eta(F,y) \rightarrow x = y$

A demonstração de Eta_1 não oferece dificuldade; trata-se de uma consequência imediata de D_{Eta} e a prova procede por casos a partir de $\exists xFin(F,x) \vee \neg\exists xFin(F,x)$. Por outro lado, a demonstração de Eta_2 requer mostrar um lema que estabelece que se o conceito dos números naturais menores do que x é equinumérico ao conceito dos números naturais menores do que y, então $x = y$:

$(Eta_{2.1})$ $[u : u < x] \approx [u : u < y] \rightarrow x = y$

Demonstração: Apêndice C.

■

Com $Eta_{2.1}$ em mãos, a demonstração de Eta_2 é imediata:

Demonstração: Suponha $Eta(F,x) \wedge Eta(F,y)$. Logo,

$$(\exists z(Fin(F,z) \wedge x = z+\mathbf{1})) \vee (\neg \exists z Fin(F,z) \wedge x = \mathbf{0}))$$

e

$$(\exists w(Fin(F,w) \wedge y = w+\mathbf{1})) \vee (\neg \exists w Fin(F,w) \wedge y = \mathbf{0}))$$

Temos, portanto, quatro casos, dentre os quais consideramos apenas o primeiro:

1. $\exists z(Fin(F,z) \wedge x = z+\mathbf{1})$ e $\exists w(Fin(F,w) \wedge y = w+\mathbf{1})$. De $\exists z(Fin(F,z) \wedge x = z+\mathbf{1}) \Rightarrow Fin(F,z) \wedge x = z+\mathbf{1}$ e $\exists w(Fin(F,w) \wedge y = w+\mathbf{1}) \Rightarrow Fin(F,w) \wedge y = w+\mathbf{1}$, obtemos $[u : u < z] \approx F$ e $[u : u < w] \approx F$. Por $\approx_1$-$\approx_3$, inferimos $[u : u < z] \approx [u : u < w]$, e por $Eta_{2.1}$, obtemos $z = w$. Logo, $z+\mathbf{1} = w+\mathbf{1} \Rightarrow x = y$.

■

Tendo legitimado a definição do operador η, podemos agora formalizar em $\mathbf{AP}_2$ o correspondente modelo-teórico do Princípio de Hume na $\Sigma_{\mathbf{F}}$-estrutura $\mathfrak{F}$:

$(P\eta)$

$$\eta F = \eta G \leftrightarrow F \approx G$$

cuja demonstração, como veremos, pressupõe a seguinte propriedade de Fin:

(Fin_1) $\neg \exists x Fin(F,x) \leftrightarrow Nat \approx F$

Fin_1 estabelece que um conceito F de números naturais é infinito se, e somente se, o conceito dos números naturais é equinumérico a F, o que, na presença de $\approx_2$ e $\approx_3$, implica que todos os conceitos infinitos de números naturais são equinuméricos e possuem, portanto, o mesmo tamanho ou cardinalidade. A demonstração de Fin_1 em $\mathbf{AP_2}$ requer uma série de teoremas preliminares, cujas provas ocuparão grande parte do restante deste capítulo; e, antes demonstrá-lo, iremos mostrar como $P\eta$ pode ser obtido unicamente a partir dele. A demonstração abaixo é uma réplica formal da prova do Teorema 3.3:

Demonstração: ($\Rightarrow$) Suponha $\eta F = \eta G$. Por D_η, obtemos $Eta(F, \eta G)$. Logo,

$$\exists x(Fin(F,x) \wedge \eta G = x+\mathbf{1}) \vee (\neg \exists x Fin(F,x) \wedge \eta G = \mathbf{0})$$

Temos dois casos:

1. $\exists x(Fin(F,x) \wedge \eta G = x+\mathbf{1}) \Rightarrow Fin(F,x) \wedge \eta G = x+\mathbf{1}$. Logo, $Eta(G, x+\mathbf{1})$ e, portanto,

   $$\exists y(Fin(G,y) \wedge x+\mathbf{1} = y+\mathbf{1}) \vee (\neg \exists y Fin(G,y) \wedge x+\mathbf{1} = \mathbf{0})$$

   Temos dois subcasos:

   (a) $\exists y(Fin(G,y) \wedge x+\mathbf{1} = y+\mathbf{1}) \Rightarrow Fin(G,y) \wedge x+\mathbf{1} = y+\mathbf{1}$. Logo, $x = y$ e, portanto, $Fin(G,x)$. De $Fin(F,x)$ e $Fin(G,x)$, inferimos $[u : u < x] \approx F$ e $[u : u < x] \approx G$ e, por $\approx_2$ e $\approx_3$, obtemos $F \approx G$.

   (b) $\neg \exists y Fin(G,y) \wedge x+\mathbf{1} = \mathbf{0}$. Logo, $sx = \mathbf{0}$, o que contradiz $\mathbf{AP_2}2$. Absurdo.

2. $\neg \exists x Fin(F,x) \wedge \eta G = \mathbf{0}$. Por Fin_1, inferimos $Nat \approx F$. De $\eta G = \mathbf{0}$, obtemos então $Eta(G, \mathbf{0})$. Logo,

   $$\exists y(Fin(G,y) \wedge \mathbf{0} = y+\mathbf{1}) \vee (\neg \exists y Fin(G,y) \wedge \mathbf{0} = \mathbf{0})$$

   Temos, novamente, dois subcasos:

   (a) $\exists y(Fin(G,y) \wedge \mathbf{0} = y+\mathbf{1}) \Rightarrow Fin(G,y) \wedge \mathbf{0} = y+\mathbf{1}$. Logo, $sy = \mathbf{0}$. Absurdo.

   (b) $\neg \exists y Fin(G,y) \wedge \mathbf{0} = \mathbf{0}$. Por Fin_1, obtemos $Nat \approx G$ e, por $\approx_2$ e $\approx_3$, obtemos $F \approx G$.

($\Leftarrow$) Suponha $F \approx G$. Temos dois casos:

1. $\exists x Fin(F,x) \Rightarrow Fin(F,x) \Rightarrow [u : u < x] \approx F$. Logo, $[u : u < x] \approx G$ (por $\approx_3$) e, portanto, $Fin(G,x) \Rightarrow Fin(G,x) \wedge x+\mathbf{1} = x+\mathbf{1} \Rightarrow \exists y(Fin(G,y) \wedge x+\mathbf{1} = y+\mathbf{1}) \Rightarrow Eta(G, x+\mathbf{1}) \Rightarrow \eta G = x+\mathbf{1}$. Logo, $\exists x(Fin(F,x) \wedge \eta G = x+\mathbf{1}) \Rightarrow Eta(F, \eta G) \Rightarrow \eta F = \eta G$.

2. $\neg \exists x Fin(F,x)$. Por Fin_1, temos $Nat \approx F$ e, portanto, $Nat \approx G$. Por Fin_1 novamente, obtemos $\neg \exists y Fin(G,y) \Rightarrow \neg \exists y Fin(G,y) \wedge \mathbf{0} = \mathbf{0} \Rightarrow Eta(G, \mathbf{0}) \Rightarrow \eta G = \mathbf{0}$. Logo, $\neg \exists x Fin(F,x) \wedge \eta G = \mathbf{0} \Rightarrow Eta(F, \eta G) \Rightarrow \eta F = \eta G$.

■

3.2.3 Demonstração de Fin_1

Como se pode notar, Fin_1 é tudo o que precisamos mostrar para nos certificarmos que $P\eta$ é, de fato, um teorema de $\mathbf{AP_2}$. Porém, embora sua formulação seja extremamente simples e seu conteúdo, intuitivamente claro, sua demonstração é algo longa e trabalhosa: teremos de introduzir dois novos conceitos auxiliares, formalizados pelas constantes relacionais *Cont* e *Enum* – de tipos $\langle\langle 0,0\rangle,\langle 0\rangle,0\rangle$ e $\langle\langle 0,0\rangle,0\rangle$, repectivamente – e estabelecer algumas de suas propriedades.

Diremos que R é uma x_1-*enumeração* de um conceito F, quando R é uma relação funcional que enumera os x_1 primeiros números que caem sob F de acordo com a ordem $<$. Formalmente:

(D_{Cont})

$$Cont(R,F,x_1) \leftrightarrow \forall x\forall y\forall z(Rxy \wedge Rxz \rightarrow y = z) \wedge \forall x(\exists y Rxy \leftrightarrow x < x_1) \wedge$$

$$\forall x\forall y(Rxy \rightarrow Fy) \wedge \forall x\forall y\forall z\forall w(Rxy \wedge Rzw \wedge x < z \rightarrow y < w) \wedge$$

$$\forall x\forall y\forall w(Fw \wedge Rxy \wedge w < y \rightarrow \exists z(z < x \wedge Rzw))$$

A primeira conjunta do lado direito de D_{Cont} expressa que R é uma relação funcional; a segunda e a terceira conjuntas expressam, respectivamente, que o "domínio" de R é o conceito dos números menores do que x_1 e que todos os objetos que caem sob o "contra-domínio" de R também caem sob F. A quarta conjunta exige que R preserve a ordem $<$, e a quinta, que a enumeração não deixe "lacunas" em F.

De maneira análoga a D_{Cont}, definimos quando uma relação R é uma *enumeração* de um conceito F; a única diferença é que o "domínio" de R é fixado como o conceito dos números naturais:

(D_{Enum})

$$Enum(R,F) \leftrightarrow \forall x\forall y\forall z(Rxy \wedge Rxz \rightarrow y = z) \wedge \forall x(\exists y Rxy \leftrightarrow Nat(x)) \wedge$$

$$\forall x \forall y (Rxy \to Fy) \wedge \forall x \forall y \forall z \forall w (Rxy \wedge Rzw \wedge x < z \to y < w) \wedge$$

$$\forall x \forall y \forall w (Fw \wedge Rxy \wedge w < y \to \exists z (z < x \wedge Rzw))$$

Se existe x_1 tal que R é uma x_1-enumeração de F, diremos que R é uma *enumeração parcial* de F; e se existe R tal que R é uma x_1-enumeração/enumeração de F, diremos que F um conceito é *x_1-enumerável/enumerável*.

É fácil verificar que toda relação que preserva $<$ é injetiva; ou seja, que a fórmula

$$\forall x \forall y \forall z \forall w (Rxy \wedge Rzw \wedge x < z \to y < w) \to \forall x \forall y \forall z (Ryx \wedge Rzx \to y = z)$$

é um teorema de $\mathbf{AP_2}$. Portanto, as seguintes fórmulas também são teoremas de $\mathbf{AP_2}$:

$(Cont_1)$ $Cont(R,F,x_1) \to \forall x \forall y \forall z \forall w (Rxy \wedge Rzw \to (x = z \leftrightarrow y = w))$

$(Enum_1)$ $Enum(R,F) \to \forall x \forall y \forall z \forall w (Rxy \wedge Rzw \to (x = z \leftrightarrow y = w))$

$Enum_1$ será útil para provarmos $Enum_3$ abaixo, que afirma que todo conceito enumerável é equinumérico ao conceito dos números naturais e cuja demonstração requer o fato de que nenhuma enumeração associa a um número x um número menor do que x:

$(Enum_2)$ $Enum(R,F) \to \forall x \forall y (Rxy \to x \leq y)$

Demonstração: Apêndice C.

■

$(Enum_3)$ $\exists R\, Enum(R,F) \to Nat \approx F$

Demonstração: Suponha $\exists R\, Enum(R,F) \Rightarrow Enum(R,F)$.

(i) Para provar $\forall x (Nat(x) \to \exists y (Fy \wedge Rxy))$, suponha $Nat(x)$. Logo, $\exists y Rxy \Rightarrow Rxy$ e, portanto, Fy. Assim, $Fy \wedge Rxy \Rightarrow \exists y (Fy \wedge Rxy)$.

