O Observador como parte da descrição

João Carlos Prandini

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE
DOUTOR EM
MATEMÁTICA

Área de Concentração: **Lógica Matemática**
Orientador: **Prof. Dr. Jacob Zimbarg Sobrinho**

— São Paulo, Outubro de 1989 —

PREFÁCIO

O princípio do *reducionismo* em Biologia sustenta serem todas as questões dessa ciência formuláveis em termos moleculares. Não parece ser necessário fazer a apologia dessa visão, face seu enorme sucesso e ampla publicidade. Apesar disso, há outras concepções que também obtiveram resultados bons: a *matemática genética*, ie, o estudo da dinâmica das populações no decurso do tempo, não está fundada numa imagem física do gen ou nas bases físicas de duplicação e transmissão da informação genética; de fato, ela se apoia em um punhado de regras formais (remontando aos tempos de Mendel) que especificam a maneira pela qual os gens são selecionados através de uma sequência de gerações, e em uma noção intuitiva de *adequação*—uma noção de conteúdo biológico e não físico; como outro exemplo, citamos a *teoria das redes neuronais* de McCulloch e Pitts, que não parte de uma descrição física de um neurônio ou das bases físicas da excitação e transmissão de potenciais de ação, alicerça-se, antes, em algumas regras que descrevem um *neurônio formal*; essa teoria teve o mérito de revelar relações insuspeitadas entre a biologia do sistema nervoso central e a teoria da computação abstrata, os computadores e a cibernética; além disso, sua flexibilidade permitiu sua aplicação em outros domínios da biologia como, por exemplo, a teoria das redes de controle genético, os problemas de diferenciação e morfogênese e uma explicação do conceito de *auto-reprodução* (ao qual este trabalho também

está ligado) que é básico na Biologia¹.

Assim, a colocação de novos paradigmas é uma das tarefas centrais da Biologia e o presente trabalho ocupa-se com um novo paradigma introduzido no final da década de 60 por H. Maturana e F. Varela (Maturana, por sinal, colaborou com McCulloch e com Pitts no final da década de 50) e por eles denominado de *autopoiese* (poiesis = produção, criação).

Segundo as próprias palavras dos autores:

“An autopoieic system is organized (defined as a unity) as a network of processes of production (transformation and destruction) of components that produces the components that: (1) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (2) constitute it (the machine) as a concrete unity in the space in which they exist by specifying the topological domain of its realization as such network².”

Em outras palavras, um sistema autopoietico é um agregado de unidades produtivas inter-relacionadas de forma que as produções que ocorram em algumas delas dependem dos produtos já produzidos em outras; além disso a integridade de algumas unidades produtivas podem depender de produtos oriundos de outras unidades, dessa maneira, a forma organizativa de um sistema autopoietico é continuamente regenerada pelo seu funcionamento e a autopoiese é identificada com essa

¹As afirmações acima estão expostas de um modo mais completo em *Foundations of Mathematical Biology*, vol 1, 1972, pp xiii-xxi, editado por Robert Rosen.

²F. J. Varela, *Principles of Biological Autonomy*, The North Holland Series in General Systems Research, vol 2, 1979, p 13, Academic Press.

organização. Esta acepção do termo organização é a mesma que a de Maturana e Varela (que abreviaremos doravante por M & V) que é, por sua vez, distinta daquela do termo *estrutura*. Segundo M & V, a estrutura de um sistema autopoietico é o modo pelo qual a organização autopoietica toma lugar no mundo, por exemplo, numa célula a autopoiese é realizada por meio de um complexo espaço-temporal de lipídios, proteínas, carbo-hidratos e ácidos nucleicos; em suma, a autopoiese pode ser entendida como um invariante de uma classe de entes dotados de autonomia.

Ainda, a título de ilustração, citamos o trecho a seguir:

“The consequences of autopoietic organization are paramount:

1. Autopoietic machines are autonomous; that is, they subordinate all changes to the maintenance of their own organization, independently of how profoundly they may otherwise be transformed in the process. Other machines, henceforth called *allopoietic machines*, have as the product of their functioning something different from themselves...
2. Autopoietic machines have individuality; that is, by keeping their organization as an invariant through its continuous production, they actively maintain an identity that is independent and yet makes possible their interaction with an observer...
3. Autopoietic machines are units because, and only because, of their specific autopoietic or-

ganization: their operations specify their own boundaries in the process of self production...

4. Autopoietic machines do not have inputs or outputs. They can be perturbed by independent events and undergo internal structural changes which compensate these perturbations. If the perturbations are repeated, the machine may undergo repeated series of internal changes, which may or may not be identical. Whichever series of internal changes takes place, however, they are always subordinated to the maintenance of the machine organization, a condition which is definitive for the autopoietic machines. Thus if there is a relation between these changes and the course of perturbations to which we may point, it pertains to the domain in which the machine is observed, but not to its organization. Although an autopoietic machine can be treated as an allopoietic machine, this treatment does not reveal its organization as an autopoietic machine. In fact, autopoietic and allopoietic descriptions of a system are complementary pairs, depending on the observer's needs³..."

A concepção da autopoiese como um invariante que caracterizaria a organização dos seres vivos foi aqui utilizada como uma possível descrição do fazer matemático. Assim, procuramos apresentar a atividade do matemático como envolvida na produção de certos produtos que por sua vez tomam parte

³F. J. Varela, *op cit*, pp 15-16.

e sustentam a própria atividade de fazer matemática. Tais processos tem lugar num sistema simbólico e, para que a analogia com as idéias da autopoiese seja maior, serão considerados compartimentalizados em diversas *unidades*. Dessa forma há a *unidade linguística*, responsável pela elaboração das fórmulas e dos conjuntos de hipóteses; há uma *unidade de formação de objetos*, onde são produzidos os objetos internos e os de fronteira, ambos por meio de ε -termos; há a *unidade das deduções*, onde se processam, entre outras coisas, as *provas autopoieticas*, análogas às do cálculo de predicados clássico; bem como uma *unidade semântica*, onde se imbricam a sintaxe e a semântica. Convém, nessa altura, destacar uma característica especial do sistema a ser desenvolvido com o propósito distingui-lo dos sistemas lógicos usuais. Normalmente, quando se introduz uma linguagem formal, fornecem-se certas regras efetivas para a formação de suas fórmulas que acabam constituindo-se num conjunto *recursivo* de expressões da linguagem. Diferentemente do usual, as fórmulas aqui não estão dadas de antemão, sua formação podendo depender de certas deduções que possam vir a ser feitas no sistema. Em outras palavras, o reconhecer se uma dada expressão é ou não uma sentença não depende apenas da expressão em si, mas também dos processos envolvidos na sua formação. Isso porque nossa definição de fórmula pode envolver o aparecimento, em uma dedução, da expressão $\Gamma : \models \exists x(A(x) \wedge Form(x))[f]$; nesse caso passamos a considerar $\varepsilon x(A(x) \wedge Form(x))[f]$ uma fórmula (mais precisamente, uma ε_Γ -fórmula). Nas expressões precedentes, A é uma fórmula (ou mesmo uma ε_Γ -fórmula) cuja única variável livre é x , enquanto que $Form$ é uma fórmula da linguagem de ZF cuja extensão no modelo na-

tural é a do conjunto dos números de Gödel das fórmulas da linguagem de ZF . Em particular, ficará evidente, numa certa altura do desenvolvimento (usando o lema 2.8.1 junto com a regra 2.3.2), que o conjunto das fórmulas da linguagem do sistema tem complexidade computacional no mínimo igual ao do conjunto dos números de Gödel das sentenças da linguagem de ZF da forma $\exists x(A(x) \wedge Form(x))$ dedutíveis a partir dos axiomas de $ZF + V = OD$. Consequentemente, é no mínimo cabível a pergunta de se saber se o conjunto das fórmulas da linguagem do sistema é recursivo, contudo não nos ocuparemos dessa questão nesse trabalho.

Serão apresentadas duas noções de provas. A primeira delas (ver definição 2.7.1) não envolve a idéia de modificações e é estritamente mais fraca que a segunda. A primeira noção é bastante análoga à do cálculo de predicados clássico embora, por dispor de maiores recursos, mais forte. Por exemplo, é possível com a primeira noção de prova deduzir a sentença $Con(ZF + V = OD)$ a partir dos axiomas de $ZF + V = OD$ (embora, para tanto, seja necessário utilizar regras ausentes no cálculo clássico). A segunda noção (ver definição 2.7.4), envolve modificações e “programas”. Um programa é uma coleção finita de expressões envolvendo a semântica e que, em se as supondo válidas, poderiam ser incorporadas como linhas de uma dedução, produzindo novos resultados que não poderiam ser alcançados com a primeira noção de prova somente. Uma modificação é um modo de introduzir-se expressões em programas, partindo-se, por exemplo, no início, de um programa vazio. Essa noção é uma tentativa de mimetizar-se a ação de incorporar-se, no fazer matemático, novas aquisições (heurísticas ou advindas por

qualquer outro processo)⁴ supostamente válidas, resultantes, talvez, de uma apreciação das deduções até então obtidas. As provas com modificações são também uma alternativa para o conceito de *prova dinâmica*, como entendido por D. Perlis em seus trabalhos na área de inteligência artificial⁵. O fato das provas com modificações serem estritamente mais fortes que as sem modificações (ver o teorema 2.10.3) garante que efetivamente ganhamos coisas novas com elas. De fato, sem ela é impossível mostrar que a constante **P** do nosso vocabulário tem as características de um *predicado verdade*. Esse resultado deve ser comparado com os obtidos de alguns anos para cá em que se procuram emprestar a maior generalidade possível a esquemas como $\vdash T([\sigma]) \leftrightarrow \sigma$ sem entrar em conflito com o teorema de Tarski. Há um artigo de S. Kripke⁶, pelo lado semântico, que é muito interessante e onde podemos encontrar uma definição bastante precisa de paradoxo, mas é no artigo de S. Feferman⁷ que encontraremos um apanhado bastante amplo dessa questão, bem como de outras relacionadas, junto com os resultados obtidos pelo autor. Tal discussão será apresentada com maiores detalhes na seção denominada “Predicado Verdade e Completude”.

Junto com **M** & **V**, acreditamos que só há linguagem quando também houver auto-referência. Vamos abaixo des-

⁴A esse respeito, vale a pena dar uma olhada nos artigos de Penelope Maddy, *Believing the axioms I, II*, *The Journal of Symbolic Logic*, vol 53, Number 2, June 1988, e vol 53, Number 3, Sept. 1988.

⁵D. Perlis, *Languages with self-reference I: Foundations*, *Artificial Intelligence*, 25, (1985), pp 301–322, e também *Languages with self-reference II: Knowledge, Belief, and Modality*, *Artificial Intelligence*, 34, (1988), pp 179–212.

⁶S. Kripke, *Outline of a theory of truth*, *Journal of Philosophy*, 1975, vol 72, pp 690–716.

⁷S. Feferman, *Toward useful type-free theories. I*, *The Journal of Symbolic Logic*, vol 49, Number 1, 1984.

crever, de um modo bastante sucinto, como M & V enunciam essa tese que também sustentamos e que é um dos motivos fundamentais do primeiro e último capítulo desse trabalho.