(ii) Para provar $\forall y (Fy \to \exists x (Nat(x) \wedge Rxy))$, suponha Fy. Como $Nat(y)$, temos $\exists z Ryz \Rightarrow Ryz$ e, portanto, Fz. Por $Enum_2$, obtemos $y \leq z \Rightarrow y < z \vee y = z$. Como $Nat(z)$, $\exists w Rzw \Rightarrow Rzw$

e, portanto, $z \leq w$ (por $Enum_2$ novamente). Se $y < z$, então $y < w$, o que, juntamente com Fy e Rzw, implica $\exists x(x < z \wedge Rxy) \Rightarrow x < z \wedge Rxy$. Logo, $\exists x(Nat(x) \wedge Rxy)$. Se, por outro lado, $y = z$, então Ryy e, portanto, $\exists x(Nat(x) \wedge Rxy)$.

De (i), (ii) e $Enum_1$, inferimos $Nat \approx F$.

■

De acordo com $Enum_3$, a estratégia para demonstrar a direção esquerda-direita de Fin_1 consistirá em mostrar que todo conceito infinito de números naturais é enumerável; i.é., em mostrar $Enum_4$:

$(Enum_4)\ \neg\exists x Fin(F,x) \rightarrow \exists R\, Enum(R,F)$

que, por sua vez, requer provar dois lemas auxiliares sobre $Cont$. O primeiro deles, $Cont_2$ abaixo, estabelece que qualquer conceito infinito F de números naturais é x_1-enumerável, para todo número natural x_1:

$(Cont_2)\ \neg\exists x Fin(F,x) \rightarrow \forall x_1 \exists R\, Cont(R,F,x_1)$

Demonstração: Suponha $\neg\exists x Fin(F,x)$. Instanciando F para $[u : \exists R\, Cont(R,F,u)]$ na variação alfabética $F\mathbf{0} \wedge \forall x_1(Fx_1 \rightarrow Fsx_1) \rightarrow \forall x_1 Fx_1$ de $\mathbf{AP_2}3$, obtemos

$[u : \exists R\, Cont(R,F,u)]\mathbf{0} \wedge$

$\forall x_1([u : \exists R\, Cont(R,F,u)]x_1 \rightarrow [u : \exists R\, Cont(R,F,u)]sx_1) \rightarrow \forall x[u : \exists R\, Cont(R,F,u)]x_1$

(i) Para provar $[u : \exists R\ Cont(R,F,u)]\mathbf{0}$, considere a relação vazia $[uu_1 : u \neq u \wedge u_1 \neq u_1]$. É imediato mostrar $Cont([uu_1 : u \neq u \wedge u_1 \neq u_1], F, \mathbf{0}) \Rightarrow \exists R\, Cont(R,F,\mathbf{0})$.

(ii) Para provar o passo indutivo, $\forall x_1([u : \exists R\ Cont(R,F,u)]x_1 \rightarrow [u : \exists R\ Cont(R,F,u)]sx_1)$, suponha $[u : \exists R\, Cont(R,F,u)]x_1 \Rightarrow \exists R\, Cont(R,F,x_1) \Rightarrow Cont(R,F,x_1)$. Vamos mostrar

$$[u : \exists R\, Cont(R,F,u)]sx_1$$

Considere o conceito dos números que caem sob F mas que não são correlacionados por

R a nenhum número natural: $[u : Fu \wedge \neg\exists xRxu]$; e suponha que esse conceto é não vazio: $\neg\exists y[u : Fu \wedge \neg\exists xRxu]y \Rightarrow (*)\ \forall y(Fy \rightarrow \exists xRxy)$.

Suponha $[u : u < x_1]x \Rightarrow x < x_1$. De $Con(R,F,x_1)$, inferimos $\exists yRxy \Rightarrow Rxy \Rightarrow Fy$. Logo, $\exists y(Fy \wedge Rxy)$ e, portanto,

$$(**)\ \forall x([u : u < x_1]x \rightarrow \exists y(Fy \wedge Rxy))$$

Suponha agora Fy. Por $(*)$, obtemos $\exists xRxy \Rightarrow Rxy \Rightarrow \exists yRxy$. De $Cont(R,F,x_1)$, inferimos $x < x_1 \Rightarrow [u : u < x_1]x$. Logo, $\exists x([u : u < x_1]x \wedge Rxy)$ e, portanto,

$$(***)\ \forall y(Fy \rightarrow \exists x([u : u < x_1]x \wedge Rxy))$$

De $(**)$, $(***)$ e $Cont_1$, inferimos então $[u : u < x_1] \approx F \Rightarrow Fin(F,x_1) \Rightarrow \exists xFin(F,x)$, o que contradiz a hipótese inicial. Logo, $\exists y[u : Fu \wedge \neg\exists xRxu]y \Rightarrow \exists y_1[u : Fu \wedge \neg\exists xRxu]y_1$. Pelo princípio do menor número, obtemos

$$\exists y_1([u : Fu \wedge \neg\exists xRxu]y_1 \wedge \forall w(w < y_1 \rightarrow \neg[u : Fu \wedge \neg\exists xRxu]w)) \Rightarrow$$
$$(\dagger)\ [u : Fu \wedge \neg\exists xRxu]y_1 \wedge \forall w(w < y_1 \rightarrow \neg[u : Fu \wedge \neg\exists xRxu]w)$$

Considere agora a relação $R' := [uu_1 : (u = x_1 \wedge u_1 = y_1) \vee Ruu_1]$. Fazendo uso de $(\dagger)$ acima, mostramos $Cont(R',F,sx_1)$:

As primeiras três conjuntas do lado direito da instância correspondente de D_{Cont} são imediatas. Para mostrar a quarta conjunta, $\forall x\forall y\forall w(Fw \wedge R'xy \wedge w < y \rightarrow \exists z(z < x \wedge R'zw))$, suponha $Fw \wedge R'xy \wedge w < y$. Logo, $(x = x_1 \wedge y = y_1) \vee Rxy$. Temos dois casos:

1. $x = x_1 \wedge y = y_1$. De $w < y$, inferimos $w < y_1$. Por $(\dagger)$, obtemos $\neg[u : Fu \wedge \neg\exists xRxu]w \Rightarrow \neg Fw \vee \exists xRxw$. Como Fw, temos $\exists xRxw \Rightarrow Rzw$ e, portanto, $R'zw \Rightarrow \exists yR'zy$. Logo, $z < sx_1$ (pela segunda conjunta no lado direito de D_{Cont}) e, assim, $z \leq x_1 \Rightarrow z \leq x$. Suponha $z = x$. De $R'xy$ e $R'xw$, obtemos $y = w$ (pela funcionalidade de R'). Absurdo (pois $w < y$, por hipótese). Logo, $z < x$ e, portanto, $\exists z(z < x \wedge R'zw) \Rightarrow \exists R\ Cont(R,F,sx_1) \Rightarrow [u : \exists R\ Cont(R,F,u)]sx_1$.

2. Rxy. De Fw, Rxy e $w < y$, obtemos $\exists z(z < x \wedge Rzw) \Rightarrow z < x \wedge Rzw$. De Rzw, inferimos $R'zw$ e, portanto, $\exists z(z < x \wedge R'zw)$.

De (i) e (ii), inferimos $\forall x_1[u : \exists R\, Cont(R,F,u)]x_1 \Rightarrow \forall x_1 \exists R\, Cont(R,F,x_1)$.

■

O segundo lema sobre *Cont* necessário na demonstração de $Enum_4$ estabelece que quaisquer duas enumerações parciais de um conceito de números naturais são compatíveis entre si, ou seja, que, dadas uma x_1-enumeração R e uma y_1-enumeração S de um conceito F, em que $x_1 \leq y_1$, R e S coincidem em todos os números menores do que x_1: elas associam os mesmos números que caem sob F aos números naturais menores do que x_1:

($Cont_3$) $Cont(R,F,x_1) \wedge Cont(S,F,y_1) \wedge x_1 \leq y_1 \rightarrow \forall x(x < x_1 \rightarrow \forall y \forall z(Rxy \wedge Sxz \rightarrow y = z))$

Demonstração: Suponha $Cont(R,F,x_1) \wedge Cont(S,F,y_1) \wedge x_1 \leq y_1$ e considere o conceito

$$G := [u : u < x_1 \rightarrow \forall y \forall z(Ruy \wedge Suz \rightarrow y = z)]$$

Instanciando F para G em $\mathbf{AP_2}3$, obtemos

$$G\mathbf{0} \wedge \forall x(Gx \rightarrow Gsx) \rightarrow \forall x Gx$$

(i) Para provar $G\mathbf{0}$, suponha $\mathbf{0} < x_1$ e $R0y \wedge S0z$. Logo, Fy e Fz, por D_{Cont}. Sem perda de generalidade, suponha $y < z$ (para proceder por absurdo). Temos, portanto, Fy, $S\mathbf{0}z$ e $y < z$. Por D_{Cont} novamente, obtemos $\exists x_2(x_2 < \mathbf{0} \wedge Sx_2y) \Rightarrow \exists x_2\, x_2 < \mathbf{0}$. Absurdo.

(ii) Para provar $\forall x(Gx \rightarrow Gsx)$, suponha $[u : u < x_1 \rightarrow \forall y \forall z(Ruy \wedge Suz \rightarrow y = z)]x$. Logo, $(*)$ $x < x_1 \rightarrow \forall y \forall z(Ruy \wedge Suz \rightarrow y = z)$. Suponha também $sx < x_1$ e $Rsxy \wedge Ssxz$. Logo, Fy e Fz (por D_{Cont}). Como $x < x_1$ (pois $sx < x_1$) e $x_1 \leq y_1$, temos $x < y_1$. Portanto, $\exists y_2 Rxy_2 \Rightarrow Rxy_2$ e $\exists z_1 Sxz_1 \Rightarrow Sxz_1$. Por $(*)$, obtemos $y_2 = z_1 \Rightarrow Sxy_2$. De Rxy_2, $Rsxy$, Sxz_1, $Ssxz$ e $x < sx$, inferimos $y_2 < y$ e $z_1 < z$ (pois R e S preservam $<$). Sem perda de generalidade, suponha $y < z$ (para proceder por absurdo). De Fy, $Ssxz$ e $y < z$, inferimos $\exists x_2(x_2 < sx \wedge Sx_2y) \Rightarrow x_2 < sx \wedge Sx_2y$. Logo, $x_2 < x \vee x_2 = x$. Se $x_2 < x$, temos Sx_2y, Sxy_2 e $x_2 < x$ e, portanto, $y < y_2$. Absurdo

(pois $y_2 < y$). Se, por outro lado, $x_2 = x$, de Sx_2y e Sx_2y_2, inferimos $y = y_2$. Absurdo. Logo, $x = z$.

De (i) e (ii), inferimos

$$\forall x[u : u < x_1 \to \forall y \forall z(Ruy \wedge Suz \to y = z)]x \;\Rightarrow\; \forall x(x < x_1 \to \forall y \forall z(Rxy \wedge Sxz \to y = z))$$

■

Finalmente, estamos aptos a demonstrar $Enum_4$. Dado um conceito infinito F de números naturais, o procedimento a ser empregado na prova consiste em considerar de uma só vez todas as (infinitas) enumerações parciais de F que $Cont_2$ afirma existir; i.é., em considerar uma relação T que representa a união dessas enumerações: T irá correlacionar todos os pares de números naturais que estão relacionados entre si por ao menos uma enumeração parcial R de F. Fazendo uso de $Cont_3$, mostraremos então que T é, de fato, uma enumeração de F:

($Enum_4$) $\neg \exists x Fin(F,x) \to \exists R\, Enum(R,F)$

Demonstração: Suponha $\neg \exists x Fin(F,x)$ e considere a relação

$$T := [uu_1 : \exists R(Cont(R,F,su) \wedge Ruu_1)]$$

Mostramos $Enum(T,F)$:

(i) Para provar que T é funcional, $\forall x \forall y \forall z(Txy \wedge Txz \to y = z)$, suponha $Txy \wedge Txz$. Logo, $\exists R(Cont(R,F,sx) \wedge Rxy) \Rightarrow Cont(R,F,sx) \wedge Rxy$ e $\exists S(Cont(S,F,sx) \wedge Sxz) \Rightarrow Cont(S,F,sx) \wedge Sxz$. De Rxy e Sxz, inferimos $y = z$, por $Cont_3$.

(ii) Para provar a direção direita-esquerda de $\forall x(\exists y Txy \leftrightarrow Nat(x))$, suponha $Nat(x)$. Por $Cont_2$, temos $\forall x_1 \exists R\, Cont(R,F,x_1)$ e, em particular, $\exists R\, Cont(R,F,sx) \;\Rightarrow\; Cont(R,F,sx)$. De $x < sx$, obtemos $\exists y Rxy \;\Rightarrow\; Rxy$. Logo, $Cont(R,F,sx) \wedge Rxy$ e, assim, $\exists R\ (Cont(R,F,sx) \wedge Rxy) \;\Rightarrow\; Txy \;\Rightarrow\; \exists y Txy$.

(iii) Para provar $\forall x \forall y(Txy \to Fy)$, suponha Txy. Logo,

$$\exists R(Cont(R,F,sx) \wedge Rxy) \;\Rightarrow\; Cont(R,F,sx) \wedge Rxy$$

Portanto, Fy.

(iv) Por fim, para provar $\forall x \forall y \forall w(Fw \wedge Txy \wedge w < y \rightarrow \exists z(z < x \wedge Twz))$, suponha $Fw \wedge Txy \wedge w < y \Rightarrow \exists R(Cont(R,F,sx) \wedge Rxy) \Rightarrow Cont(R,F,sx) \wedge Rxy$. De Fw, Rxy e $w < y$, inferimos $\exists z(z < x \wedge Rzw) \Rightarrow z < x \wedge Rzw$. Seja $R' := [uu_1 : Ruu_1 \wedge u \leq z]$; i.é., R com domínio restrito aos números menores ou iguais a z. É imediato mostrar $Cont(R',F,sz) \wedge R'zw \Rightarrow \exists R(Cont(R,F,sz) \wedge Rzw) \Rightarrow Tzw$. Logo, $z < x \wedge Tzw \Rightarrow \exists z(z < x \wedge Tzw)$.

■

A partir de $Enum_3$ e $Enum_4$, obtemos imediatamente a direção esquerda-direita de Fin_1, que afirma que se F é um conceito infinito de números naturais, então ele é equinumérico ao conceito dos números naturais:

$(Fin_{1.1})$ $\neg \exists x Fin(F,x) \rightarrow Nat \approx F$

Para completar a demonstração de Fin_1, resta agora demonstrar apenas a direção direita-esquerda:

$(Fin_{1.2})$ $Nat \approx F \rightarrow \neg \exists x Fin(F,x)$

e, para isso, é suficiente demonstrar $Fin_{1.2.1}$:

$(Fin_{1.2.1})$ $\neg \exists x Nat \approx [u : u < x_1]$

Demonstração: Suponha $\exists x_1\ Nat \approx [u : u < x_1]$. Logo, $\exists x_1 \big[u_1 : Nat \approx [u : u < u_1]\big] x_1$. Pelo princípio do menor número,

$$\exists x_1 \Big(\big[u_1 : Nat \approx [u : u < u_1]\big] x_1 \wedge \forall w \big(w < x_1 \rightarrow \neg \big[u_1 : Nat \approx [u : u < u_1]\big] w \big) \Big) \Rightarrow$$

$$(*)\ \big[u_1 : Nat \approx [u : u < u_1]\big] x_1 \wedge \forall w \big(w < x_1 \rightarrow \neg \big[u_1 : Nat \approx [u : u < u_1]\big] w \big)$$

Logo, $\exists R\ Nat \approx_R [u : u < x_1] \Rightarrow Nat \approx_R [u : u < x_1]$.

De $Nat(x_1)$, inferimos $\exists y([u : u < x_1]y \wedge Rx_1y) \Rightarrow \exists y\ y < x_1$. Logo, $x_1 \neq \mathbf{0}$. Portanto, $\exists w\ x_1 =$

$sw \Rightarrow x_1 = sw$. De $w < x_1$, inferimos $\exists z(Nat(z) \wedge Rzw) \Rightarrow Nat(z) \wedge Rzw$. Seja

$$S := [uu_1 : (u < z \wedge Ruu_1) \vee (z \leq u \wedge Rsuu_1)]$$

Vamos mostrar $Nat \approx_S [u : u < w]$:

(i) Para provar $\forall x(Nat(x) \rightarrow \exists y([u : u < w]y \wedge Sxy))$, suponha $Nat(x)$. Temos dois casos:

1. $x < z$. De $Nat(x)$, inferimos $\exists y([u : u < x_1]y \wedge Rxy) \Rightarrow [u : u < x_1]y \wedge Rxy$. Logo, $x < z \wedge Rxy \Rightarrow Sxy$. De $y < x_1$ e $x_1 = sw$, obtemos $y \leq w$. Suponha $y = w$. De Rxw e Rzw, inferimos $x = z$, o que contradiz $x < z$. Logo, $y < w \Rightarrow [u : u < w]y$ e, portanto, $\exists y([u : u < w]y \wedge Sxy)$.

2. $z \leq x$. Como $Nat(sx)$, temos $\exists y([u : u < x_1]y \wedge Rsxy) \Rightarrow [u : u < x_1]y \wedge Rsxy$. Logo, $z \leq x \wedge Rsxy \Rightarrow Sxy$. Como $y < x_1$, $y \leq w$. Suponha $y = w$. De $Rsxw$ e Rzw, inferimos $sx = z \Rightarrow x < z$, o que contradiz $z \leq x$. Logo, $y < w$ e, portanto, $\exists y([u : u < w]y \wedge Sxy)$.

(ii) Para provar $\forall y([u : u < w]y \rightarrow \exists x(Nat(x) \wedge Sxy))$, suponha $[u : u < w]y$. Logo, $[u : u < x_1]y \Rightarrow \exists x(Nat(x) \wedge Rxy) \Rightarrow Rxy$. Temos dois casos:

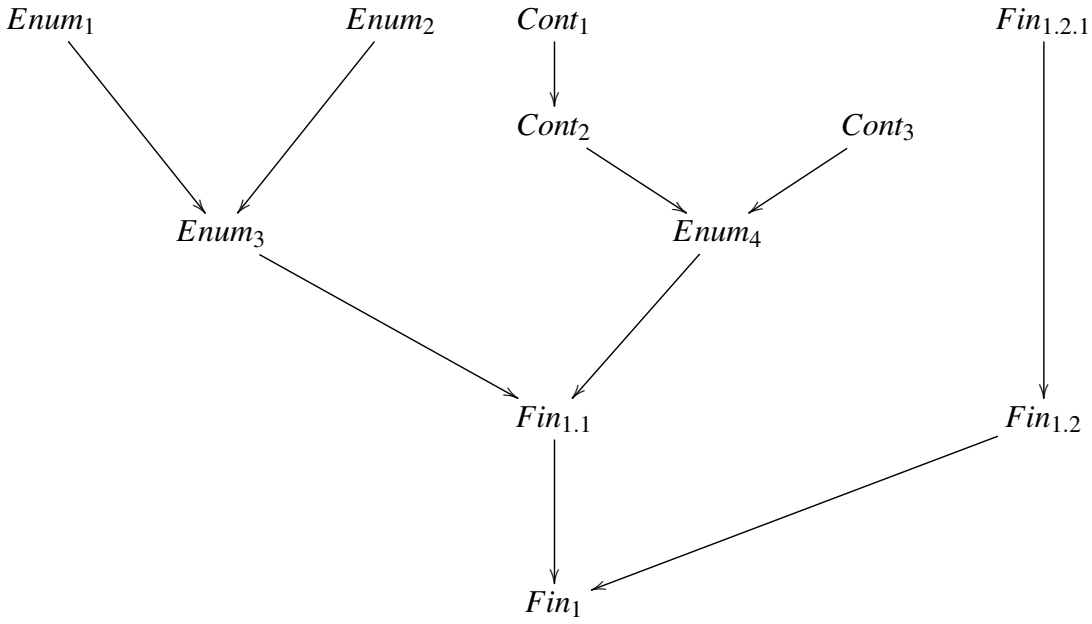
1. $x < z$. Logo, Sxy e, portanto, $\exists x(Nat(x) \wedge Sxy)$.

2. $z \leq x$. Se $z = x$, de Rzw e Rzy, inferimos $w = y$. Absurdo (pois $y < w$, por hipótese). Logo, $z < x$, de onde inferimos $\exists w_1\ x = sw_1 \Rightarrow x = sw_1$. Portanto, $z \leq w_1 \wedge Rsw_1y \Rightarrow Sw_1y \Rightarrow \exists x(Nat(x) \wedge Sxy)$.

(iii) A prova de $\forall x \forall y \forall z \forall w(Sxy \wedge Szw \rightarrow (x = z \leftrightarrow y = w)$ não oferece dificuldade.

De (i), (ii) e (iii), inferimos $Nat \approx [u : u < w] \Rightarrow [u_1 : Nat \approx [u : u < u_1]]w$. Mas $w < x_1$. Portanto, $\neg[u_1 : Nat \approx [u : u < u_1]]$ (por $(*)$). Absurdo. Logo, $\neg\exists x_1\ Nat \approx [u : u < x_1]$.

■



Estrutura da Demonstração de Fin_1

3.2.4 Finalização da Prova

A partir dos desenvolvimentos acima, provamos o seguinte resultado de consistência relativa:

Teorema 3.6. **AF** é interpretável em uma extensão definicional de $\mathbf{AP}_2$.

Prova: Sejam $\Sigma_{\mathbf{P}^{df}}$ a assinatura resultante da introdução dos símbolos $+, <, \leq, Nat, \mathbf{1}, \approx, \eta, Cont$ e $Enum$ a $\Sigma_{\mathbf{P}}$ e $\mathbf{AP}_{\mathbf{2}}^{df}$, a extensão definicional correspondente. Seja I a seguinte interpretação de $\mathcal{L}(\Sigma_{\mathbf{F}})$ em $\mathcal{L}(\Sigma_{\mathbf{P}^{df}})$:

(i) $\forall^I = Nat$

(ii) $\#^I = \eta$

Como $\vdash_{\mathbf{AP}_{\mathbf{2}}^{df}} \exists x Nat(x)$ e $\vdash_{\mathbf{AP}_{\mathbf{2}}^{df}} \forall F(\forall x(Fx \to Nat(x)) \to Nat(\eta F))$ (pois $\vdash_{\mathbf{AP}_{\mathbf{2}}^{df}} \forall x Nat(x)$), I é uma interpretação de $\mathcal{L}(\Sigma_{\mathbf{F}})$ em $\mathbf{AP}_{\mathbf{2}}^{df}$.

Para provar que I é uma interpretação de **AF** em $\mathbf{AP}_{\mathbf{2}}^{df}$, basta notar que, devido ao fato de $\forall x Nat(x)$ ser um teorema de $\mathbf{AP}_{\mathbf{2}}^{df}$, a interpretação de PH por I é equivalente à seguinte

variante de $P\eta$:

$$\forall F \forall G(\eta F = \eta G \leftrightarrow \exists R(\forall x(Fx \rightarrow \exists y(Gy \wedge Rxy)) \wedge \forall y(Gy \rightarrow \exists x(Fx \wedge Rxy)) \wedge$$

$$\forall x \forall y \forall z \forall w(Rxy \wedge Rzw \rightarrow (x = z \leftrightarrow y = w))))$$

■

Como consequência do teorema acima e dos teoremas 1.38 e 1.46, obtemos:

Corolário 3.7. Se $\mathbf{AP_2}$ é consistente, então **AF** é consistente.