Segundo M & V, o fundamental para a existência da vida é a manutenção da autopoiese. Dessa forma, uma unidade autopoietica, enquanto tal, poderá apenas ter as interações com o meio que não causem a perda da sua organização autopoietica; em caso contrário a unidade será destruída, morta. Assim, existe um domínio de interações que uma dada unidade pode manter sem a perda de sua autopoiese; tal domínio é chamado *domínio cognitivo* da unidade em questão. Contudo nada impede que interações com outras unidades autopoieticas façam parte do domínio cognitivo de uma dada unidade autopoietica “A”, digamos; reciprocamente, “A”, por sua vez, pode ser parte do domínio cognitivo de unidades autopoieticas do domínio cognitivo da própria “A”. Seja “B” uma tal unidade. Pode então ocorrer que a conduta de “A” torne-se uma fonte de deformação para “B”; por outro lado, “B”, em razão da deformação advinda de “A” e com o fito de manter sua autopoiese, pode ter um comportamento compensatório e tornar-se uma fonte de deformação para “A”, num processo recorrente. Essa sequência que acabamos de descrever (“A” fonte de deformação para “B”; “B”, compensando a deformação, torna-se fonte de deformação para “A”; “A” compensa essa deformação e torna-se, novamente, fonte para “B”; etc...) é chamada por M & V de sequência de *interações comunicativas* e a coleção de tais interações forma o que é chamado *domínio linguístico*. Não obstante, é possível falar-se em domínio linguístico sem que exista aquilo que chamamos de linguagem. M & V afirmam mesmo que podem ocorrer

interações comunicativas entre seres que sequer possuam sistema nervoso⁸. Contudo um sistema autopoietico capaz de interagir com seus próprios estados pode tratar seus estados linguísticos como fontes de deformações, isto é, tratar sua linguagem como algo que deva ser por ela descrito. Nas palavras de M & V:

“Lo fundamental en el caso humano, es que el observador ve que *las descripciones pueden ser hechas tratando a otras descripciones como si fueran objetos o elementos del dominio de interacciones*⁹...”

“...Pero para nuestros propósitos podemos identificar la característica clave del lenguaje que modifica de manera tan radical los dominios conductuales humanos haciendo posibles nuevos fenómenos como la reflexión y la conciencia. Esta característica es que el lenguaje permite al que opera en él *describirse a sí mismo* y a su circunstancia¹⁰.”

Quando há a linguagem no sentido exposto acima, o sistema passa, segundo M & V, a operar como um *observador*¹¹.

O primeiro capítulo desse trabalho aborda precisamente este ponto. A definição fundamental é a de *estrutura auto-referente* (ver definição 1.3.2) que passaremos a identificar com o observador.

É perfeitamente concebível que um sistema capaz de ser um observador interaja com descrições linguísticas de si próprio,

⁸H. Maturana e F. Varela, *El árbol del conocimiento*, Editorial Universitaria, 1987, p 129.

⁹H. Maturana e F. Varela, *op. cit.*, p 139.

¹⁰H. Maturana e F. Varela, *op. cit.*, p 139.

¹¹H. Maturana e F. Varela, *Autopoiesis and Cognition*, Boston Studies in the Philosophy of Science, vol 42, 1980, p 121.

além disso, ele pode interagir com descrições de si mesmo entrando em relação com descrições de si próprio, e assim por diante. A isso, M & V chamam de *domínio de auto-observação*. Por meio do teorema da completude (ver o teorema 2.8.3), é possível de obter-se, internamente, uma teoria completa; restringiremos nossos observadores àqueles cuja teoria subjacente seja esta; mesmo assim, para obtermos descrições como acima, é necessário a introdução de níveis superiores, ou seja, para que o sistema possa falar de si mesmo enquanto observador, é preciso falar a partir de níveis superiores, assunto do último capítulo. Nele está o teorema do “observador como parte da descrição” que consubstancia essas afirmações. Quanto a descrições do observador interagindo com descrições de si mesmo, apresentamos no mesmo capítulo final o modo como isso pode ser feito através de uma cadeia de sistemas encaixados; o passo fundamental da passagem do primeiro nível para o segundo é repetido na passagem do segundo para o terceiro, e assim por diante.

A Jacob Zimbarg Sobrinho, de quem provieram as idéias e o estímulo indispensáveis para a concretização deste trabalho, meus agradecimentos nas dimensões que somente um José de terras e eras distantes poderia ter.

Conteúdo

1	Linguagens Auto-Referentes	1
1.1	Introdução	1
1.2	Predicados	2
1.3	Estruturas Auto-Referentes	5
2	Sistemas Autopoiéticos	9
2.1	Introdução	9
2.2	Regras de formação de objetos	11
2.3	Formação de fórmulas e conjuntos de hipóteses	13
2.4	Regras sintáticas de dedução	16
2.5	Regras de semântica	17
2.6	Regras de definibilidade e P regras	19
2.7	Provas Autopoiéticas	22
2.8	Predicado Verdade e Completude	25
2.9	Estruturas Auto-Referentes	39
2.10	Consistência	41
3	Sistemas de Ordem Superior	55
3.1	Introdução	55
3.2	Extensão de mais um nível	56
3.3	Invariância do Finito	60
3.4	Um modelo conveniente com sua interpretação	63

Capítulo 1

Linguagens Auto-Referentes

1.1 Introdução

Conforme o que dissemos na introdução, entenderemos a atividade matemática como sendo uma atividade linguística no sentido lá empregado, ou seja, onde o essencial está no descrever o que descreve. Neste capítulo tornaremos essas idéias mais precisas. Na sua primeira seção introduzimos a noção de *predicado*, que é uma coleção de fórmulas de linguagem; trata-se de um primeiro passo na direção que pretendemos trilhar, o segundo será tomado na segunda seção pela definição 1.3.2, onde passaremos a considerar os próprios predicados como os elementos do domínio de uma estrutura; estaremos assim fazendo as fórmulas (e, em consequência, os predicados, conforme o conceito de *extensão* de um predicado, introduzido logo após a definição 1.2.1) referirem-se a elementos puramente linguísticos.

1.2 Predicados

Começemos com a definição de predicado.

DEFINIÇÃO 1.2.1 *Sejam T uma teoria consistente numa linguagem de primeira ordem $\mathcal{L}$ e P uma coleção de fórmulas com $n + 1$ variáveis livres, $n \geq 0$, que, para efeitos de normalização, serão sempre $v_0, \dots, v_n$. Diremos que P é um **predicado** de T se, e somente se, as condições abaixo estiverem satisfeitas:*

1. *Se $\varphi(v_0, \dots, v_n)$ está em P e se*

$$T \vdash \varphi(v_0, \dots, v_n) \leftrightarrow \psi(v_0, \dots, v_n),$$

então $\psi(v_0, \dots, v_n)$ está em P .

2. *Se $\varphi(v_0, \dots, v_n)$ e $\psi(v_0, \dots, v_n)$ estão em P , então*

$$T \vdash \varphi(v_0, \dots, v_n) \leftrightarrow \psi(v_0, \dots, v_n).$$

*Dá-se a T o nome de **teoria companheira** de P .*

Vale a pena observar a dependência, na definição acima, do predicado P com a teoria T e com a lógica empregada, afinal P depende tanto de T quanto das deduções feitas a partir de T . Para efeitos de simplificação da escrita, referir-nos-emos quase que com exclusividade a predicados unários; será sempre claro no contexto de que não se trata de uma restrição essencial, quando tal ocorrer, avisos convenientes estarão presentes.

Se $\mathbf{M} = \langle D, \in \rangle$ é um modelo de T , P é um predicado unário e $\varphi(v_0)$ está em P , seja $S^\varphi = \{x \in D : \mathbf{M} \models \varphi(v_0)[v_0/x]\}$; caso $\psi(v_0)$ esteja também em P , como $T \vdash$

$\varphi(v_0) \leftrightarrow \psi(v_0)$, temos $M \models \varphi(v_0) \leftrightarrow \psi(v_0)[v_0/x] \forall x \in D$; portanto $S^\varphi = S^\psi$ e faz sentido falar-se sobre S^P que chamaremos de **extensão** do predicado P .

Também é possível aplicar a predicados operações booleanas como $\wedge$ e $\neg$, por exemplo, $P \wedge Q$ é formado por fórmulas equivalentes a $\varphi(v_0) \wedge \psi(v_0)$, onde $\varphi(v_0)$ está em P e $\psi(v_0)$ está em Q .

DEFINIÇÃO 1.2.2 *Dados um predicado P e uma estrutura M diremos que M é adequada para P (ou então que é P -adequada) se, e somente se, para todos $\varphi(v_0)$ e $\psi(v_0)$ de P vale $M \models \forall v_0(\varphi(v_0) \leftrightarrow \psi(v_0))$.*

Naturalmente, se M é um modelo de T , então M é adequada para todos os predicados de T .

Denotemos por $\mathcal{U}$ (de universal) o predicado onde esteja a fórmula $v_0 = v_0$; com essa notação temos o

LEMA 1.2.1 *M é modelo de T se, e somente se, M é $\mathcal{U}$ -adequada.*

PROVA:

Basta provar que se M é $\mathcal{U}$ -adequada então ela é modelo de T .

Seja então S uma sentença dedutível a partir de T ; é fácil ver que $T \vdash v_0 = v_0 \leftrightarrow S \wedge v_0 = v_0$; portanto

$$M \models \forall v_0(v_0 = v_0 \leftrightarrow S \wedge v_0 = v_0),$$

ou seja $M \models S$ e é portanto um modelo de T . $\square$

DEFINIÇÃO 1.2.3 *Uma coleção Σ de predicados duma teoria T será chamada de sistema de predicados. Quando*

toda fórmula está em algum predicado de Σ , diremos que o sistema é **completo**. Quando a estrutura $\mathbf{M}$ é P -adequada para todo o predicado P de Σ , diremos que ela é Σ -adequada.

Assim, levando-se em conta o lema 1.2.1, basta que $\mathcal{U}$ esteja em Σ e que $\mathbf{M}$ seja $\mathcal{U}$ -adequada para que $\mathbf{M}$ seja Σ -adequada.

DEFINIÇÃO 1.2.4 Um sistema completo de predicados Σ é dito **extensional**, se, e somente se, para todo o par P, Q de Σ e para toda estrutura $\mathbf{M}$ que seja Σ -adequada, se $S^P = S^Q$ então $P = Q$.

TEOREMA 1.2.1 Para um sistema completo de predicados Σ duma teoria T são equivalentes:

1. Σ é extensional.
2. T é completa.

PROVA:

Suponha que Σ é extensional, $\mathbf{M}$ é um modelo de T e S é uma sentença verdadeira em $\mathbf{M}$; vamos provar que S é consequência de T .

Seja P o predicado onde está $v_0 = v_0 \wedge S$; daí $S^P = D = S^{\mathcal{U}}$ e assim $P = \mathcal{U}$; logo $T \vdash v_0 = v_0 \leftrightarrow v_0 = v_0 \wedge S$; portanto $T \vdash S$.

Suponha que T é completa, $\varphi(v_0)$ está em P , $\psi(v_0)$ está em Q , $\mathbf{M}$ é uma estrutura Σ -adequada onde valha $S^P = S^Q$; provaremos que $P = Q$.

Como para $\mathbf{M}$, $S^P = S^Q$, vale $\mathbf{M} \models \forall v_0 (\varphi(v_0) \leftrightarrow \psi(v_0))$; como T é completa e $\mathbf{M}$ é modelo de T (pelo lema 1.2.1 e por ser o sistema completo), decorre

$$T \vdash \forall v_0 (\varphi(v_0) \leftrightarrow \psi(v_0));$$

portanto $P = Q$. $\square$

1.3 Estruturas Auto-Referentes

Nesta seção apresentamos os predicados como os objetos sobre os quais o discurso matemático se refere. Além disso, formaremos predicados sobre os próprios predicados (daí o nome auto-referência), ou seja, devemos falar sobre predicados que *satisfazem* predicados. É preciso então que a linguagem que empreguemos para formalizar este processo tenha um símbolo relacional binário para a noção de satisfatibilidade. Ao invés de utilizarmos algum símbolo do tipo $sat(x, y)$, é mais natural empregar o já consagrado $\in$; daí a definição abaixo.