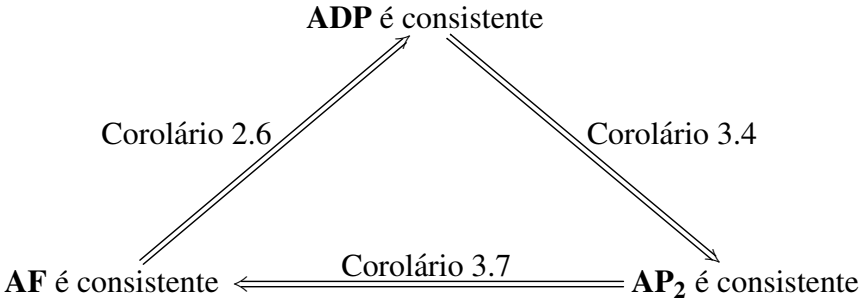
E como consequência dos corolários 2.6, 3.4 e 3.7, obtemos finalmente a equiconsistência entre **AF**, **ADP** e $\mathbf{AP_2}$:

Teorema 3.8. EQUICONSISTÊNCIA ENTRE **AF**, **ADP** E $\mathbf{AP_2}$.

(i) **AF** é consistente se e somente se **ADP** é consistente.

(ii) **AF** é consistente se e somente se $\mathbf{AP_2}$ é consistente.

(iii) **ADP** é consistente se e somente se $\mathbf{AP_2}$ é consistente.



Estrutura da prova do Teorema 3.8

Conclusão

Neste trabalho, apresentamos uma proposta de sistematização do Teorema de Frege.

A partir do diagnóstico de que a lógica de segunda ordem padrão é inadequada para formalizar a linguagem da Aritmética de Frege, propomos, no Capítulo 1, uma caracterização geral de linguagens de segunda ordem apropriadas para esse próposito, recorrendo a tipos para descrever uniformemente o comportamento sintático do operador de cardinalidade # e de outros símbolos definidos envolvidos na prova do Teorema de Frege. Fornecermos também uma semântica e um sistema dedutivo adequados para essas linguagens e estabelecemos algumas definições e resultados metateóricos acerca desse sistema; dentre os quais, destacamos a definição contextual do operador de abstração (Definição 1.22) e o Teorema da Interpretação (Teorema 1.47).

Embora tenhamos restringido nosso enfoque a aspectos técnicos relacionados ao Teorema de Frege, enfatizamos que as linguagens e sistema dedutivo formulados podem ser empregados para formalizar precisamente e manipular quaisquer teorias fregeanas. Esse aspecto representa um subproduto importante deste trabalho, no sentido de fornecer um arcabouço geral no qual essas teorias podem ser discutidas e relacionadas entre si – e também relacionadas com outras teorias formais.

No Capítulo 2, formulamos o Teorema de Frege tecnicamente como uma relação entre a Aritmética de Frege e a Aritmética de Dedekind-Peano e apresentamos uma prova detalhada do teorema, seguindo, tanto quanto possível, o percurso delineado por Frege nas seções 70–83 do *Grundlagen*.

No Capítulo 3, abordamos o tema da consistência da Aritmética de Frege, apresentando

uma elaboração detalhada da prova formulada por Boolos da consistência de **AF** relação a Aritmética de Peano de Segunda Ordem, e utilizando esse resultando para estabelecer a equi-consistência entre ambas as teorias e a Aritmética de Dedekind-Peano.

Apêndice A

A1. O Operador de Abstração

Teorema 1.23. Se A é $V_j v_i$, então $\vdash \forall v_i\big(A\big[[v_i : B]/V_j\big] \leftrightarrow B\big)$; ou seja, $\vdash \forall v_i([v_i : B]v_i \leftrightarrow B)$.

Prova:

1.	$[v_i : B]v_i, \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i \vdash [v_i : B]v_i$	
2.	$[v_i : B]v_i, \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i \vdash \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i$	
3.	$[v_i : B]v_i, \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i \vdash \forall v_i(V_j v_i \leftrightarrow B)$	2, lógica proposicional
4.	$[v_i : B]v_i, \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i \vdash V_j v_i$	2, lógica proposicional
5.	$[v_i : B]v_i, \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i \vdash V_j v_i \leftrightarrow B$	3, $E\forall$
6.	$[v_i : B]v_i, \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i \vdash B$	4, 5, lógica proposicional
7.	$[v_i : B]v_i \vdash \exists V_j(\forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i)$	1, Definição 1.22
8.	$[v_i : B]v_i \vdash B$	6, 7, $E\exists$
9.	$\vdash [v_i : B]v_i \rightarrow B$	8, *TD*

1.	$B, \forall v_i(V_j v_i \leftrightarrow B) \vdash B$	
2.	$B, \forall v_i(V_j v_i \leftrightarrow B) \vdash \forall v_i(V_j v_i \leftrightarrow B)$	
3.	$B, \forall v_i(V_j v_i \leftrightarrow B) \vdash V_j v_i \leftrightarrow B$	2, $E\forall$
4.	$B, \forall v_i(V_j v_i \leftrightarrow B) \vdash V_j v_i$	1,3, lógica proposicional
5.	$B, \forall v_i(V_j v_i \leftrightarrow B) \vdash \forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i$	2, 4, lógica proposicional
6.	$B, \forall v_i(V_j v_i \leftrightarrow B) \vdash \exists V_j(\forall v_i(V_j v_i \leftrightarrow B) \wedge V_j v_i)$	5, $I\exists$
7.	$B, \forall v_i(V_j v_i \leftrightarrow B) \vdash [v_i : B]v_i$	6, Definição 1.22

8.	$B \vdash \exists V_j \forall v_i (V_j v_i \leftrightarrow B)$	$L10$
9.	$B \vdash [v_i : B] v_i$	7, 8, $E\exists$
10.	$\vdash B \rightarrow [v_i : B] v_i$	9, TD

Logo, $\vdash [v_i : B] v_i \leftrightarrow B$ e, portanto, $\vdash \forall v_i ([v_i : B] v_i \leftrightarrow B)$.

■

Teorema 1.25. Sejam $A, B \subseteq \mathrm{F}(\Sigma)$ tais que v_i é livre em B. Se nenhuma das variáveis livres em B é ligada em A, então

$$\forall v_i (V_j v_i \leftrightarrow B) \vdash A \leftrightarrow A\big[[v_i : B]/V_j\big]$$

Prova: Indução sobre as subfórmulas de A.

1. **Base**

   (a) A é uma fórmula atômica e V_j não ocorre em A. Logo, $A\big[[v_i : B]/V_j\big] = A$ e o resultado é obtido imediatamente de $\vdash A \leftrightarrow A$.

   (b) A é uma fórmula atômica e V_j ocorre em A. Logo,

   $$A\big[[v_i : B]/V_j\big] = \exists V_k (\forall v_i (V_k v_i \leftrightarrow B) \wedge A[V_k/V_j])$$

   Assim,

1.	$\forall v_i (V_j v_i \leftrightarrow B), A \vdash A$	
2.	$\forall v_i (V_j v_i \leftrightarrow B), A \vdash \forall v_i (V_j v_i \leftrightarrow B)$	
3.	$\forall v_i (V_j v_i \leftrightarrow B), A \vdash \forall v_i (V_j v_i \leftrightarrow B) \wedge A$	1, 2, lógica proposicional
4.	$\forall v_i (V_j v_i \leftrightarrow B), A \vdash \forall v_i (V_j v_i \leftrightarrow B) \wedge A[V_k/V_j][V_j/V_k]$	3, $A = A[V_k/V_j][V_j/V_k]$
5.	$\forall v_i (V_j v_i \leftrightarrow B), A \vdash \exists V_k (\forall v_i (V_k v_i \leftrightarrow B) \wedge A[V_k/V_j])$	2, $I\exists$
6.	$\forall v_i (V_j v_i \leftrightarrow B), A \vdash A\big[[v_i : B]/V_j\big]$	5, Definição 1.22
7.	$\forall v_i (V_j v_i \leftrightarrow B) \vdash A \rightarrow A\big[[v_i : B]/V_j\big]$	6, TD

   Sejam

   $$\Gamma = \Big\{ \forall v_i (V_j v_i \leftrightarrow B), A\big[[v_i : B]/V_j\big], \forall v_i (V_k v_i \leftrightarrow B) \wedge A[V_k/V_j] \Big\}$$

e

$$\Gamma' = \left\{ \forall v_i(V_j v_i \leftrightarrow B), A\left[[v_i : B]/V_j\right] \right\}$$

1.	$\Gamma \vdash \forall v_i(V_j v_i \leftrightarrow B)$	
2.	$\Gamma \vdash \forall v_i(V_k v_i \leftrightarrow B) \wedge A[V_k/V_j]$	
3.	$\Gamma \vdash \forall v_i(V_k v_i \leftrightarrow B)$	2, lógica proposicional
4.	$\Gamma \vdash A[V_k/V_j]$	2, lógica proposicional
5.	$\Gamma \vdash \forall v_i(V_j v_i \leftrightarrow V_k v_i)$	1, 3, lógica de primeira ordem
6.	$\Gamma \vdash A[V_k/V_j] \rightarrow A[V_k/V_j][V_j/V_k]$	5, *L*9, *MP*
7.	$\Gamma \vdash A[V_k/V_j][V_j/V_k]$	4, 6, *MP*
8.	$\Gamma \vdash A$	7, $A[V_k/V_j][V_j/V_k] = A$
9.	$\Gamma' \vdash A\left[[v_i : B]/V_j\right]$	
10.	$\Gamma' \vdash \exists V_k(\forall v_i(V_k v_i \leftrightarrow B) \wedge A[V_k/V_j])$	9, Definição 1.22
11.	$\Gamma' \vdash A$	8, 10, $E\exists$
12.	$\forall v_i(V_j v_i \leftrightarrow B) \vdash A\left[[v_i : B]/V_j\right] \rightarrow A$	11, *TD*

Portanto, $\forall v_i(V_j v_i \leftrightarrow B) \vdash A \leftrightarrow A\left[[v_i : B]/V_j\right]$.

2. **Passo Indutivo**

(a) A é $(\neg C)$. Logo, $A\left[[v_i : B]/V_j\right] = \left(\neg C\left[[v_i : B]/V_j\right]\right)$. Pela hipótese de indução,

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash C \leftrightarrow C\left[[v_i : B]/V_j\right]$$

Portanto,

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash \neg C \leftrightarrow \neg C\left[[v_i : B]/V_j\right]$$

(b) A é $(C \rightarrow D)$. Logo, $A\left[[v_i : B]/V_j\right] = \left(C\left[[v_i : B]/V_j\right] \rightarrow D\left[[v_i : B]/V_j\right]\right)$. Pela hipótese de indução,

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash C \leftrightarrow C\left[[v_i : B]/V_j\right]$$

e

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash D \leftrightarrow D\left[[v_i : B]/V_j\right]$$

Portanto,

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash (C \to D) \leftrightarrow \big(C\big[[v_i : B]/V_j\big] \to D\big[[v_i : B]/V_j\big]\big)$$

(c) A é $\forall \mathbf{v} C$. Como, por hipótese, nenhuma das variáveis livres em B é ligada em A, $\mathbf{v}$ não é livre em B. Se $\mathbf{v} = V_j$, $A\big[[v_i : B]/V_j\big] = A$ e o resultado é obtido imediatamente de $\vdash A \leftrightarrow A$. Se, por outro lado, $\mathbf{v} \neq V_j$, então $\mathbf{v}$ não é livre em $\forall v_i(V_j v_i \leftrightarrow B)$. Pela hipótese de indução,

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash C \leftrightarrow C\big[[v_i : B]/V_j\big]$$

Aplicando $I\forall$, obtemos

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash \forall \mathbf{v}\big(C \leftrightarrow C\big[[v_i : B]/V_j\big]\big)$$

Portanto,

$$\forall v_i(V_j v_i \leftrightarrow B) \vdash \forall \mathbf{v} C \leftrightarrow \forall \mathbf{v} C\big[[v_i : B]/V_j\big]$$

■

A2. Teorema da Interpretação

A prova do Teorema da Interpretação (Teorema 1.47) requer três lemas preliminares:

Lema A2.1. Sejam Σ e Σ' assinaturas de segunda ordem, K' uma Σ'-teoria e I uma interpretação de $\mathcal{L}(\Sigma)$ em K'. Seja $t \in \mathrm{T}(\Sigma)_\tau$. Se todas as variáveis em t estão entre $\mathbf{v}_1, \ldots, \mathbf{v}_n$ e $\mathbf{v}_i \in \mathcal{V}_{\tau_i}$, para todo $1 \leq i \leq n$, então

$$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} U_\tau(t^I)$$

Prova: Indução sobre o comprimento de t.