DEFINIÇÃO 1.3.1 *Por linguagem auto-referente entenderemos a de ZF enriquecida, eventualmente, por algumas constantes convenientes.*

A seguir a definição onde os predicados fazem parte do universo de discurso.

DEFINIÇÃO 1.3.2 *Sejam uma linguagem auto-referente $\mathcal{L}$ e uma estrutura $M = \langle D, E, c_j : j \in J \rangle$; M é dita uma estrutura auto-referente se existir um sistema extensional Σ tal que:*

1. $D \subseteq \Sigma$.
2. M é Σ -adequada.
3. $M \models \forall v_0 (v_0 \in v_1 \leftrightarrow \varphi(v_0)) [v_1/P]$, onde φ está em P e P está em D .

O sentido do último item da definição acima, que é uma forma forte do axioma da compreensão, é o de identificar $\in$

com a noção de satisfatibilidade; mais abaixo, no lema 1.3.1, veremos que ela desempenhará um papel essencial nesse sentido. Como já vimos, M é necessariamente um modelo de T , uma teoria completa a partir da qual Σ foi obtida, consequentemente, a forma forte do axioma da compreensão dada pela sentença $\exists v_1 \forall v_0 (v_0 \in v_1 \leftrightarrow \varphi(v_0))$ tem que estar em T toda vez que φ esteja em P , um elemento de D^1 .

LEMA 1.3.1 *Dados um sistema extensional Σ e uma estrutura auto-referente M , vale em M o axioma da extensionalidade.*

PROVA:

Devemos mostrar que

$$M \models \forall v_0 (v_0 \in v_1 \leftrightarrow v_0 \in v_2) \rightarrow v_1 = v_2 [v_1/P; v_2/Q];$$

ou seja, de que se $\forall X \in D, XEP$ se, e somente se, XEQ ; então $P = Q$. Como Σ é extensional e M é Σ -adequada, para se obter $P = Q$ é bastante mostrar que, dados φ de P e ψ de Q arbitrários, vale $M \models \forall v_0 (\varphi(v_0) \leftrightarrow \psi(v_0))$; contudo, pelo item 3 da definição 1.3.2, tem-se XEP se, e somente se, $M \models \varphi(v_0)[v_0/X]$ enquanto que XEQ se, e somente se, $M \models \psi(v_0)[v_0/X]$; assim se vale $\forall X \in D, XEP$ se, e somente se, XEQ ; vale $M \models \forall v_0 (\varphi(v_0) \leftrightarrow \psi(v_0))$. $\square$

Portanto, tendo em vista o último item da definição 1.3.2, podemos afirmar que numa estrutura auto-referente todo seu elemento é definível por meio de uma fórmula; o teorema abaixo é uma espécie de recíproca desse fato.

¹Nem toda fórmula φ está em algum predicado de D , por exemplo, se ela for $v_0 \notin v_0$, o predicado $\mathcal{R}$ que a contenha não poderá estar em D , devido ao conhecido paradoxo de Russell.

TEOREMA 1.3.1 *Seja $M = \langle D, E, c_j : j \in J \rangle$ uma estrutura onde vale o axioma da extensionalidade e onde todo elemento de D seja definível por meio de uma fórmula de $\mathcal{L}$, ie, dado um elemento a de D , há uma fórmula $\varphi(v_0)$ tal que $M \models \exists! v_0 \varphi(v_0)[v_0/a]$; então M é isomorfa a uma estrutura auto-referente.*

PROVA:

Dado $a \in D$, vamos obter um predicado P_a que lhe corresponda. Em primeiro lugar, tomemos como T a teoria completa constituída pelas sentenças verdadeiras em M . Seja agora $\varphi(v_0)$ uma fórmula que define a ; diremos que uma fórmula $\psi(v_0)$ está em P_a se, e somente se, $T \vdash \psi(v_0) \leftrightarrow \exists v_1 (v_0 \in v_1 \wedge \varphi(v_1))$, ou seja, P_a é $P^{\sigma(v_0)}$, onde $\sigma(v_0)$ é $\exists v_1 (v_0 \in v_1 \wedge \varphi(v_1))$. Como T é completa, o sistema Σ que ela define é extensional, de acordo com o teorema 1.2.1, e o conjunto $\Delta = \{P_a : a \in D\}$ é uma parte de Σ , consequentemente o item 1 da definição 1.3.2 está satisfeito. Notemos ainda que $P_a = P_b \implies a = b$, pois $P_a = P_b$ acarreta $M \models \forall v_0 (\exists v_1 (v_0 \in v_1 \wedge \varphi(v_1)) \leftrightarrow \exists v_1 (v_0 \in v_1 \wedge \theta(v_1)))$, onde $\theta(v_0)$ define b ; como vale a extensionalidade, resulta que $a = b$. A seguir coloquemos $P_a E_\Delta P_b$ se, e somente se, $a E b$ e façamos c_j corresponder-se a P_{c_j} ; desse modo fica estabelecido um isomorfismo entre M e $N = \langle \Delta, E_\Delta, P_{c_j} : j \in J \rangle$. Resta mostrar que N é auto-referente. Como M é um modelo de T , também o será N que é portanto Σ -adequada; seja agora $\psi(v_0)$ de P_a de Δ como descrito acima; certamente então

$$N \models \psi(v_0) \leftrightarrow \exists v_1 (v_0 \in v_1 \wedge \varphi(v_1))[v_0/X; v_1/P_a],$$

ou seja, $\forall X \in \Delta, \mathbf{N} \models \psi(v_0)[v_0/X]$ **se, e somente se,** $\forall X \in \Delta, X E_{\Delta} P_a$; **portanto vale** $\mathbf{N} \models \exists v_1 \forall v_0 (v_0 \in v_1 \leftrightarrow \varphi(v_0))[v_1/P_a]$. $\square$

Em decorrência do teorema anterior, chamaremos de **sistema auto-referente** toda estrutura onde valha a extensibilidade e a definibilidade, além disso, se $\mathbf{M}$ é uma tal estrutura e $a \in D$, colocamos

$$P_a = \{\varphi(v_0) : \mathbf{M} \models \forall v_0 (v_0 \in v_1 \leftrightarrow \varphi(v_0))[v_1/a]\}.$$

Capítulo 2

Sistemas Autopoiéticos

2.1 Introdução

Neste capítulo será desenvolvido um sistema formal por nós denominado sistema autopoiético. As seções iniciais cuidam basicamente de sua sintaxe, as seguintes, das noções de dedução, de resultados especiais obtidos no sistema e de sua consistência. Cada seção será normalmente precedida por comentários que esperamos sejam esclarecedores.

A sintaxe do nosso sistema formal utilizará os seguintes átomos, a partir dos quais formaremos as expressões significativas:

- $=, \in$ como símbolos predicativos.
- S (Sucessor) como símbolo funcional de posto 1.
- v_i , para i natural, como variáveis.
- $\vee, \wedge, \rightarrow, \leftrightarrow$ como conectivos.
- $\exists, \forall$ como quantificadores.

- $\neg$ como negação.
- 0 como constante individual.
- $:\vdash$ e $:\models$.
- ε para emprego em ε -termos.
- $(,), [,], \{, \}$.
- **P.**

Uma operação básica da sintaxe será a de recursivamente combinar n componentes $c_1, \dots, c_n$ por meio de $\{ \text{ e } \}$ obtendo-se a componente $\{c_1, \dots, c_n\}$; átomos serão considerados componentes de nível 0.

Podemos, em particular, utilizando-nos da concepção de Kuratowski de par ordenado, referirmo-nos a pares de componentes, a funções entre conjuntos de componentes, empregar enfim na metamatemática o jargão conjuntístico usual, o que faremos doravante sem maiores comentários.

Por uma atribuição de variáveis entenderemos uma componente que seja uma função cujo domínio seja inteiramente constituído por variáveis; normalmente elas serão denotadas por f, g, h etc...

A concatenação de n símbolos S consecutivos seguidos por 0 será substituída por n e chamada de numeral. A abreviação “ n é um numeral” será $Num(n)$.

A seguir apresentaremos as regras sintáticas de nosso sistema. Procuraremos apresentá-las dentro do espírito dos chamados sistemas autopoieticos, ie, inserindo-as em unidades que cumprem papéis específicos e que interagem entre si dentro de uma relação “circular” em que os produtos de uma delas

são necessários e dependem de produtos produzidos numa outra. Por essa razão o leitor não deve se preocupar quando se deparar com definições que pressuponham coisas ainda não definidas: elas serão introduzidas posteriormente e tudo é recursivamente consistente.

2.2 Regras de formação de objetos

Dividiremos os *objetos* em dois grupos: *internos* e *de fronteira*. De um modo grosseiro, os objetos internos correspondem àqueles que se situam no interior da unidade autopoiética, ou ainda, aos elementos do domínio duma estrutura, se quisermos usar da linguagem da teoria dos modelos, enquanto que aqueles de fronteira situam-se na fronteira da unidade, podendo também serem pensados como classes próprias do universo. As regras abaixo, como já tivemos ocasião de mencionar, utilizarão noções definidas posteriormente; passemos, então, aos enunciados.

REGRA 2.2.1 *Sendo n um numeral e Γ um conjunto de hipóteses, n será um objeto interno relativamente a Γ , ou, dizendo de um modo alternativo, um Γ -objeto interno. Abreviaremos o que acabamos de dizer sob a seguinte forma:*

$$Num(n) \ \& \ SH(\Gamma) \implies IO(\Gamma; n).$$

Com as notações acima, não é difícil de se entender o enunciado da próxima regra.

REGRA 2.2.2 $SH(\Gamma) \implies IO(\Gamma; 0).$

A próxima regra garante-nos, em particular, que o ε -termo duma sentença existencial cuja veracidade decorre de Γ é um Γ -objeto interno; mas ela é, na verdade, muito mais geral do que acabamos de dizer, porquanto ela permitirá que o ε -termo possa ser o de uma fórmula associada a uma atribuição de variáveis, estando aqui a noção de fórmula na sua acepção generalizada, ie, permitindo que nela ocorram as ε_Γ -fórmulas. Seu enunciado é como segue:

REGRA 2.2.3 *Se A é uma fórmula, $\Gamma : \models \exists v_m A(v_m)[f]$ e todas as ε -fórmulas que ocorram em A são ε_Γ -fórmulas, então o ε -termo de $\exists v_m A(v_m)[f]$ é um Γ -objeto interno, em símbolos, $IO(\Gamma; \varepsilon v_m A[f])$.*

O conteúdo da regra abaixo é bastante semelhante ao da acima, apenas que, no presente caso, o ε -termo passa a ser um Γ -objeto de fronteira, caso tenhamos a não existência.

REGRA 2.2.4 *Se A é uma fórmula, $\Gamma : \models \neg \exists v_m A(v_m)[f]$ e todas as ε -fórmulas que ocorram em A são ε_Γ -fórmulas, então o ε -termo de $\exists v_m A(v_m)[f]$ é um Γ -objeto de fronteira, em símbolos, $BO(\Gamma; \varepsilon v_m A[f])$.*

É simples o significado das duas regras que se seguem.

REGRA 2.2.5 $SH(\Gamma) \implies IO(\Gamma; \wedge)$ e também para $\vee, \rightarrow, \leftrightarrow, \neg, \forall, \exists, \in, =, S, P$.

REGRA 2.2.6 $SH(\Gamma) \implies IO(\Gamma; v_m)$.

Também encararemos as fórmulas como objetos internos, por isso, a

REGRA 2.2.7 $SH(\Gamma) \ \& \ F_{\mathcal{L}}(A) \ \& \ \text{todas as } \varepsilon\text{-fórmulas que ocorram em } A \text{ são } \varepsilon_{\Gamma}\text{-fórmulas} \implies IO(\Gamma; A).$

A última regra desta secção afirma que objetos internos são fechados para a operação de “formação de conjuntos” por meio de $\{$ e $\}$.

REGRA 2.2.8 *Se $a_0, \dots, a_n$ são Γ -objetos internos, então $\{a_0, \dots, a_n\}$ também é um Γ -objeto interno.*

2.3 Formação de fórmulas e conjuntos de hipóteses

Passemos agora à definição das *fórmulas* do sistema. A primeira das regras desta secção diz-nos que as fórmulas atômicas usuais da linguagem da Teoria dos Conjuntos estão entre elas. Nada há de surpreendente nisso. Contudo outras fórmulas terão um estatuto parecido quando a fórmula A da segunda regra for da linguagem da Teoria dos Conjuntos. Nessa segunda regra, $Form(v_k)$ é uma expressão da linguagem da Teoria dos Conjuntos cuja interpretação *standard* é o predicado satisfeito pelos números de Gödel de fórmulas da linguagem da Teoria dos Conjuntos, obtido através duma gödelização a ser definida na secção 2.6. Tais serão as chamadas ε -fórmulas, ou, mais precisamente, as ε_{Γ} -fórmulas. Embora acima tenhamos tomado A como uma fórmula da linguagem da Teoria dos Conjuntos, a segunda regra abaixo não nos obriga a tanto, podendo A conter em sua composição outras ε -fórmulas. A introdução dessas fórmulas *não standard* permitir-nos-á mostrar, no teorema 2.8.3, que a constante (extralinguística) P

atua como uma espécie de *predicado verdade*, além de permitir que “nomes” de fórmulas possam ser utilizados também na linguagem na construção de fórmulas e não apenas como termos, como é usual. O teorema 2.8.3 também esclarece porque não se pode colocar $\mathbf{P}$ como parte da linguagem envolvida nas deduções sintáticas: contradiríamos o teorema da indefinibilidade da verdade de Tarski. Cumpre notar que a introdução dessas fórmulas *não standard* dá-se simultaneamente com toda uma metamatemática também *não standard*, assim, por exemplo, afirmações para ε -fórmulas do tipo “ σ é uma sentença” deve ser entendida como $Sent(v_0)[v_0/\sigma]$, onde $Sent$ é a aritmetização do predicado dos números de Gödel das sentenças da linguagem da Teoria dos Conjuntos.

A terceira e a quarta regras são de entendimento simples e seguem o padrão habitual de formação de fórmulas. Na quinta, associamos às fórmulas da linguagem atribuições de variáveis, que devem ser pensadas como “fórmulas com suas variáveis livres amarradas a certos objetos internos” e que diremos fazer parte duma linguagem um pouco mais ampliada Λ . A sexta regra diz-nos que o passo zero para formação de hipóteses é pela consideração do conjunto vazio e a última delas ensina como engrossar conjunto de hipóteses.

REGRA 2.3.1 $F_{\mathcal{L}}(\in v_m v_n) \ \& \ F_{\mathcal{L}}(= v_m v_n)$.

REGRA 2.3.2 Se A é uma fórmula, todas as ε -fórmulas de A são ε_{Γ} -fórmulas,

$$\Gamma : \models \exists v_k (A(v_k) \wedge Form(v_k))[f],$$

então

$$\varepsilon v_k (A(v_k) \wedge Form(v_k))[f]$$

2.3. FORMAÇÃO DE FÓRMULAS E CONJUNTOS DE HIPÓTESES 15

é uma fórmula. Neste caso dizemos que

$$\varepsilon v_k(A(v_k) \wedge \text{Form}(v_k))[f]$$

é uma ε_Γ -fórmula; quando não fizermos menção explícita a Γ , diremos que se trata de uma ε -fórmula.

Dizemos que v_m é uma variável livre de

$$\varepsilon v_k(A(v_k) \wedge \text{Form}(v_k))[f]$$

quando

$$\Gamma : \models \text{Livre}(v_0, v_1)[v_0/\varepsilon v_k(A(v_k) \wedge \text{Form}(v_k))[f]; v_1/v_m].$$

Dizemos que v_m não é uma variável livre de

$$\varepsilon v_k(A(v_k) \wedge \text{Form}(v_k))[f]$$

quando

$$\Gamma : \models \neg \text{Livre}(v_0, v_1)[v_0/\varepsilon v_k(A(v_k) \wedge \text{Form}(v_k))[f]; v_1/v_m].$$

REGRA 2.3.3 $F_{\mathcal{L}}(A) \ \& \ F_{\mathcal{L}}(B) \implies F_{\mathcal{L}}(\wedge AB) \ \& \ F_{\mathcal{L}}(\neg A)$, e analogamente para os outros conectivos.

REGRA 2.3.4 $F_{\mathcal{L}}(A) \implies F_{\mathcal{L}}(\forall v_m A)$, e analogamente para $\exists$.

Nesse caso, diremos que v_m não é livre em A .

REGRA 2.3.5 $IO(\Gamma; f) \ \& \ F_{\mathcal{L}}(A) \implies F_{\Lambda}(\Gamma; A[f])$, onde todas as ε -fórmulas de A são ε_Γ -fórmulas.

REGRA 2.3.6 $SH(0)$.

REGRA 2.3.7 $F_{\Lambda}(\Gamma; A[f]) \ \& \ SH(\Gamma) \ \& \ \text{toda variável livre de } A \text{ está no domínio de } f \implies SH(\Gamma \cup \{A[f]\})$.

2.4 Regras sintáticas de dedução

Apresentaremos nesta seção as regras de inferência do sistema. Não há aqui grandes novidades, a não ser em casos que possam envolver ε -fórmulas; por exemplo, na terceira regra abaixo, as fórmulas A e B ou são da linguagem usual da Teoria dos Conjuntos, ou envolvem ε -fórmulas; nesse último caso, tais fórmulas deverão ser na realidade ε_Γ -fórmulas, além disso, as afirmações metamatemáticas deverão ser tomadas em sua acepção generalizada, por exemplo, ainda na terceira regra, " v_m não é livre em A " deve ser entendida como $\Gamma : \vdash \neg \text{Livre}(v_0, v_1)[v_0/A; v_1/v_m]$ quando tal se fizer necessário. Doravante comentários como o que acabamos de fazer ficarão implícitos.

REGRA 2.4.1 *Se A é um axioma da teoria $ZF + V=OD$ ou uma tautologia, então $\vdash A$ (ou, um de uma forma um pouco mais geral, se Γ é um conjunto de hipóteses, $\Gamma : \vdash A$).*

REGRA 2.4.2 *De $\vdash A$, podemos inferir $\vdash \forall x A$ (valem as mesmas considerações feitas na regra anterior).*

REGRA 2.4.3 *$F_{\mathcal{L}}(A) \ \& \ F_{\mathcal{L}}(B) \ \& \ v_m$ não é livre em A , então se $\Gamma : \vdash A \rightarrow B$ vale $\Gamma : \vdash A \rightarrow \forall v_m B$.*

REGRA 2.4.4 *$F_{\mathcal{L}}(A) \ \& \ F_{\mathcal{L}}(B) \ \& \ v_m$ não é livre em A , então se $\Gamma : \vdash B \rightarrow A$ vale $\Gamma : \vdash \exists v_m B \rightarrow A$.*

REGRA 2.4.5 $\Gamma : \vdash A \ \& \ \Gamma : \vdash A \rightarrow B \implies \Gamma : \vdash B$.

REGRA 2.4.6 *Se Γ é um conjunto de sentenças e A está em Γ , então $\Gamma : \vdash A$.*

REGRA 2.4.7 *Se B é uma sentença e $\Gamma \vdash A$, então $\Gamma, B \vdash A$.*

REGRA 2.4.8 *$\Gamma, A \vdash B$ e A é uma sentença $\implies \Gamma \vdash A \rightarrow B$, desde que $A \rightarrow B$ seja uma Γ -fórmula.*

2.5 Regras de semântica

Abordaremos agora as regras da unidade semântica. Nesta seção ocupar-nos-emos, portanto, com expressões do tipo $\Gamma \models A[f]$, que intuitivamente significa que a veracidade da fórmula A , com suas variáveis livres avaliadas por f , decorre da veracidade de todas as fórmulas de Γ . A maioria das regras abaixo tem contrapartida informal bastante evidente.

REGRA 2.5.1 *Se A é uma fórmula; todas as ε -fórmulas que ocorram em A são ε_Γ -fórmulas; v_m não está no dom f ; $\Gamma \models A[f]$ e $IO(\Gamma; a)$ então $\Gamma \models A[v_m/a; f]$.*

REGRA 2.5.2 *Se A é uma fórmula; todas as ε -fórmulas que ocorram em A são ε_Γ -fórmulas; v_m não é livre em A e $\Gamma \models A[v_m/a; f]$ então $\Gamma \models A[f]$.*

A regra abaixo contém uma abreviação, $\Gamma \models a = b$, que não faz propriamente parte da linguagem, mas a qual recorreremos frequentemente por seu caráter simplificador da escrita e facilitador da leitura. O que ela está abreviando é $\Gamma \models x = y[x/a; y/b]$. De um modo semelhante, a abreviação $\Gamma \models a \in b$ substitui a expressão $\Gamma \models x \in y[x/a; y/b]$.

REGRA 2.5.3 *Caso tenhamos $\Gamma \models a = b$, $\Gamma \models A[v_m/a; f]$, então teremos $\Gamma \models A[v_m/b; f]$.*

A próxima regra diz que os ε -termos são os termos “por excelência” para a satisfatibilidade de afirmações existenciais.

REGRA 2.5.4 $\Gamma : \models \exists v_m A[f]$ & todas as ε -fórmulas que ocorram em A são ε_Γ -fórmulas $\implies \Gamma : \models A[v_m/\varepsilon v_m A[f]; f]$.

Não surpreende a regra que se segue, se tivermos em conta o significado do quantificador universal.

REGRA 2.5.5 Se a for um Γ -objeto interno, $\Gamma : \models \forall v_m A[f]$ e todas as ε -fórmulas que ocorram em A são ε_Γ -fórmulas, então $\Gamma : \models A[v_m/a; f]$.

A seguir uma versão semântica do Modus Ponens.

REGRA 2.5.6 Se $f \subseteq g$, $\Gamma : \models A[f]$ e $\Gamma : \models (A \rightarrow B)[g]$, então $\Gamma : \models B[g]$.

REGRA 2.5.7 Se $A[f]$ está em Γ , então $\Gamma : \models A[f]$.

Expandir um conjunto de fórmulas Γ , no caso em a veracidade de $\Gamma : \models A[f]$ está já estabelecida, não pode trazer prejuízos, daí a a regra que se segue.

REGRA 2.5.8 Se $\Gamma : \models B[g]$, $SH(\Gamma \cup \{A[f]\})$, e todas as ε -fórmulas que ocorram em A são ε_Γ -fórmulas, então $\Gamma, A[f] : \models B[g]$.

Finalmente, preencher eventuais variáveis livres de um modo compatível deve manter a validade da versão do “teorema da dedução” incluída na

REGRA 2.5.9 Se $f \cup g \subseteq h$, valer que $\Gamma, A[f] : \models B[g]$ e $F_\Lambda(\Gamma; (A \rightarrow B)[h])$ e, além disso, todas as ε -fórmulas que ocorram em A ou B são ε_Γ -fórmulas, então $\Gamma : \models (A \rightarrow B)[h]$.