1. **Base**

(a) t é $\mathbf{v} \in \mathcal{V}_\tau$. Logo, existe $1 \leq i \leq n$ tal que $\tau = \tau_i$ e $\mathbf{v} = \mathbf{v}_i$. Portanto,

$$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} U_\tau(\mathbf{v})$$

(b) t é $R \in \mathcal{R}_\tau$. Nenhuma variável ocorre em R e $\vdash_{K'} U_\tau(R^I)$ (pela cláusula (ii) da Definição 1.43). Portanto,

$$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} U_\tau(R^I)$$

2. **Passo Indutivo**

t é $ft_1 \ldots t_m$, em que $f \in \mathcal{F}_{\langle \tau_1, \ldots, \tau_m, \tau \rangle}$ e $t_i \in \mathrm{T}(\Sigma)_{\tau_i}$, para todo $1 \leq i \leq m$. Como todas as variáveis em t estão entre $\mathbf{v}_1, \ldots, \mathbf{v}_n$, todas as variáveis em t_i estão entre $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Pela hipótese de indução, temos

$$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} U_{\tau_i}(t_i^I)$$

Pela cláusula (iii) da Definição 1.43, pelo Teorema da Variação Alfabética e pela regra $E\forall$ (aplicada m vezes), obtemos

$$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} U_{\tau_1}(t_1^I) \to \cdots \to U_{\tau_n}(t_m^I) \to U_\tau(f^I t_1^I \ldots t_m^I)$$

e aplicando *MP* m vezes, concluímos

$$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} U_\tau(f^I t_1^I \ldots t_m^I)$$

■

Lema A2.2. Seja I uma interpretação de $\mathcal{L}(\Sigma)$ em K'. Para todo $\tau \in \mathcal{T}_2$, $\vdash_{K'} \exists \mathbf{v}^\tau U_\tau(\mathbf{v}^\tau)$.

Prova:

1. $\tau = 0$. Logo, $U_\tau(\mathbf{v}^\tau)$ é a fórmula $U(v_j)$ e $\vdash_{K'} \exists v_j U(v_j)$ (pela cláusula (i) da Definição 1.43 e pelo Teorema da Variação Alfabética).

2. $\tau = \langle \tau_1, \ldots, \tau_n \rangle$, em que $\tau_i = 0$, para todo $1 \leq i \leq n$. Logo, $U_\tau(\mathbf{v}^\tau)$ é a fórmula

$$\forall v_{j_1} \ldots v_{j_n} (V_k v_{j_1} \ldots v_{j_n} \to U(v_{j_1}) \wedge \cdots \wedge U(v_{j_n}))$$

Pelo axioma $L10$, temos

$$\vdash_{K'} \exists V_k \forall v_{j_1} \ldots v_{j_n} (V_k v_{j_1} \ldots v_{j_n} \leftrightarrow U(v_{j_1}) \wedge \cdots \wedge U(v_{j_n}))$$

e, portanto, $\vdash_{K'} \exists \mathbf{v}^\tau U_\tau(\mathbf{v}^\tau)$.

■

Lema A2.3. Sejam K uma Σ-teoria, K' uma Σ'-teoria e I uma interpretação de K em K'. Se $A \in \mathrm{AL}(\Sigma)$, i.é., se A é um axioma lógico de K, então $\vdash_{K'} A^{(I)}$.

Prova: Indução sobre o número de generalizações sobre instâncias dos esquemas $L1$–$L10$.

1. **Base:** A é uma instância de $L1$–$L10$.

   (a) A é uma instância de $L1$, ou seja, A é $B \to (C \to B)$. Logo, A^I é $B^I \to (C^I \to B^I)$, que é uma instância de $L1$ (em K'): $\vdash_{K'} A^I$. Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_n$ todas as variáveis livres em A, em que $\mathbf{v}_i \in \mathcal{V}_{\tau_i}$, para todo $1 \leq i \leq n$. Portanto,

   $$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} A^I$$

   Aplicando TD n vezes, concluímos $\vdash_{K'} A^{(I)}$.

   (b) A é uma instância de $L2$. Similar ao caso (a).

   (c) A é uma instância de $L3$. Similar ao caso (a).

   (d) A é uma instância de $L4$, ou seja, A é $\forall \mathbf{v} B \to B[t/\mathbf{v}]$, em que $\mathbf{v} \in \mathcal{V}_\tau$, $t \in \mathrm{T}(\Sigma)_\tau$ e $\mathbf{v}$ é substituível por t em B. Logo, A^I é $\forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I) \to B^I[t^I/\mathbf{v}]$. Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_n$ todas as variáveis livres em A, em que $\mathbf{v}_i \in \mathcal{V}_i$, para todo $1 \leq i \leq n$. Portanto,

   $$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_1}(\mathbf{v}_n), \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I) \vdash_{K'} \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I)$$

   Como $\mathbf{v}$ é substituível por t em B, $\mathbf{v}$ é substituível por t^I em B^I. Logo,

   $$(*)\ U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_1}(\mathbf{v}_n), \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I) \vdash_{K'} U_\tau(t^I) \to B^I[t^I/\mathbf{v}]$$

   (por $E\forall$). Como $\mathbf{v}$ é substituível por t em B, todas as variáveis em t estão entre $\mathbf{v}_1, \ldots, \mathbf{v}_n$ e, pelo Lema A2.1, $U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_1}(\mathbf{v}_n) \vdash_{K'} U_\tau(t^I)$. Logo,

   $$(**)\ U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_1}(\mathbf{v}_n), \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I) \vdash_{K'} U_\tau(t^I)$$

   Aplicando MP a $(*)$ e $(**)$, obtemos

   $$U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_1}(\mathbf{v}_n), \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I) \vdash_{K'} B^I[t^I/\mathbf{v}]$$

   e aplicando TD $n+1$ vezes, concluímos $\vdash_{K'} A^{(I)}$.

(e) A é uma instância de $L5$, ou seja, A é $B \to \forall \mathbf{v} B$, em que $\mathbf{v} \in \mathcal{V}_\tau$ e não é livre em B. Logo, A^I é $B^I \to \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I)$. De $B^I, U_\tau(\mathbf{v}) \vdash_{K'} B^I$, obtemos $B^I \vdash_{K'} U_\tau(\mathbf{v}) \to B^I$ (por TD) e, como $\mathbf{v}$ não é livre em B^I, inferimos $B^I \vdash_{K'} \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I)$ (por $I\forall$). Aplicando TD, temos $\vdash_{K'} A^I$ e, portanto, $\vdash_{K'} A^{(I)}$.

(f) A é uma instância de $L6$, ou seja, A é $\forall \mathbf{v}(B \to C) \to (\forall \mathbf{v} B \to \forall \mathbf{v} C)$, em que $\mathbf{v} \in \mathcal{V}_\tau$. Logo, A^I é $\forall \mathbf{v}(U_\tau(\mathbf{v}) \to (B^I \to C^I)) \to (\forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I) \to \forall \mathbf{v}(U_\tau(\mathbf{v}) \to C^I))$. Seja

$$\Gamma = \{\forall \mathbf{v}(U_\tau(\mathbf{v}) \to (B^I \to C^I)), \forall \mathbf{v}(U_\tau(\mathbf{v}) \to B^I), U_\tau(\mathbf{v})\}$$

Logo,

$\Gamma \vdash_{K'} U_\tau(\mathbf{v})$,

$\Gamma \vdash_{K'} U_\tau(\mathbf{v}) \to (B^I \to C^I)$ (por $E\forall$) e

$\Gamma \vdash_{K'} U_\tau(\mathbf{v}) \to B^I$ (por $E\forall$).

Por MP, obtemos $\Gamma \vdash_{K'} C^I$; e, por TD, $\Gamma - \{U_\tau(\mathbf{v})\} \vdash_{K'} U_\tau(\mathbf{v}) \to C^I$. Como $\mathbf{v}$ não é livre em $\Gamma - \{U_\tau(\mathbf{v})\}$, inferimos $\Gamma - \{U_\tau(\mathbf{v})\} \vdash \forall \mathbf{v}(U_\tau(\mathbf{v}) \to C^I)$ (por $I\forall$). Aplicando TD 2 vezes, temos $\vdash_{K'} A^I$ e, portanto, $\vdash_{K'} A^{(I)}$.

(g) A é uma instância de $L7$. Similar ao caso (a).

(h) A é uma instância de $L8$. Similar ao caso (a).

(i) A é uma instância de $L9$, ou seja, A é

$$\forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m}) \to (B \to B[V_l|V_k])$$

Logo, A^I é

$$\forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m})^I \to (B^I \to B^I[V_l|V_k])$$

em que $\forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m})^I$ é a fórmula

$$\forall v_{j_1}(U(v_{j_1}) \to \ldots \forall v_{j_m}(U(v_{j_m}) \to (V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m})) \ldots)$$

Sejam $\mathbf{v}_1, \ldots, \mathbf{v}_n$ todas as variáveis livres em A, em que $\mathbf{v}_i \in \mathcal{V}_{\tau_i}$, para todo $1 \leq i \leq n$. Logo, existem $1 \leq p, q \leq n$ tais que $\mathbf{v}_p = V_k$ e $\mathbf{v}_q = V_l$ (portanto $\tau_p = \tau_q$). Seja $\Gamma = \{U_{\tau_p}(\mathbf{v}_p), U_{\tau_q}(\mathbf{v}_q), \forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m})^I\}$. Logo,

$\Gamma, V_k v_{j_1} \ldots v_{j_m} \vdash_{K'} U_{\tau_p}(\mathbf{v}_p)$, ou seja,

$\Gamma, V_k v_{j_1} \ldots v_{j_m} \vdash_{K'} \forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \rightarrow U(v_{j_1}) \wedge \cdots \wedge U(v_{j_m}))$

Pela regra $E\forall$ (aplicada m vezes), obtemos

$\Gamma, V_k v_{j_1} \ldots v_{j_m} \vdash_{K'} V_k v_{j_1} \ldots v_{j_m} \rightarrow U(v_{j_1}) \wedge \cdots \wedge U(v_{j_m})$

que, por *MP*, implica

$\Gamma, V_k v_{j_1} \ldots v_{j_m} \vdash_{K'} U(v_{j_1}) \wedge \cdots \wedge U(v_{j_m})$

Portanto, $\Gamma, V_k v_{j_1} \ldots v_{j_m} \vdash_{K'} U(v_{j_i})$, para todo $1 \leq i \leq m$.

Como $\forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m})^I \in \Gamma$,

$\Gamma, V_k v_{j_1} \ldots v_{j_m} \vdash_{K'} \forall v_{j_1}(U(v_{j_1}) \rightarrow \ldots \forall v_{j_m}(U(v_{j_m}) \rightarrow (V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m})) \ldots)$

e, por $E\forall$ e *MP* (aplicadas m vezes), obtemos

$$\Gamma, V_k v_{j_1} \ldots v_{j_m} \vdash_{K'} V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m}$$

Portanto, $\Gamma \vdash_{K'} V_k v_{j_1} \ldots v_{j_m} \rightarrow V_l v_{j_1} \ldots v_{j_m}$ (por *MP* e *TD*).