2.6 Regras de definibilidade e $\mathbf{P}$ regras

Nesta seção apresentaremos essencialmente dois tipos de regras. Um deles está ligado à semântica de alguns objetos internos e corresponde a uma certa *internalização*, por assim dizer, duma gödelização da sintaxe. Tal gödelização será obtida da seguinte forma: em primeiro lugar, associamos aos átomos da linguagem alguns números naturais de maneira permanente, a seguir, por meio duma regra (a primeira desta seção), diremos quais fórmulas são satisfeitas pelos numerais; utilizando então da correspondência átomo-número natural, será possível exibir as fórmulas satisfeitas pelos átomos; após isso, concatenando os numerais por meio de n -uplas ordenadas (que são definíveis por fórmulas), definem-se as fórmulas satisfeitas por essas sequências de numerais, como é usual, sendo portanto possível definir as fórmulas satisfeitas pelas fórmulas; quanto às ε -fórmulas, estas serão definidas por fórmulas às quais os ε -termos satisfazem, no espírito da regra 2.6.3. Sobre essa última regra, é necessário dizer que o símbolo $\leq_D$ nela presente é uma fórmula da linguagem da teoria dos conjuntos que bem ordena o Universo, consoante a hipótese $V = OD$; o que ela em última análise diz é que o ε -termo definido por uma fórmula A é o menor elemento do Universo que satisfaz A . Outro tipo de regra, por exemplo, as regras 2.6.4, 2.6.5 e 2.6.6, diz respeito à relação existente entre ser dedutível no sistema e estar dentro do predicado $\mathbf{P}$; ou então entre ser verdadeiro e estar dentro do predicado $\mathbf{P}$. Finalmente as regras 2.6.7 e 2.6.8 dizem algo sobre esse mesmo predicado $\mathbf{P}$: toda “fórmula” dedutível classicamente da teoria $ZF + V = OD$ está em $\mathbf{P}$, e que $\mathbf{P}$ é um conjunto dedutivamente fechado de sentenças; es-

tamos com isso também dizendo que $Prov(ZF + V = OD, v_0)$ da regra 2.6.7 significa, na gödelização adotada, que v_0 deve ser avaliado como número de Gödel duma sentença dedutível a partir dos axiomas de $ZF + V = OD$, enquanto que na regra 2.6.8, $Provl(v_1, v_0)$ significa que a avaliação de v_0 é dedutível a partir do conjunto de sentenças avaliado por v_1 . Passemos agora aos enunciados formais.

REGRA 2.6.1 *Sejam $D_0(v_0)$ a fórmula $\forall v_1(\neg(v_1 \in v_0))$ e $D_{S_n}(v_0)$, a fórmula*

$$\exists v_1(D_n(v_1) \wedge \forall v_2(v_2 \in v_0 \leftrightarrow v_2 \in v_1 \vee v_2 = v_1))$$

então $:\models D_n(v_0)[v_0/n]$.

REGRA 2.6.2 *Se $a_0, \dots, a_n$ são Γ -objetos internos) e $b = \{a_0, \dots, a_n\}$, então*

$$\Gamma : \models \forall v_1(v_1 \in v_0 \leftrightarrow v_1 = v_2 \vee \dots \vee v_1 = v_{n+2})[v_0/b; v_2/a_0; \dots; v_{n+2}/a_n].$$

REGRA 2.6.3 *Se $IO(\Gamma; \varepsilon v_m A[f])$ e todas as ε -fórmulas de A são ε_Γ -fórmulas, então*

$$\Gamma : \models (\exists v_m(v_m = v_0 \wedge A(v_m)) \wedge \forall v_m(A(v_m) \rightarrow v_0 \leq_D v_m))[v_0/\varepsilon v_m A[f]; f].$$

REGRA 2.6.4 *Se todas as ε -fórmulas de A são ε_Γ -fórmulas e $\Gamma : \models A \in \mathbf{P}$, então $\Gamma : \models A$.*

REGRA 2.6.5 *Se todas as ε -fórmulas de A são ε_Γ -fórmulas; Γ é um conjunto de sentenças e $\Gamma : \models A \in \mathbf{P}$, então $\Gamma : \vdash A$.*

REGRA 2.6.6 *Se A é uma sentença; todas as ε -fórmulas de A são ε_Γ -fórmulas para $\Gamma = 0$; $\vdash A$, então $:\models A \in \mathbf{P}$.*

REGRA 2.6.7

$$:\models \forall v_0 (Prov(ZF + V = OD, v_0) \rightarrow v_0 \in v_1)[v_1/\mathbf{P}].$$

REGRA 2.6.8 $:\models \forall v_0 (Prov_l(v_1, v_0) \rightarrow v_0 \in v_1)[v_1/\mathbf{P}].$

Para α em $\{\wedge, \neg, \forall, \in, =, v\}$, seja $D_\alpha(v_0)$ a fórmula $D_{[\alpha]}(v_0)$; notações análogas valem para os outros conectivos e quantificadores.

Colocamos $D_{v_m}(v_0)$ como

$$\exists v_1 \exists v_2 (D_v(v_1) \wedge D_m(v_2) \wedge v_0 = \langle v_1, v_2 \rangle).$$

Por definição, $D_{\in v_m v_n}(v_0)$ é

$$\exists v_1 \exists v_2 \exists v_3 (D_\in(v_1) \wedge D_{v_m}(v_2) \wedge D_{v_n}(v_3) \wedge v_0 = \langle v_1, v_2, v_3 \rangle).$$

Para $= v_m v_n$, adotamos uma representação análoga.

Quando A é uma ε -fórmula, digamos

$$\varepsilon v_m (B(v_m) \wedge Form(v_m)),$$

adotamos como $D_A(v_0)$ a fórmula correspondente a $B(v_m) \wedge Form(v_m)$ pela regra 2.6.3, ou seja,

$$\begin{aligned} & \exists v_m (v_m = v_0 \wedge (B(v_m) \wedge Form(v_m))) \wedge \\ & \forall v_m ((B(v_m) \wedge Form(v_m)) \rightarrow v_0 \leq_D v_m). \end{aligned}$$

A $D_{\wedge AB}(v_0)$ fazemos corresponder

$$\exists v_1 \exists v_2 \exists v_3 (D_\wedge(v_1) \wedge D_A(v_2) \wedge D_B(v_3) \wedge v_0 = \langle v_1, v_2, v_3 \rangle).$$

De modo análogo tratamos o caso $\neg$, bem como os outros conectivos e os quantificadores.

REGRA 2.6.9 *Se α é um átomo, variável ou fórmula, então*

$$:\models D_\alpha(v_0)[v_0/\alpha].$$

2.7 Provas Autopoiéticas

Nesta seção apresentaremos dois conceitos de dedução no sistema. O primeiro deles, definido logo abaixo, é análogo ao do cálculo de predicados clássico, não obstante é algo mais forte, tanto assim que é possível derivar-se a partir dos axiomas de $ZF + V = OD$ a sentença $Con(ZF + V = OD)$; contudo sua dedução será apresentada numa instância que julgamos mais apropriada, ver o teorema 3.2.1 e os comentários que o seguem. A segunda noção de prova, com modificações, é apresentada na definição 2.7.4. Ela supõe as definições de “programa” e Π -derivações, onde Π é um programa. A idéia subjacente é de que um programa é um conjunto finito de asserções de caráter semântico, cada uma das quais poderia ser utilizada no curso de uma dedução (também chamada de uma Π -derivação); a definição 2.7.4, por sua vez, diz-nos como podemos ir formando programas, partindo-se do programa vazio. O que essa definição almeja consubstanciar é um processo, ou uma computação, cujas regras, ou cujo programa, pudessem ser enriquecidas no seu transcorrer e a partir dos produtos que venham sendo obtidos; achamos que essa posição é mais próxima do que realmente ocorre no fazer matemático, mas não entraremos em discussão tão complexa nesse trabalho, chamaremos apenas à atenção o teorema 2.10.3, que evidencia o fato de serem as provas com modificações mais ricas do que aquelas que as prescindem.

DEFINIÇÃO 2.7.1 *Por uma prova autopoiética (pura) entendemos uma seqüência de linhas*

$$L_1, \dots, L_n$$

satisfazendo as seguintes condições:

1. *Se L_i é $IO(\Gamma; a)$, então ela é obtida pelo uso de uma das regras de formação de objetos da seção 2.2, sendo necessariamente precedida por linhas constituídas pelos condicionantes da formação de $IO(\Gamma; a)$; por exemplo, se a linha for $IO(\Gamma; n)$, antes de L_i encontraremos uma linha que é $Num(n)$ e uma outra que é $SH(\Gamma)$.*
2. *Se L_i é $F_L(A)$, então ela é obtida pelo uso de uma das regras de formação de fórmulas da seção 2.3, sendo necessariamente precedida por linhas constituídas pelos condicionantes da formação de L_i .*
3. *Se L_i é $F_\Lambda(\Gamma; A[f])$, então ela é obtida pelo uso de uma das regras de formação de fórmulas da seção 2.3, sendo necessariamente precedida por linhas constituídas pelos condicionantes da formação de L_i .*
4. *Se L_i é $SH(\Gamma)$, então ela é obtida pelo uso de uma das regras de formação de conjuntos de hipóteses da seção 2.3, sendo necessariamente precedida por linhas constituídas pelos condicionantes da formação de L_i .*
5. *Se L_i é $\Gamma \vdash A$, então ela é obtida pelo uso de uma das regras sintáticas da seção 2.4, ou então pela regra 2.6.5, sendo necessariamente precedida por linhas constituídas pelos condicionantes da formação de L_i .*
6. *Se L_i é $\Gamma \models A[f]$, então ela é obtida pelo uso de uma das regras semânticas da seção 2.5, ou pelo uso de uma das regras da seção 2.6, sendo necessariamente precedida por linhas constituídas pelos condicionantes da formação de L_i .*

Abaixo, a definição de programa.

DEFINIÇÃO 2.7.2 *Um programa é um conjunto finito de pares da forma $\Gamma/A[f]$ que satisfaz as condições seguintes:*

1. Γ é um conjunto finito da forma $\{B_1[g_1], \dots, B_n[g_n]\}$.
2. Tanto A como $B_1, \dots, B_n$ são fórmulas.
3. Tanto f como $g_1, \dots, g_n$ são atribuições de variáveis.
4. Toda variável livre numa B_i está no domínio de g_i , o mesmo vale para A e f .

Descrevemos, a seguir, o modo pelo qual um programa pode influenciar o curso de uma dedução autopoietica.

DEFINIÇÃO 2.7.3 *Seja Π um programa. Por uma Π -derivação entendemos uma sequência de linhas*

$$L_1, \dots, L_n$$

satisfazendo uma das seguintes condições:

1. A linha obedece a uma das exigências da definição 2.7.1.
2. A linha é $\Gamma : \models A[f]$, está em Π e é precedida por linhas que são $SH(\Gamma)$ e $F_\Lambda(A[f])$.

As noções anteriores tomam parte do conceito mais forte de prova incorporado na próxima

DEFINIÇÃO 2.7.4 *Definimos uma prova autopoietica auto-modificável como uma sequência finita da forma*

$$L_1, D_1, \Pi_1; \dots; L_n, D_n, \Pi_n,$$

onde cada uma das triplas L_i, D_i, Π_i satisfaz uma das seguintes condições:

1. i é 1, L_i é $IO(0)$, D_i é 0 e Π_i é 0.
2. $i > 1$, e L_i satisfaz uma das condições da definição 2.7.1.
3. $i > 1$, L_i é $\Gamma : \models A[f]$, existem duas linhas anteriores que são $SH(\Gamma)$ e $F_\Lambda(\Gamma; A[f])$; e Γ , $A[f]$, D_i e Π_i satisfazem as seguintes condições:
 - (a) Γ é um conjunto finito da forma $\{B_1[g_1], \dots, B_n[g_n]\}$.
Seja Σ o programa $\Pi_{i-1} \cup \{ : \models B_1[g_1], \dots, : \models B_n[g_n] \}$
 - (b) D_i é uma Σ derivação cuja última linha é $: \models A[f]$.
 - (c) Π_i é $\Pi_{i-1} \cup \{ \Gamma : \models A[f] \}$.

Quando não estivermos interessados em enfatizar o tipo de dedução autopoietica em questão, falaremos em *prova autopoietica*.

2.8 Predicado Verdade e Completude

Mostraremos a seguir que a constante **P** possui a propriedade que um *predicado verdade* deve ter e (em consequência) a que uma *teoria completa* tem; mais precisamente, apresentaremos uma dedução dentro do sistema de

$$: \models \sigma \leftrightarrow \sigma \in \mathbf{P},$$

para uma sentença qualquer σ , bem como de

$$: \models \forall x (Sent(x) \rightarrow x \in y \vee \forall z (Neg(x, z) \rightarrow z \in y)) [y/\mathbf{P}],$$

onde $Sent(x)$ significa (na gödelização) que x é número de Gödel de uma sentença, enquanto que $Neg(x, y)$ deve ser interpretado como a relação segundo a qual y é número de Gödel

da negação da sentença cujo número de Gödel é x ; assim estamos dizendo, para uma dada sentença, que ela, ou sua negação, estará em $\mathbf{P}$.

Como se sabe, não é possível dentro de um contexto onde as sentenças tem “nomes”, ie, a cada sentença σ está associado um termo $[\sigma]$ que a nomeia, e onde vale o teorema do ponto fixo (segundo o qual para cada fórmula $\varphi(x)$ é possível obter uma sentença σ tal que a expressão $\sigma \leftrightarrow \varphi([\sigma])$ seja dedutível no sistema), manter-se em toda a generalidade um esquema do tipo $T([\sigma]) \leftrightarrow \sigma$, sem incorrer em uma contradição. Tal é o teorema da indefinibilidade da verdade de Tarski. Mesmo assim, muito esforço vem sido desenvolvido para contornar essa impossibilidade, especialmente nos últimos anos. Não é meu propósito aqui fornecer um histórico detalhado do que vem sendo feito, desejo apenas fazer um apanhado superficial daquelas que me parecem ser as mais importantes linhas de ataque desse problema. Para efeitos da exposição abaixo, podemos supor que nosso contexto é o da aritmética de Peano, numa linguagem de primeira ordem L_0 .

Uma primeira alternativa consiste em introduzir na linguagem em que estávamos trabalhando um novo predicado T junto com o esquema $T([\sigma]) \leftrightarrow \sigma$, para todas as sentenças σ da linguagem original, ie, aquelas que não contenham T . A consistência de uma tal extensão decorre facilmente da consideração do modelo obtido pela interpretação de T , no modelo natural, como o conjunto dos números de Gödel das sentenças verdadeiras na aritmética. Essencialmente, o que apresentamos foi o primeiro degrau da hierarquia de Tarski, e, se não nos incomodarmos em trabalhar com teorias tipadas, essa é uma solução viável. Contudo se quisermos manter a simpli-

cidade, ie, não usar tipos, a proibição de T figurar em σ é altamente indesejável.

Outra tentativa viável passa pela adoção de modelos com semânticas a 3 valores, como as de Łukasiewicz e Kleene. Para se ter uma idéia um pouco mais precisa, vamos reproduzir as tabelas-verdade a partir das quais a semântica é construída. Para os conectivos $\wedge$ e $\vee$, e o símbolo da negação $\neg$, não há diferenças entre as semânticas de Łukasiewicz e Kleene e as tabelas são:

p	q	$p \wedge q$	$p \vee q$
t	t	t	t
t	f	f	t
t	u	u	t
f	t	f	t
f	f	f	f
f	u	f	u
u	t	u	t
u	f	f	u
u	u	u	u

p	$\neg p$
t	f
f	t
u	u

Não obstante, há diferenças entre os condicionais e os bicondicionais. Adotando $\rightarrow$ e $\leftrightarrow$ para o condicional e o bicondicional de Kleene, e $\supset$ e $\equiv$ para o condicional e o bicondicional de Łukasiewicz, respectivamente, obtemos as tabelas abaixo:

p	q	$p \rightarrow q$	$p \leftrightarrow q$	$p \supset q$	$p \equiv q$
t	t	t	t	t	t
t	f	f	f	f	f
t	u	u	u	u	u
f	t	t	f	t	f
f	f	t	t	t	t
f	u	t	u	t	u
u	t	t	u	t	u
u	f	u	u	u	u
u	u	u	u	t	t

A extensão da conjunção simples para a infinitária faz-se de acordo com a seguinte definição:

$$\bigwedge_{i \in I} p_i = \begin{cases} t & \text{se todo } p_i \text{ é } t, \\ f & \text{se algum } p_i \text{ é } f, \\ u & \text{caso contrário.} \end{cases}$$

Se considerarmos então uma estrutura $\mathbf{M} = \langle D, \dots \rangle$, obteremos em $\mathbf{M}$ uma semântica a 3 valores usando as tabelas-verdade acima exibidas e definindo o valor de $\forall x \varphi(x)$ como sendo o de $\bigwedge_{\delta \in D} \varphi(\delta)$, onde δ é o nome do elemento d do domínio D . Naturalmente, o valor-verdade de $\exists x \varphi(x)$ é definido como o de $\neg \forall x \neg \varphi(x)$.

É fundamental, para a obtenção do teorema que logo enunciaremos, que a semântica com que se esteja trabalhando seja *monotônica*, ou seja, colocando-se em $\{t, f, u\}$ a ordem parcial $u \leq t$, $u \leq f$, $t \leq t$, $f \leq f$ e $u \leq u$, se tivermos $p \leq q$ e $r \leq s$ então teremos $p \wedge r \leq q \wedge s$, o mesmo valendo para os outros conectivos e para a negação. Por essa razão, o resultado que vem a seguir não é válido para a semântica de Łukasiewicz.

TEOREMA 2.8.1 *Seja A_0 a aritmética de Peano expressa numa linguagem de primeira ordem L_0 . Denotemos por L a linguagem L_0 enriquecida com um novo predicado T , e por A a teoria A_0 junto com o esquema $T([\sigma]) \leftrightarrow \sigma$, onde σ é uma sentença arbitrária de L (podendo, portanto, incluir T). Então há um modelo a 3 valores de A , ie, o valor semântico de $T([\sigma])$ é o mesmo que o de σ .*

O resultado que acabamos de enunciar é um caso bastante particular do chamado FIXED-MODEL THEOREM, cujo enunciado completo pode ser encontrado na página 91 do artigo de Feferman, **Toward useful type-free theories. I**, The Journal of Symbolic Logic, Vol 49, Number 1, 1984.

Há, porém, aspectos muito criticáveis no teorema acima. Para iniciar, diferentemente do que ocorre com a lógica de Łukasiewicz, não é simples axiomatizar-se as fórmulas K-válidas, ie, aquelas que em qualquer estrutura possuam o valor-verdade t. Wang e Scott forneceram uma axiomatização, às custas, contudo, da introdução dos novos símbolos auxiliares $\vdash$ e $\vdash\!\!\!\vdash$. O símbolo $\vdash$ é utilizado para formar expressões do tipo $\varphi \vdash \psi$ (não estamos empregando $\vdash$ no sentido usual, segundo o qual ψ decorre sintaticamente da premissa φ) que serão entendidas como K-válidas quando o valor-verdade de ψ for t sempre que o de φ for t. O símbolo $\vdash\!\!\!\vdash$, por sua vez, abrevia $\varphi \vdash \psi$ e $\psi \vdash \varphi$. Contudo nem $\vdash$, nem $\vdash\!\!\!\vdash$, são realmente conectivos, quer dizer, eles não podem aparecer nas fórmulas φ ou ψ . Com essas restrições, é possível, no entanto, deduzir o esquema $T([\sigma]) \vdash\!\!\!\vdash \sigma$. Mas, mesmo assim, temos impecilhos como o fato da sentença $\sigma \vee \neg \sigma$ deixar de ser válida, além da impossibilidade, comentada acima, de não

se poder empregar $\vdash$ e $\vdash\!\!\!\dashv$ como conectivos.

Uma outra rota foi tentada por Feferman e Aczel, que introduziram um novo conectivo $\equiv$ em adição aos clássicos (incluindo o bicondicional $\leftrightarrow$) junto com vários axiomas, segundo os quais $\equiv$ assume muitas das características de um bicondicional, por exemplo, $\equiv$ é uma relação de equivalência, bem como são válidas as seguintes equivalências:

$$\begin{aligned}(\neg\varphi) &\equiv (0 = 0) \leftrightarrow \varphi \equiv (0 \neq 0) \\(\neg\varphi) &\equiv (0 \neq 0) \leftrightarrow \varphi \equiv (0 = 0) \\ \forall x\varphi(x) &\equiv (0 = 0) \leftrightarrow \forall x(\varphi(x) \equiv (0 = 0)),\end{aligned}$$

apenas para citar alguns exemplos.

Nesse sistema, é possível provar a consistência de

$$T([\sigma]) \equiv \sigma.$$

A vantagem dessa abordagem é que permanecemos num modelo a 2 valores, sem prejuízo de nenhuma regra empregada no cálculo dos predicados clássico, apenas com a ressalva evidente que devemos nos restringir aos axiomas que regem $\equiv$ quando do seu emprego. Mas, como sói acontecer, nem tudo são maravilhas nesse sistema. Por exemplo, não é válida a regra de inferência “natural” $\varphi, \varphi \equiv \psi / \psi$, além disso, a própria interpretação de $\equiv$ é problemática. Feferman mesmo coloca a questão de como interpretar $\equiv$, reconhecendo que “our intuitions are not firm in this respect” (página 99 do artigo de Feferman supracitado), ao comentar sobre uma possível leitura de $\equiv$.

Ainda outra alternativa está no emprego de não apenas um, mas de dois predicados-verdade complementares T e $\bar{T}$, cuja intensão é a de formalizar os esquemas $T([\sigma]) \leftrightarrow \sigma$ e

$\bar{T}([\sigma]) \leftrightarrow \neg\sigma$. Assumem-se os esquemas

$$\begin{aligned} &\neg(T(x) \wedge \bar{T}(x)) \\ &T([\varphi]) \leftrightarrow \varphi^+ \\ &\bar{T}([\varphi]) \leftrightarrow \varphi^-, \end{aligned}$$

onde φ^+ e φ^- são “aproximações” para φ e $\neg\varphi$, respectivamente, e definidas indutivamente da seguinte maneira: (i) φ^+ é φ , se φ é atômica; (ii) se φ é de L_0 (tanto T como $\bar{T}$ não são de L_0) e é atômica, então φ^- é $\neg\varphi$, $(T)^-$ é $\bar{T}$ e $(\bar{T})^-$ é T ; (iii) $(\neg\varphi)^+$ é φ^- e $(\neg\varphi)^-$ é φ^+ ; (iv) $(\varphi \wedge \psi)^+$ é $\varphi^+ \wedge \psi^+$ e $(\varphi \wedge \psi)^-$ é $\varphi^- \wedge \psi^-$; (v) $(\varphi \vee \psi)^+$ é $\varphi^+ \vee \psi^+$ e $(\varphi \vee \psi)^-$ é $\varphi^- \vee \psi^-$; (vi) $(\forall x\varphi(x))^+$ é $\forall x\varphi^+$ e $(\forall x\varphi(x))^-$ é $\exists x\varphi^-$; (vii) $(\exists x\varphi(x))^+$ é $\exists x\varphi^+$ e $(\exists x\varphi(x))^-$ é $\forall x\varphi^-$.