Pelo mesmo raciocínio acima, obtemos também

$$\Gamma \vdash_{K'} V_l v_{j_1} \ldots v_{j_m} \rightarrow V_k v_{j_1} \ldots v_{j_m}$$

Logo, $\Gamma \vdash_{K'} V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m}$ e, portanto,

$$\Gamma \vdash_{K'} \forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \leftrightarrow V_l v_{j_1} \ldots v_{j_m})$$

(aplicando $I\forall$ m vezes). Por *L9* (em K') e *MP*, obtemos $\Gamma \vdash_{K'} B^I \rightarrow B^I[V_l|V_k]$, que, por *TD*, implica

$$U_{\tau_p}(\mathbf{v}_p), U_{\tau_q}(\mathbf{v}_q) \vdash_{K'} A^I$$

Logo, $U_{\tau_1}(\mathbf{v}_1), \ldots, U_{\tau_n}(\mathbf{v}_n) \vdash_{K'} A^I$. Aplicando *TD* n vezes, concluímos $\vdash_{K'} A^{(I)}$.

(j) A é uma instância de $L10$, ou seja, A é $\exists V_k \forall v_{j_1} \ldots \forall v_{j_m}(V_j v_{j_1} \ldots v_{j_m} \leftrightarrow B)$, em que $v_{j_1}, \ldots, v_{j_m}$ são livres e V_k não é livre em B. Logo, A^I é (equivalente a) a fórmula

$$\exists V_k(U_\tau(V_k) \wedge \forall v_{j_1}(U(v_{j_1}) \rightarrow \ldots \forall v_{j_m}(U(v_{j_m}) \rightarrow (V_k v_{j_1} \ldots v_{j_m} \leftrightarrow B^I)) \ldots))$$

em que $\tau = \langle \tau_1, \ldots, \tau_m \rangle$ e $\tau_i = 0$, para todo $1 \leq i \leq m$. Mas A^I é equivalente à seguinte instância de $L10$ (em K'):

$$\exists V_k \forall v_{j_1} \ldots \forall v_{j_m}(V_k v_{j_1} \ldots v_{j_m} \leftrightarrow U(v_{j_1}) \wedge \cdots \wedge U(v_{j_m}) \wedge B^I)$$

Logo, $\vdash_{K'} A^I$ e, portanto, $\vdash_{K'} A^{(I)}$.

2. **Passo Indutivo:** Existem um axioma lógico B de K (i.e., $B \in \mathrm{AL}(\Sigma)$) e uma variável $\mathbf{v} \in \mathcal{V}_\tau$ tais que A é $\forall \mathbf{v} B$. Logo, A^I é $\forall \mathbf{v}(U_\tau(\mathbf{v}) \rightarrow B^I)$. Sejam $\mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_n}$ todas as variáveis livres em A e $\mathbf{v}_{j_1}, \ldots, \mathbf{v}_{j_m}$ todas as variáveis livres em B, em que $\mathbf{v}_{k_i} \in \mathcal{V}_{\tau_{k_i}}$, para todo $1 \leq i \leq n$; e $\mathbf{v}_{j_i} \in \mathcal{V}_{\tau_{j_i}}$, para todo $1 \leq i \leq m$. Portanto, $\{\mathbf{v}_{j_1}, \ldots, \mathbf{v}_{j_m}\} \subseteq \{\mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_n}, \mathbf{v}\}$ e $\mathbf{v} \neq \mathbf{v}_{k_i}$, para todo $1 \leq i \leq n$. Pela hipótese de indução, $\vdash_{K'} B^{(I)}$, ou seja,

$$\vdash_{K'} U_{\tau_{j_1}}(\mathbf{v}_{j_1}) \rightarrow \cdots \rightarrow U_{\tau_{j_m}}(\mathbf{v}_{j_m}) \rightarrow B^I$$

Como $\{\mathbf{v}_{j_1}, \ldots, \mathbf{v}_{j_m}\} \subseteq \{\mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_n}, \mathbf{v}\}$, aplicando *MP* m vezes, obtemos

$$U_{\tau_{k_1}}(\mathbf{v}_{k_1}), \ldots, U_{\tau_{k_n}}(\mathbf{v}_{k_n}), U_\tau(\mathbf{v}) \vdash_{K'} B^I$$

e, por *TD*,

$$U_{\tau_{k_1}}(\mathbf{v}_{k_1}), \ldots, U_{\tau_{k_n}}(\mathbf{v}_{k_n}) \vdash_{K'} U_\tau(\mathbf{v}) \rightarrow B^I$$

Como $\mathbf{v} \neq \mathbf{v}_{k_i}$, para todo $1 \leq i \leq n$, por $I\forall$, inferimos

$$U_{\tau_1}(\mathbf{v}_{k_1}), \ldots, U_{\tau_{k_n}}(\mathbf{v}_{k_n}) \vdash_{K'} \forall \mathbf{v}(U_\tau(\mathbf{v}) \rightarrow B^I)$$

e aplicando *TD* n vezes, concluímos $\vdash_{K'} A^{(I)}$.

■

Teorema 1.47. TEOREMA DA INTERPRETAÇÃO. Sejam K uma Σ-teoria axiomatizável, K' uma Σ'-teoria e I uma interpretação de $\mathcal{L}(\Sigma)$ em K'. Se $\vdash_{K'} A^{(I)}$, para todo axioma próprio A de K, então I é uma interpretação de K em K'.

Prova: Considere o conjunto $S = \{A \in \mathrm{F}(\Sigma) : \vdash_{K'} A^{(I)}\}$. Mostramos que S inclui os axiomas de K e é fechado sob *MP*:

1. Se A é um axioma lógico de K, $\vdash_{K'} A^{(I)}$ (pelo Lema A2.3).

2. Se A é um axioma próprio de K, $\vdash_{K'} A^{(I)}$ (pela hipótese do teorema).

3. Suponha que $A \to B, A \in S$. Logo, $\vdash_{K'} (A \to B)^{(I)}$ e $\vdash_{K'} A^{(I)}$. Sejam $\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_l}$ todas as variáveis livres em $A \to B$, $\mathbf{v}_{j_1}, \ldots, \mathbf{v}_{j_m}$ todas variáveis livres em A e $\mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_n}$ todas as variáveis livres em B, em que $\mathbf{v}_{i_p} \in \mathcal{V}_{\tau_{i_p}}$, para todo $1 \leq p \leq l$; $\mathbf{v}_{j_p} \in \mathcal{V}_{\tau_{j_p}}$, para todo $1 \leq p \leq m$ e $\mathbf{v}_{k_p} \in \mathcal{V}_{\tau_{k_p}}$, para todo $1 \leq p \leq n$. Portanto, $\{\mathbf{v}_{j_1}, \ldots, \mathbf{v}_{j_m}, \mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_n}\} = \{\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_l}\}$ e

$$(*) \ \vdash_{K'} U_{\tau_{i_1}}(\mathbf{v}_{i_1}) \to \cdots \to U_{\tau_{i_l}}(\mathbf{v}_{i_l}) \to (A^I \to B^I)$$

$$(**) \ \vdash_{K'} U_{\tau_{j_1}}(\mathbf{v}_{j_1}) \to \cdots \to U_{\tau_{j_1}}(\mathbf{v}_{j_m}) \to A^I$$

Aplicando *MP* l vezes em $(*)$, obtemos

$$U_{\tau_{i_1}}(\mathbf{v}_{i_1}), \ldots, U_{\tau_{i_l}}(\mathbf{v}_{i_l}) \vdash_{K'} A^I \to B^I$$

e, como $\{\mathbf{v}_{j_1}, \ldots, \mathbf{v}_{j_m}\} \subseteq \{\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_l}\}$, aplicando *MP* m vezes em $(**)$, obtemos

$$U_{\tau_{i_1}}(\mathbf{v}_{i_1}), \ldots, U_{\tau_{i_l}}(\mathbf{v}_{i_l}) \vdash_{K'} A^I$$

Logo,

$$U_{\tau_{i_1}}(\mathbf{v}_{i_1}), \ldots, U_{\tau_{i_l}}(\mathbf{v}_{i_l}) \vdash_{K'} B^I$$

Sejam $\mathbf{v}_{q_1}, \ldots, \mathbf{v}_{q_r}$ todas as variáveis em $\{\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_l}\}$ que não são livres em B^I, em que $\mathbf{v}_{q_p} \in \mathcal{V}_{\tau_{q_p}}$, para todo $1 \leq p \leq r$. Logo, $\{\mathbf{v}_{k_1}, \ldots, \mathbf{v}_{k_n}, \mathbf{v}_{p_1}, \ldots, \mathbf{v}_{p_r}\} = \{\mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_l}\}$ e, portanto,

$$U_{\tau_{q_1}}(\mathbf{v}_{q_1}), \ldots, U_{\tau_{q_r}}(\mathbf{v}_{q_r}), U_{\tau_{k_1}}(\mathbf{v}_{k_1}), \ldots, U_{\tau_{k_n}}(\mathbf{v}_{k_n}) \vdash_{K'} B^I$$

Aplicando *TD n* vezes, obtemos

$$(*)\ U_{\tau_{q_1}}(\mathbf{v}_{q_1}), \ldots, U_{\tau_{q_r}}(\mathbf{v}_{q_r}) \vdash_{K'} B^{(I)}$$

Mas, pelo Lema A2.2, $U_{\tau_{q_1}}(\mathbf{v}_{q_1}), \ldots, U_{\tau_{q_{r-1}}}(\mathbf{v}_{q_{r-1}}) \vdash_{K'} \exists \mathbf{v}_{q_r} U_{\tau_{q_r}}(\mathbf{v}_{q_r})$, de onde inferimos

$$U_{\tau_{q_1}}(\mathbf{v}_{q_1}), \ldots, U_{\tau_{q_{r-1}}}(\mathbf{v}_{q_{r-1}}) \vdash_{K'} B^{(I)}$$

por $(*)$ e $E\exists$. Repetindo esse processo $r-1$ vezes, concluímos $\vdash_{K'} B^{(I)}$.

■

Apêndice B

B1. Demonstração de $Anc1$-12

($Anc1$) $Rxy \to R^*xy$

Demonstração: Suponha Rxy e $Her(F,R) \wedge In(R,F,x)$. Por D_{In}, $\forall x_1(Rxx_1 \to Fx_1)$ e, em particular, $Rxy \to Fy$. Logo, Fy e, portanto, $\forall F(Her(F,R) \wedge In(R,F,x) \to Fy) \Rightarrow R^*xy$.

■

($Anc2$) $Her(F,R) \wedge In(R,F,x) \to \forall y(R^*xy \to Fy)$

Demonstração: Suponha $Her(F,R) \wedge In(F,R,x)$ e R^*xy. Por D_*,

$$\forall F(Her(F,R) \wedge In(R,F,x) \to Fy)$$

e, em particular, $Her(F,R) \wedge In(R,F,x) \to Fy$. Logo, Fy e, portanto, $\forall y(R^*xy \to Fy)$.

■

($Anc3$) $Fx \wedge Her(F,R) \to \forall y(R^*xy \to Fy)$

Demonstração: Suponha $Fx \wedge Her(F,R)$ e Rxx_1. De $Her(F,R)$, Rxx_1 e Fx, obtemos Fx_1. Logo, $\forall x_1(Rxx_1 \to Fx_1) \Rightarrow In(R,F,x)$. Por $Anc2$, inferimos $\forall y(R^*xy \to Fy)$.

■

($Anc4$) $R^*xy \to \exists zRzy$

Demonstração: Instanciando F para $[u : \exists zRzu]$ em $Anc2$, obtemos

$$Her([u : \exists zRzu],R) \wedge In(R,[u : \exists zRzu],x) \to \forall y(R^*xy \to [u : \exists zRzu]y)$$

(i) Para provar $Her([u:\exists zRzu],R)$, suponha Rx_1x_2 e $[u:\exists zRzu]x_1$. De Rx_1x_2, obtemos $\exists zRzx_2 \Rightarrow [u:\exists zRzu]x_2$.

(ii) Para provar $In(R,[u:\exists zRzu],x)$, suponha Rxx_1. Logo, $\exists zRzx_1 \Rightarrow [u:\exists zRzu]x_1$.

De (i) e (ii), inferimos $\forall y(R^*xy \rightarrow [u:\exists zRuz]y) \Rightarrow \forall y(R^*xy \rightarrow \exists zRzy)$.