É possível provar que a extensão de A_0 e de L_0 com os novos símbolos e esquemas produz uma extensão conservativa de A_0 . A forma análoga desses resultados para o axioma generalizado da compreensão (de Frege), ie, para

$$\begin{aligned} x \in \{y | \varphi(y)\} &\leftrightarrow \varphi^+(x) \\ x \notin \{y | \varphi(y)\} &\leftrightarrow \varphi^-(x), \end{aligned}$$

já havia sido obtida anteriormente por Gilmore, veja **The consistency of partial set theory without extensionality**, *Axiomatic Set Theory, Proceedings of Symposia in Pure Mathematics*, 1974, Vol 13, Part II, AMS, Providence, RI, pp 147—153. D. Perlis também publicou um resultado análogo em seu artigo **Languages with self-reference I: Foundations**, *Artificial Intelligence*, 25, 1985, pp 301–322.

Os resultados que aqui apresentaremos foram obtidos por uma outra vertente. Partiu-se de algumas propriedades que uma certa constante **P**, que está ligada com as provas do sistema (as regras da seção 2.6), tem e, se utilizarmos o conceito

apropriado de dedução (veja a seção 2.7 e o teorema 2.10.3), consegue-se provar que $\mathbf{P}$ é um predicado verdade, ver o teorema 2.8.2. Vamos agora entrar nos resultados técnicos necessários.

O primeiro resultado abaixo é evidente levando-se em conta que as regras do sistema contém todas as do cálculo de primeira ordem e que nele são dedutíveis os axiomas da teoria $ZF + V = OD$.

LEMA 2.8.1 *Se S é um teorema de $ZF + V = OD$; então $\vdash S$ tem prova autopoietica.*

O lema abaixo era de se esperar, será aplicado futuramente e serve como ilustração de uma prova autopoietica simples.

LEMA 2.8.2 *Sejam A e B fórmulas, f e g atribuições tais que $f|_{\text{dom } f \cap \text{dom } g} = g|_{\text{dom } f \cap \text{dom } g}$. Se $\models A[f]$ e $\models B[g]$, então $\models (A \wedge B)[f \cup g]$.*

PROVA:

1. $\vdash (A \rightarrow (B \rightarrow (A \wedge B)))$ (Regra 2.4.1).
2. $\vdash \forall \vec{x} (A \rightarrow (B \rightarrow (A \wedge B)))$ (Regra 2.4.2).
3. $\models \forall \vec{x} (A \rightarrow (B \rightarrow (A \wedge B))) \in \mathbf{P}$ (Regra 2.6.6).
4. $\models \forall \vec{x} (A \rightarrow (B \rightarrow (A \wedge B)))$ (Regra 2.6.4).
5. $\models (A \rightarrow (B \rightarrow (A \wedge B)))[f \cup g]$ (Regra 2.5.5).
6. $\models (B \rightarrow (A \wedge B))[f \cup g]$ (Regra 2.5.6).
7. $\models (A \wedge B)[f \cup g]. \square$

O resultado a seguir é a metade fácil do teorema 2.8.2, onde se mostra que $\mathbf{P}$ tem o caráter de predicado verdade.

LEMA 2.8.3 $:\models (x \in y \rightarrow \varphi)[x/\varphi; y/P]$.

PROVA:

Da regra 2.5.7 com $\Gamma = \{\varphi \in P\}$ obtemos

$$\varphi \in P : \models \varphi \in P.$$

Utilizando agora a regra 2.6.4, podemos escrever

$$\varphi \in P : \models \varphi.$$

Aplicando nesse ponto a regra 2.5.9, chegamos a

$$:\models \varphi \in P \rightarrow \varphi. \square$$

Na direção do teorema 2.8.2, é útil o lema cujo enunciado segue abaixo:

LEMA 2.8.4 $\{\varphi; \neg\varphi\} : \models x \in y[x/\perp; y/P]$.

PROVA:

Da tautologia $\varphi \rightarrow (\neg\varphi \rightarrow x \in y)$, obtemos

$$:\models \varphi \rightarrow (\neg\varphi \rightarrow x \in y)[x/\perp; y/P].$$

Utilizando nesse ponto a regra 2.5.8, podemos escrever

$$\varphi : \models \varphi \rightarrow (\neg\varphi \rightarrow \perp \in P).$$

Por Modus Ponens, ie, regra 2.5.6, temos

$$\varphi : \models \neg\varphi \rightarrow \perp \in P.$$

Pela regra 2.5.8,

$$\varphi, \neg\varphi : \models \neg\varphi \rightarrow \perp \in P.$$

Novamente por Modus Ponens,

$$\varphi, \neg\varphi : \models \perp \in P,$$

como queríamos. $\square$

Convém observar que poderíamos ter mostrado o resultado mais geral $\{\varphi; \neg\varphi\} : \models S \in \mathbf{P}$, onde S é uma sentença qualquer.

Agora, apresentamos a metade que falta para a demonstração do teorema 2.8.2. É de se notar que a demonstração que virá a seguir possui auto-modificação.

LEMA 2.8.5 $: \models (\varphi \rightarrow x \in y)[x/\varphi; y/\mathbf{P}]$.

PROVA:

Da linha $: \models \varphi$ e do lema anterior obtemos

$$\neg\varphi : \models (\varphi \rightarrow x \in y)[x/\perp; y/\mathbf{P}]$$

e

$$\neg\varphi : \models \varphi,$$

empregando-se no último caso a regra 2.5.8.

Consequentemente, $\neg\varphi : \models x \in y[x/\perp; y/\mathbf{P}]$.

Utilizando a regra 2.6.5, obtemos $\neg\varphi : \vdash \perp$.

Ou ainda, usando a regra 2.4.8, $: \vdash \neg\varphi \rightarrow \perp$, que por sua vez é logicamente equivalente a $: \vdash \varphi$.

Logo, aplicando a regra 2.6.6, obtemos

$$: \models x \in y[x/\varphi; y/\mathbf{P}].$$

Assim $: \models (\varphi \rightarrow x \in y)[x/\varphi; y/\mathbf{P}]. \square$

A partir dos resultados anteriores, podemos então enunciar o

TEOREMA 2.8.2 *Pode-se provar no sistema, para uma sentença qualquer φ que*

$$: \models \varphi \leftrightarrow \varphi \in \mathbf{P}$$

(Chamaremos este resultado de TEOREMA DO PREDICADO VERDADE)

No que se segue, vamos apresentar os resultados envolvidos na prova do teorema 2.8.3, onde se exhibe o caráter de teoria completa da constante **P**.

LEMA 2.8.6 *A expressão $:\models x \in y \vee z \in y[x/\varphi; y/\mathbf{P}; z/\neg\varphi]$, tem prova autopoiética.*

PROVA:

Usando a tautologia

$$(\varphi \leftrightarrow x \in y) \rightarrow ((\neg\varphi \leftrightarrow z \in y) \rightarrow x \in y \vee z \in y)$$

e repetindo os passos 1–5 do lema 2.8.2, obtemos

$$\begin{aligned} &:\models (\varphi \leftrightarrow x \in y) \rightarrow \\ &((\neg\varphi \leftrightarrow z \in y) \rightarrow x \in y \vee z \in y)[x/\varphi; y/\mathbf{P}; z/\neg\varphi] \end{aligned}$$

Utilizando os lemas 2.8.3 e 2.8.5, obtemos

$$:\models x \in y \vee z \in y)[x/\varphi; y/\mathbf{P}; z/\neg\varphi],$$

como desejado. $\square$

Nota: Na prova que apresentaremos do teorema 2.8.3, utilizar-nos-emos do lema precedente. Contudo, no contexto de sua prova, a fórmula φ empregada será uma ε -fórmula, ou seja, é necessário que possamos formar tautologias com as ε -fórmulas, o que é o mesmo que dizer que devemos incorporá-las no conjunto de fórmulas da sintaxe. Não sabemos de nenhuma maneira de provar o teorema 2.8.3 sem o emprego das ε -fórmulas.

LEMA 2.8.7 *A expressão*

$$\begin{aligned} & \exists x(Sent(x) \wedge x \notin y \wedge \\ & \exists z(Neg(x, z) \wedge z \notin y))[y/\mathbf{P}] : \models (x \notin y \wedge \\ & z \notin y)[x/\varphi; y/\mathbf{P}; z/\neg\varphi], \end{aligned}$$

onde φ é o ε -termo de

$$\exists x(Sent(x) \wedge x \notin y \wedge \exists z(Neg(x, z) \wedge z \notin y))[y/\mathbf{P}],$$

tem dedução autopoietica auto modificável.

PROVA:

Vamos admitir como hipótese

$$: \models \exists x(Sent(x) \wedge x \notin y \wedge \exists z(Neg(x, z) \wedge z \notin y))[y/\mathbf{P}].$$

De acordo com a regra 2.5.4, podemos escrever

$$: \models (Sent(x) \wedge x \notin y \wedge \exists z(Neg(x, z) \wedge z \notin y))[y/\mathbf{P}; x/\varphi],$$

não é difícil então concluir

$$: \models Sent(x)[x/\varphi],$$

$$: \models x \notin y[x/\varphi; y/\mathbf{P}]$$

e

$$: \models \exists z(Neg(x, z) \wedge z \notin y)[x/\varphi; y/\mathbf{P}].$$

A última expressão habilita-nos a escrever

$$: \models Neg(x, z) \wedge z \notin y[x/\varphi; y/\mathbf{P}; z/\psi],$$

onde ψ é o ε -termo da fórmula da penúltima expressão acima.

Portanto, usando a regra 2.5.2, obtemos

$$:\models \text{Neg}(x, z)[x/\varphi; z/\psi].$$

Porém, como $:\vdash \text{Neg}(x, z) \wedge D_{\varphi}(x) \rightarrow D_{\neg\varphi}(z)$, obtemos

$$:\models (\text{Neg}(x, z) \wedge D_{\varphi}(x) \rightarrow D_{\neg\varphi}(z))[x/\varphi; z/\psi].$$

Utilizando $:\models \text{Neg}(x, z)[x/\varphi; z/\psi]$ e a regra 2.6.9 aplicada a φ , obtemos

$$:\models D_{\neg\varphi}(z)[z/\neg\psi].$$

Da mesma forma, levando-se em conta que

$$:\vdash D_{\neg\varphi}(z) \wedge D_{\neg\varphi}(w) \rightarrow z = w,$$

obtemos

$$:\models z = w[z/\psi; w/\neg\varphi].$$

Aplicando a regra 2.5.3 à $:\models \text{Neg}(x, z)[x/\varphi; z/\psi]$, concluimos

$$:\models \text{Neg}(x, z)[x/\varphi; z/\neg\varphi].$$

Procedendo de modo análogo com a fórmula $z \notin y$ de $:\models \text{Neg}(x, z) \wedge z \notin y[x/\varphi; y/\mathbf{P}; z/\psi]$, podemos escrever

$$:\models z \notin y[z/\neg\varphi; y/\mathbf{P}].$$

Aplicando o lema 2.8.2 às linhas 12 e 4, concluimos

$$:\models x \notin y \wedge z \notin y[x/\varphi; y/\mathbf{P}; z/\neg\varphi].$$

Assim, de acordo com a definição 2.7.4, obtemos o resultado. $\square$

Já possuímos por agora os instrumentos necessários para a demonstração do

TEOREMA 2.8.3 (da Completude) *A expressão*

$$:\models \forall x(Sent(x) \rightarrow x \in y \vee \forall z(Neg(x, z) \rightarrow z \in y))[y/P]$$

tem prova autopoietica.