■

($Anc5$) $R^*xy \wedge R^*yz \rightarrow R^*xz$

Demonstração: Suponha $R^*xy \wedge R^*yz$ e $Her(F,R) \wedge In(R,F,x)$. Por $Anc2$, $\forall y(R^*xy \rightarrow Fy)$ e, portanto, Fy. Por $Her(F,R)$ e pela variação alfabética

$$Fy \wedge Her(F,R) \rightarrow \forall z(R^*yz \rightarrow Fz)$$

de $Anc3$, obtemos $\forall z(R^*yz \rightarrow Fz)$. Logo, Fz e, portanto, $\forall F(Her(F,R) \wedge In(R,F,x) \rightarrow Fz) \Rightarrow R^*xz$.

■

($Anc6$) $R^*xy \wedge Ryz \rightarrow R^*xz$

Demonstração: Suponha $R^*xy \wedge Ryz$. Por $Anc1$, obtemos R^*yz e, por $Anc5$, R^*xz.

■

($Anc7$) $R^{*=}xx$

Demonstração: Imediata da definição $D_{*=}$.

■

($Anc8$) $R^*xy \rightarrow R^{*=}xy$

Demonstração: Imediata da definição $D_{*=}$.

■

($Anc9$) $Fx \wedge Her(F,R) \rightarrow \forall y(R^{*=}xy \rightarrow Fy)$

Demonstração: Suponha $Fx \wedge Her(F,R)$ e $R^{*=}xy$. Por $Anc3$, obtemos $(*)$ $\forall y(R^*xy \rightarrow Fy)$ e, por $D_{*=}$, $R^*xy \vee x = y$. Se R^*xy, então Fy (por $(*)$); e se $x = y$, Fy (pois Fx). Portanto, $\forall y(R^{*=}xy \rightarrow Fy)$.

■

($Anc10$) $R^{*=}xy \wedge Ryz \rightarrow R^*xz$

Demonstração: Suponha $R^{*=}xy \wedge Ryz$. Por $D_{*=}$, $R^*xy \vee x = y$. Se R^*xy, então R^*xz (por $Anc6$). Se, por outro lado, $x = y$, de Ryz, inferimos Rxz e, portanto, R^*xz (por $Anc1$).

■

($Anc11$) $R^{*=}xy \wedge Ryz \rightarrow R^{*=}xz$

Demonstração: Suponha $R^{*=}xy \wedge Ryz$. Por $Anc10$, obtemos R^*xz e, portanto, $R^{*=}xz$.

■

($Anc12$) $R^*xy \rightarrow \exists z(R^{*=}xz \wedge Rzy)$

Demonstração: Instanciando F para $[u : \exists z(R^{*=}xz \wedge Rzu)]$ em $Anc2$, obtemos

$$Her([u : \exists z(R^{*=}xz \wedge Rzu)],R) \wedge In(R,[u : \exists z(R^{*=}xz \wedge Rzu)],x) \rightarrow$$

$$\forall y(R^*xy \rightarrow [u : \exists z(R^{*=}xz \wedge Rzu)]y)$$

(i) Para provar $Her([u : \exists z(R^{*=}xz \wedge Rzu)],R)$, suponha Rx_1x_2 e $[u : \exists z(R^{*=}xz \wedge Rzu)]x_1$. Logo, $\exists z(R^{*=}xz \wedge Rzx_1) \Rightarrow R^{*=}xz \wedge Rzx_1$, de onde obtemos $R^{*=}xx_1$ (por $Anc11$). Portanto, $R^{*=}xx_1 \wedge Rx_1x_2 \Rightarrow \exists z(R^{*=}xz \wedge Rzx_2) \Rightarrow [u : \exists z(R^{*=}xz \wedge Rzu)]x_2$.

(ii) Para provar $In(R,[u : \exists z(R^{*=}xz \wedge Rzu)],x)$, suponha Rxx_1. Por $Anc7$, $R^{*=}xx$. Logo, $R^{*=}xx \wedge Rxx_1 \Rightarrow \exists z(R^{*=}xz \wedge Rzx_1) \Rightarrow [u : \exists z(R^{*=}xz \wedge Rzu)]x_1$.

De (i) e (ii), inferimos $\forall y(R^*xy \rightarrow [u : \exists z(R^{*=}xz \wedge Rzu)]y) \Rightarrow \forall y(R^*xy \rightarrow \exists z(R^{*=}xz \wedge Rzy))$.

■

B2. Demonstração de **ADP**4.1.1

(**ADP**4.1.1) $F \approx G \wedge Fx \wedge Gx \rightarrow F^{-x} \approx G^{-y}$

Demonstração: Suponha $F \approx G \wedge Fx \wedge Gy$. Logo, $\exists R\, F \approx_R G \;\Rightarrow\; F \approx_R G$. Temos dois casos:

1. Rxy. É imediato mostrar $F^{-x} \approx_R G^{-y}$. Logo, $F^{-x} \approx G^{-y}$.

2. $\neg Rxy$. De Fx, Gy e $F \approx_R G$, obtemos

$$\exists w(Rxw \wedge Gw) \;\Rightarrow\; Rxw \wedge Gw$$

e

$$\exists z(Rzy \wedge Fz) \;\Rightarrow\; Rzy \wedge Fz$$

Considere a relação

$$R' := [uu_1 : (u \neq x \wedge u_1 \neq y \wedge Ruu_1) \vee (u = z \wedge u_1 = w)]$$

(em termos conjunstísticos, R' é a relação que resulta de R pela substituição dos pares $\langle x, w\rangle$ e $\langle z, y\rangle$ pelo par $\langle z, w\rangle$). Mostraremos que $F^{-x} \approx_{R'} G^{-y}$:

(i) Para provar $\forall x_1(F^{-x}x_1 \rightarrow \exists y_1(R'x_1y_1 \wedge G^{-y}y_1))$, suponha $F^{-x}x_1$. Logo, $Fx_1 \wedge x_1 \neq x$. Temos dois subcasos:

(a) Rx_1y. De Rx_1y e Rzy, inferimos $x_1 = z$. Logo, $x_1 = z \wedge w = w \;\Rightarrow\; R'x_1w$. Como Rxw e $\neg Rxy$, $w \neq y$. Portanto, $Gw \wedge w \neq y \;\Rightarrow\; G^{-y}w \;\Rightarrow\; R'x_1w \wedge G^{-y}w$. Quantificando sobre w, obtemos

$$\exists y_1(R'x_1y_1 \wedge G^{-y}y_1)$$

(b) $\neg Rx_1y$. De Fx_1, inferimos $\exists y_1(Rx_1y_1 \wedge Gy_1) \;\Rightarrow\; Rx_1y_1 \wedge Gy_1$. Como $\neg Rx_1y$, $y_1 \neq y$ e, portanto, $x_1 \neq x \wedge y_1 \neq y \wedge Rx_1y_1 \;\Rightarrow\; R'x_1y_1$. De Gy_1 e $y_1 \neq y$, obtemos $G^{-y}y_1$. Logo,

$$\exists y_1(R'x_1y_1 \wedge G^{-y}y_1)$$

(ii) Para provar $\forall y_1(G^{-y}y_1 \to \exists x_1(R'x_1y_1 \wedge F^{-x}x_1))$, suponha $G^{-y}y_1$. Logo, $Gy_1 \wedge y_1 \neq y$. Temos novamente dois subcasos:

(a) Rxy_1. De Rxy_1 e Rxw, inferimos $y_1 = w$. Logo, $z = z \wedge y_1 = w \Rightarrow R'zy_1$. Como Rzy e $\neg Rxy$, $z \neq x$. Portanto, $Fz \wedge z \neq x \Rightarrow F^{-x}z \Rightarrow R'zy_1 \wedge F^{-x}z$. Quantificando sobre z, obtemos

$$\exists x_1(R'x_1y_1 \wedge F^{-x}x_1)$$

(b) $\neg Rxy_1$. De Gy_1, inferimos $\exists x_1(Rx_1y_1 \wedge Fx_1) \Rightarrow Rx_1y_1 \wedge Fx_1$. Como $\neg Rxy_1$, $x_1 \neq x$ e, portanto, $x_1 \neq x \wedge y_1 \neq y \wedge Rx_1y_1 \Rightarrow R'x_1y_1$. De Fx_1 e $x_1 \neq x$, obtemos $F^{-x}x_1$. Logo,

$$\exists x_1(Rx_1y_1 \wedge F^{-x}x_1)$$

(iii) Para provar $R'x_1y_1 \wedge R'x_2y_2 \to (x_1 = x_2 \leftrightarrow y_1 = y_2)$, suponha $R'x_1y_1 \wedge R'x_2y_2$. Logo,

$$(x_1 \neq x \wedge y_1 \neq y \wedge Rx_1y_1) \vee (x_1 = z \wedge y_1 = w)$$

e

$$(x_2 \neq x \wedge y_2 \neq y \wedge Rx_2y_2) \vee (x_2 = z \wedge y_2 = w)$$

Temos quatro subcasos:

(a) $(x_1 \neq x \wedge y_1 \neq y \wedge Rx_1y_1)$ e $(x_2 \neq x \wedge y_2 \neq y \wedge Rx_2y_2)$. Se $x_1 = x_2$, então Rx_1y_1 e Rx_1y_2. Logo, $y_1 = y_2$. Se, por outro lado, $y_1 = y_2$, então Rx_1y_1 e Rx_2y_1. Logo, $x_1 = x_2$. Portanto, $x_1 = x_2 \leftrightarrow y_1 = y_2$.

(b) $(x_1 \neq x \wedge y_1 \neq y \wedge Rx_1y_1)$ e $(x_2 = z \wedge y_2 = w)$. Se $x_1 = x_2$, então Rzy_1 e Rzy. Logo, $y_1 = y$. Absurdo. Se, por outro lado, $y_1 = y_2$, então Rx_1w e Rxw. Logo, $x_1 = w$. Absurdo.

(c) $(x_1 = z \wedge y_1 = w)$. Similar ao caso (b).

(d) $(x_1 = z \wedge y_1 = w)$ e $(x_2 = z \wedge y_2 = w)$. De $x_1 = z = x_2$ e $y_1 = w = y_2$, obtemos $x_1 = x_2 \leftrightarrow y_1 = y_2$.

■

Apêndice C

$(Eta_{2.1})$ $[u : u < x] \approx [u : u < y] \to x = y$

Demonstração: Seja

$$G := \left[u_1 : \forall y([u : u < u_1] \approx [u : u < y] \to u_1 = y)\right]$$

Instanciando F para G em $\mathbf{AP_2}3$, obtemos

$$G\mathbf{0} \wedge \forall x(Gx \to Gsx) \to \forall xGx$$

(i) Para provar $G\mathbf{0}$, suponha $[u : u < \mathbf{0}] \approx [u : u < y]$. Logo, $\exists R[u : u < \mathbf{0}] \approx_R [u : u < y] \Rightarrow [u : u < \mathbf{0}] \approx_R [u : u < y]$, e, portanto,

$$\forall w([u : u < y]w \to \exists z(Rzw \wedge [u : u < \mathbf{0}]z))$$

Suponha $\mathbf{0} \neq y$ (para proceder por absurdo). Logo, $\exists w\, w < y \Rightarrow w < y \Rightarrow [u : u < y]w$. Portanto, $\exists z(Rzw \wedge [u : u < \mathbf{0}]z) \Rightarrow [u : u < \mathbf{0}]z \Rightarrow z < \mathbf{0}$. Absurdo.

(ii) Para provar $\forall x(Gx \to Gsx)$, suponha $\left[u_1 : \forall y([u : u < u_1] \approx [u : u < y] \to u_1 = y)\right]x \Rightarrow$ $(*)$ $\forall y([u : u < x] \approx [u : u < y] \to x = y)$. Suponha também $[u : u < sx] \approx [u : u < y]$. Logo, $\exists R[u : u < sx] \approx_R [u : u < y] \Rightarrow [u : u < sx] \approx_R [u : u < y]$ e, portanto,

$$(*)\ \forall z([u : u < sx]z \to \exists w(Rzw \wedge [u : u < z]w))$$

$$(**)\ \forall w([u : u < y]w \to \exists z(Rzw \wedge [u : u < sx]z))$$

De $x < sx$ e $(*)$, inferimos $\exists w(Rxw \wedge [u : u < y]w) \Rightarrow Rxw \wedge [u : u < y]w \Rightarrow w < y \Rightarrow \exists w\, w < y$. Logo, $y \neq \mathbf{0}$ e, portanto, $\exists y_1\, y = sy_1 \Rightarrow y = sy_1 \Rightarrow y_1 < y$. De $(**)$, inferimos então

$$\exists z(Rzy_1 \wedge [u : u < sx]z) \Rightarrow Rzy_1 \wedge [u : u < sx]z$$

Temos dois casos:

1. Rxy_1. É imediato mostrar $[u : u < x] \approx_R [u : u < y_1] \;\Rightarrow\; [u : u < x] \approx [u : u < y_1]$.

2. $\neg Rxy_1$. Considere a relação

$$R' := [uu_1 : (u \neq x \wedge u \neq z \wedge Ruu_1) \vee (u = z \wedge u_1 = w)]$$

(em termos conjuntísticos, R' é a relação que resulta de R pela substituição dos pares $\langle x, w\rangle$ e $\langle z, y_1\rangle$ pelo par $\langle z, w\rangle$). É imediato mostrar $[u : u < x] \approx_{R'} [u : u < y_1] \;\Rightarrow\; [u : u < x] \approx [u : u < y_1]$ (ver a demonstração de **ADP**4.1.1 no Apêndice B2).

Pelos casos 1 e 2, obtemos $[u : u < x] \approx [u : u < y_1]$ e, por $(*)$ acima (instanciando y para y_1), obtemos $x = y_1 \;\Rightarrow\; sx = sy_1 \;\Rightarrow\; sx = y$.

De (i) e (ii), inferimos então $\forall x Gx$, i.é.,

$$\forall x\big[u_1 : \forall y([u : u < u_1] \approx [u : u < y] \rightarrow u_1 = y)\big]x \;\Rightarrow\; \forall x \forall y([u : u < x] \approx [u : u < y] \rightarrow x = y)$$

■

$(Enum_2)\; Enum(R, F) \rightarrow \forall x \forall y(Rxy \rightarrow x \leq y)$

Demonstração: Suponha $Enum(R, F)$, Rxy e $y < x$ (para proceder por absurdo). Portanto, $Rxy \wedge y < x \;\Rightarrow\; [u : \exists x(Rxu \wedge u < x)]y \;\Rightarrow\; \exists y[u : \exists x(Rxu \wedge u < x)]y$. Pelo princípio do menor número, obtemos

$$\exists z([u : \exists x(Rxu \wedge u < x)]z \wedge \forall w(w < z \rightarrow \neg[u : \exists x(Rxu \wedge u < x)]w)) \;\Rightarrow$$

$$(*)\; [u : \exists x(Rxu \wedge u < x)]z \wedge \forall w(w < z \rightarrow \neg[u : \exists x(Rxu \wedge u < x)]w)$$

Assim, $[u : \exists x(Rxu \wedge u < x)]z \;\Rightarrow\; \exists x(Rxz \wedge z < x) \;\Rightarrow\; Rx_1z \wedge z < x_1$. De $Nat(z)$, inferimos $\exists y_1 Rzy_1 \;\Rightarrow\; Rzy_1$ (por D_{Enum}) e, de Rzy_1, Rx_1z e $z < x_1$, inferimos $y_1 < z$ (por D_{Enum} novamente). Logo, $Rzy_1 \wedge y_1 < z \;\Rightarrow\; \exists x(Rxy_1 \wedge y_1 < x) \;\Rightarrow\; [u : \exists x(Rxu \wedge u < x)]y_1$. Mas, como z é o menor número que cai sob $[u : \exists x(Rxu \wedge u < x)]$ e $y_1 < z$, também temos $\neg[u : \exists x(Rxu \wedge u < x)]y_1$ (por $(*)$). Absurdo. Logo, $x \leq y$.

■

Referências Bibliográficas

[1] BELNAP, N. On Rigorous Definitios. *Philosophical Studies: An International Journal for Philosophy in the Analytic 72* (1993), 151–146.

[2] BOOLOS, G. Reading the *Begriffsschrift*. *Mind 94* (1985), 331–344. Reimpresso em [8], pp. 155–170, e em [14], pp. 163–181. Citado em [8].

[3] BOOLOS, G. Saving Frege from Contradiction. *Proceedings of the Aristotelian Society 87* (1986/87), 137–151. Reimpresso em [8], pp. 171–182, e em [14], pp. 438–452. Citado em [8].

[4] BOOLOS, G. The Consistency of Frege's *Foundations of Arithmetic*. In *On Being and Saying: Essays in Honor of Richard Cartwright*, J. J. Thomson, Ed. MIT Press, 1987, pp. 137–151. Reimpresso em [8], pp. 183–201, e em [14], pp. 211–233. Citado em [8].

[5] BOOLOS, G. The Standard of Equality of Numbers. In *Meaning and Method: Essays in Honor of Hilary Putnam*, G. Boolos, Ed. Cambridge University Press, 1990, pp. 261–277. Reimpresso em [8], pp. 202–219, e em [14], pp. 243–286. Citado em [8].

[6] BOOLOS, G. Frege's Theorem and the Peano Postulates. *The Bulletin of Symbolic Logic 1* (1995), 317–326. Reimpresso em [8], pp. 291–300.

[7] BOOLOS, G. On the Proof of Frege's Theorem. In *Benacerraf and his Critics*, A. Morton and S. Stich, Eds. Blackwell, 1996, pp. 143–159. Reimpresso em [8], pp. 275–290. Citado em [8].

[8] BOOLOS, G. *Logic, Logic & Logic*. Harvard University Press, 1998.

[9] BOOLOS, G., AND HECK JR., R. *Die Grundlagen der Arithmetik*, §§82-83. In *The Philosophy of Mathematics Today*, M. Schirn, Ed. Oxford University Press, 1997, pp. 407–428. Reimpresso em [8], pp. 315–338, e em [27], pp. 69–89. Citado em [8].

[10] BURGESS, J. P. Review of Crispin Wright's *Frege's Conception of Numbers as Objects*. *Philosophical Review 93* (1984), 638–640.

[11] CHURCH, A. A Formulation of the Simple Theory of Types. *The Journal of Symbolic Logic 5* (1940), 56–68.

[12] COOK, R. T., Ed. *The Arché Papers on the Mathematics of Abstraction*. Springer, 2007.

[13] DEDEKIND, R. *Was sind und was sollen die Zahlen?* Vieweg, 1888. Traduzido para o inglês por Wooster W. Beman sob o título *The Nature and Meaning of Numbers* em *Essays on the Theory of Numbers*. Open Court, 1901, pp. 21–58.

[14] DEMOPOULOS, W., Ed. *Frege's Philosophy of Mathematics*. Harvard University Press, 1995.

[15] DUMMETT, M. *Frege's Philosophy of Mathematics*. Harvard University Press, 1991.

[16] ENDERTON, H. B. *A Mathematical Introduction to Logic*, 2 ed. Academic Press, 2001.

[17] EPSTEIN, R. L., AND CARNIELLI, W. A. *Computabilidade, Funções Computáveis, Lógica e os Fundamentos da Matemática*. Editora UNESP, 2005.

[18] FREGE, G. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Louis Nebert, 1879. Traduzido para o inglês por S. Bauer Mengelberg sob o título *Begriffsschrift, a formula language, modeled upon that arithmetic, for pure thought* em [38], pp. 1–82.

[19] FREGE, G. *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Wilhelm Koebner, 1884. Traduzido para o inglês por Dale Jacquette sob o título *The Foundations of Arithmetic: A Logical-Mathematical Investigation into the Concept of Number*. Pearson Longman, 2007.

[20] FREGE, G. *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet,* 2 vols. Verlag von H. Pohle, 1893, 1903. Traduzido parcialmente para o inglês por Montgomery Furth sob o título *The Basic Laws of Arithmetic*. University of California Press, 1964.

[21] HALE, B. Reals by Abstraction. *Philosophia Mathematica 8* (2000), 100–123. Reimpresso em [12], pp. 175–196.

[22] HALE, B., AND WRIGHT, C. *The Reason's Proper Study: Essays Twards a Neo-Fregean Philosophy of Mathematics*. Oxford University Press, 2001.

[23] HALE, B., AND WRIGHT, C. Logicism in the Twenty-first Century. In *The Oxford Handbook of Philosophy of Mathematics*, S. Shapiro, Ed. Oxford University Press, 2005, pp. 166–202.

[24] HECK JR., R. The Development of Arithmetic in Frege's *Grundgesetze der Arithmetik*. *Journal of Symbolic Logic 58* (1993), 579–601. Reimpresso em [14], pp. 257–294, e em [27], pp. 40–68.

[25] HECK JR., R. Finitude and Hume's Principle. *Journal of Philosophical Logic* (1997), 589–617. Reimpresso em [12], pp. 61–84, e em [27], pp. 237–266.

[26] HECK JR., R. Frege's Theorem: An Introduction. *The Harvard Review of Philosophy* (1997), 56–73.

[27] HECK JR., R. *Frege's Theorem*. Oxford University Press, 2011.

[28] MENDELSON, E. *Introduction to Mathematical Logic*, 5 ed. Chapman and Hall/CRC, 2009.

[29] PARSONS, C. Frege's Theory of Numbers. In *Philosophy in America*, M. Black, Ed. Cornell University Press, 1965, pp. 180–203. Reimpresso em [14], pp. 182–210, e em [30], pp. 150–175.

[30] PARSONS, C. *Mathematics in Philosophy: Selected Essays*. Cornell University Press, 1983.

[31] PEANO, G. *Arithmetices principia, nova methodo exposita*. Fratres Bocca, 1889. Traduzido para o inglês por Jean van Heijenoort sob o título *The principles of arithmetic, presented by a new method* em [38], pp. 83–97.

[32] RUFFINO, M. Why Frege would not be a neo-Fregean. *Mind 445* (2003), 51–78.

[33] RUSSELL, B. On Denoting. *Mind 14* (1905), 479–493.

[34] SHAPIRO, S. *Foundations without Foundationalism: A Case for Second-Order Logic*. Clarendon Press, 1991.

[35] SHOENFIELD, J. R. *Mathematical Logic*. Addison–Wesley, 1967.

[36] SUPPES, P. *Introduction to Logic*. Van Nostrand, 1957.

[37] VAN BENTHEM, J., AND DOETS, K. Higher-Order Logic. In *Handbook of Philosophical Logic: Elements of Classical Logic* (*vol. I*), D. M. Gabbay and F. Guenthner, Eds. Springer, 2001, pp. 189–243.

[38] VAN HEIJENOORT, J., Ed. *From Frege to Gödel: A Sourcebook in Mathematical Logic, 1879–1931*. Harvard University Press, 1967.

[39] WHITEHEAD, A. N., AND RUSSELL, B. *Principia Mathematica*, vol. I. Cambridge University Press, 1910.

[40] WRIGHT, C. *Frege's Conception of Numbers as Objects*. Aberdeen University Press, 1983.

[41] WRIGHT, C. On the Philosophical Significance of Frege's Theorem. In *Logic, Language and Thought*, R. Heck Jr., Ed. Oxford University Press, 1997, pp. 201–245. Reimpresso em [22], pp. 272–306. Citado em [22].

[42] WRIGHT, C. Is Hume's Principle Analytic. In *The Reason's Proper Study: Essays Twards a Neo-Fregean Philosophy of Mathematics*, C. Wright and B. Hale, Eds. Oxford University Press, 2001, pp. 307–332. Reimpresso em [12], pp. 17–44.

[43] ZALTA, E. N. Frege's Theorem and Foundations for Arithmetic. In *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., fall 2013 ed. The Stanford Encyclopedia of Philosophy, 2013.