PROVA:

Como $x \notin y \wedge z \notin y \rightarrow (x \in y \vee z \in y \rightarrow u \in y)$ é uma tautologia, usando os lemas 2.8.7 e 2.8.6, obtemos

$$\begin{aligned} & \exists x(Sent(x) \wedge x \notin y \wedge \\ & \exists z(Neg(x, z) \wedge z \notin y))[y/P] : \models u \in y[y/P; u/\perp] \end{aligned}$$

ou ainda,

$$\exists x(Sent(x) \wedge x \notin y \wedge \exists z(Neg(x, z) \wedge z \notin y))[y/P] : \models \perp \in P.$$

Contudo os lemas 2.8.3, 2.8.4 e 2.8.5 podem ser adaptados para o caso em que φ é uma sentença da linguagem de ZF com as mesmas demonstrações; por isso, $:\models \perp \in P \rightarrow \perp$ tem prova autopoietica; dessa forma

$$:\models \exists x(Sent(x) \wedge x \notin y \wedge \exists z(Neg(x, z) \wedge z \notin y)) \rightarrow \perp [y/P]$$

tem prova autopoietica, donde decorre

$$:\models \forall x(Sent(x) \rightarrow x \in y \vee \forall z(Neg(x, z) \rightarrow z \in y))[y/P],$$

como queríamos. $\square$

2.9 Estruturas Auto-Referentes

Definindo-se $T(x)$ como $Sent(x) \wedge x \in \mathbf{P}$, obtemos pelo teorema 2.8.3—internamente—uma teoria completa $T(x)$. A partir dela podemos definir um sistema completo de predicados Σ ; basta repetir internamente, por meio da gödelização adotada, as definições correspondentes; por exemplo, para definirmos “ P é um predicado k -ário”, lançamos mão da fórmula abaixo (onde $Eq(v_0, v_1, v_2)$ é uma fórmula de ZF cuja interpretação standard é “ v_2 é o número de Gödel da equivalência entre v_0 e v_1 ”, e $Form(k, \varphi)$ tem por interpretação “ φ é uma fórmula com k variáveis livres”)

$$\begin{aligned} & \exists \varphi (\varphi \in P) \wedge \forall \varphi (\varphi \in P \rightarrow Form(k, \varphi)) \wedge \\ & \forall \varphi \forall \psi \forall \theta (Eq(\varphi, \psi, \theta) \wedge T(\theta) \wedge \varphi \in P \rightarrow \psi \in P) \wedge \\ & \forall \varphi \forall \psi \forall \theta (Eq(\varphi, \psi, \theta) \wedge \varphi \in P \wedge \psi \in P \rightarrow T(\theta)); \end{aligned}$$

seja $Pred(k, P)$ a fórmula apenas descrita, então Σ pode ser definido como $P \in \Sigma \leftrightarrow \exists k (k \in \omega \wedge Pred(k, P))$. Assim, através da linguagem de ZF , pela maneira usual, podemos internalizar a noção de “satisfatibilidade de uma fórmula $\varphi(\vec{x})$ numa estrutura $\mathbf{M}$ para uma avaliação de variáveis $[\vec{x}/\vec{d}]$ ”, que escreveremos, abusando algo da linguagem, como $\mathbf{M} \models \varphi(\vec{x})[\vec{x}/\vec{d}]$, bem como “ $\mathbf{M}$ é uma estrutura auto referente” ($EstAR(\mathbf{M})$)¹ e o que mais interessar possa; de acordo com essas convenções podemos provar o seguinte

¹ $EstAR(\mathbf{M})$ é algo equivalente a

$$\begin{aligned} & \exists D \exists E (\mathbf{M} = \langle D, E \rangle \wedge D \subseteq \Sigma \wedge E \subseteq D \times D \wedge \forall S (T(S) \leftrightarrow Sat(\mathbf{M}, S)) \wedge \\ & \forall \varphi \forall P \forall \theta (P \in D \wedge \varphi \in P \wedge Eq(\varphi(v_0), v_0 \in v_1, \theta(v_1)) \rightarrow \mathbf{M} \models \theta(v_1)[v_1/P]). \end{aligned}$$

TEOREMA 2.9.1 *Se M é uma estrutura auto referente e S é uma sentença então $\models M \models S$ se, e somente se, $\models S \in P$ (No contexto desse teorema, “ M é uma estrutura auto-referente” é o mesmo que afirmar “ $\models EstAR(v_0)[v_0/M]$ tem prova autopoietica”).*

PROVA:

Se, no decurso de uma prova, aparecerem as linhas $\models EstAR(M)$ e $\models M \models S$, obtemos, sem muito esforço, $\models S \in P$, já que isso está implícito na fórmula $EstAR$.

Quando a linha $\models S \in P$ ocorre, é possível obter $\models T(S)$, donde $\models M \models S$, se $\models EstAR(M)$. $\square$

Assumindo a existência de um modelo $N = \langle D, E \rangle$ de $ZF + V = OD$, é possível mostrar a existência de uma subestrutura elementar M de N que seja auto-referente no sentido da definição 1.3.2. De fato, N satisfaz o axioma da extensionalidade, e o axioma $V = OD$ garante também uma boa ordenação de D via uma fórmula $x \leq y$; basta então definir $\mathcal{D}_M$ da forma seguinte: para cada fórmula $\varphi(x)$, tal que $N \models \exists x \varphi(x)$, seja d_φ o menor elemento de D que satisfaz $\varphi(x)$; claramente d_φ satisfaz $\exists x(\varphi(x) \wedge \forall y(\varphi(y) \rightarrow x \leq y))$, ou seja, d_φ é definível por uma fórmula; tomando $\mathcal{D}_M$ a coleção de tais d_φ , e $M = \langle \mathcal{D}_M, E|_{\mathcal{D}_M} \rangle$, M satisfaz o critério de Tarski-Vaught, sendo portanto uma subestrutura elementar de N , ou seja, M é (isomorfo a) uma estrutura auto-referente, de acordo com o teorema 1.3.1.

Expressando por uma fórmula $Str(x)$ a descrição acima feita de M , ie, tomando $Str(x)$ equivalente a algo semelhante a

$$\exists y \exists z (x = \langle y, z \rangle \wedge \forall w (Ax^{ZF + V=OD}(w) \rightarrow x \models w \wedge \dots)),$$

e formalizando o argumento do parágrafo precedente na teoria dos conjuntos, teremos provado no sistema

$$:\vdash \text{Con}(ZF + V = OD) \rightarrow \exists x \text{Str}(x).$$

TEOREMA 2.9.2 (da estrutura auto-referente) *A expressão abaixo é dedutível.*

$$:\models \text{Con}(ZF + V = OD) \rightarrow \exists M \text{Str}(M).$$

Não obstante, para obtermos $:\models \exists M \text{EstAR}(M)$ é ainda necessário provar que $:\models M \models S \leftrightarrow S \in \mathbf{P}$. Não é isso que o resultado acima afirma, porém a existência duma estrutura M que satisfaça $\text{EstAR}(M)$ será obtida posteriormente no teorema 3.2.2. Também será futuramente provado que $\text{Con}(ZF + V = OD)$ é derivável autopoiéticamente, isso permitir-nos-ia enunciar o teorema precedente na seguinte forma:

TEOREMA 2.9.3 (da estrutura auto-referente) *A expressão abaixo é dedutível*

$$:\models \exists M \text{Str}(M).$$

2.10 Consistência

Esta seção é consagrada a dois objetivos. O primeiro deles é obter a consistência do sistema relativamente a uma certa teoria. Como já tivemos oportunidade de observar na introdução deste trabalho, em vista da possibilidade (ver o teorema 3.2.1 e o parágrafo a ele subsequente) de $\text{Con}(ZF + V = OD)$ ser dedutível no sistema, essa teoria deve ser algo mais

forte que $ZF + V = OD$. Será mostrado que uma teoria como $ZF + V = OD + \text{“Existe um modelo natural de } ZF + V = OD\text{”}$ é o bastante. Outro objetivo é o de provar que as noções de provas, sem e com modificações, não são equivalentes, sendo a segunda estritamente mais forte.

Com o propósito de demonstrar-se a consistência do sistema, introduziremos as definições e resultados que se seguem; a idéia é mostrar que, no modelo natural, a aplicação das regras autopoieticas preserva a verdade, dessa forma, uma derivação autopoietica não poderá produzir contradições, mostrando assim a consistência do sistema. Nas definições e resultados a seguir, muitas noções e fatos enunciados em um lugar dependem recursivamente de outras noções e fatos enunciados em um outro lugar, contudo não haverá “petição de princípio”, como ficará claro no teorema 2.10.1.

Seja então $M = \langle D, \in \rangle$ o modelo natural da teoria $ZF + V = OD$. Durante o transcorrer de uma derivação autopoietica vamos definir uma função de denotação δ , que associa a objetos internos do sistema, certos elementos de D , o domínio da estrutura M , e uma noção de satisfação em M . Ela é como segue:

DEFINIÇÃO 2.10.1 *Suponhamos que L_i seja uma linha do tipo descrito abaixo, ocorrendo em uma derivação autopoietica, definimos a função δ e a noção de satisfatibilidade como segue:*

1. $\delta(P) = P^M$, o conjunto verdade de M .
2. Se L_i é obtida por meio da regra 2.2.1, colocamos

$$n \in \text{dom } \delta$$

e

$$\mathbf{M} \models D_n(x)[x/\delta(n)].$$

3. Se L_i é obtida por meio da regra 2.2.2, colocamos

$$0 \in \text{dom } \delta$$

e

$$\mathbf{M} \models D_0(x)[x/\delta(0)].$$

4. Se L_i é obtida por meio da regra 2.2.3, ela foi precedida pela linha

$$\Gamma : \models \exists v_m A[f],$$

sendo todas as ε -fórmulas de A ε_Γ -fórmulas. Caso

$$\mathbf{M} \models \Gamma,$$

$$f \in \text{dom } \delta$$

e

$$\mathbf{M} \models \exists v_m A[f],$$

colocaremos $\varepsilon v_m A[f]$ em $\text{dom } \delta$. Nesse caso o valor de δ no ε -termo é definido de modo que

$$\begin{aligned} \mathbf{M} \models \exists v_m (v_m = x \wedge A(v_m)) \wedge \\ \forall v_m (A(v_m) \rightarrow x \leq v_m)[x/\delta(\varepsilon v_m A[f]); \delta \circ f] \end{aligned}$$

5. Se L_i é obtida por meio da regra 2.2.5, colocaremos

$$\alpha \in \text{dom } \delta$$

onde

$$\alpha \in \{\wedge, \vee, \rightarrow, \leftrightarrow, \neg, \forall, \exists, \in, =, S\}.$$

$\delta(\alpha)$ é definido como

$$\mathbf{M} \models D_\alpha(x)[x/\delta(\alpha)].$$

6. Se L_i é obtida por meio da regra 2.2.6, a definição é totalmente análoga à anterior com as adaptações convenientes.

7. L_i é obtida por meio da regra 2.2.7.

Caso $M \models \Gamma$, colocaremos $A \in \text{dom } \delta$ e $\delta(A)$ satisfaz

$$M \models D_A(x)[x/\delta(A)].$$

8. L_i é obtida por meio da regra 2.2.8.

Caso $a_0, \dots, a_n \in \text{dom } \delta$, colocaremos

$$\{a_0, \dots, a_n\} \in \text{dom } \delta$$

e

$$\delta(\{a_0, \dots, a_n\}) = \{\delta(a_0), \dots, \delta(a_n)\},$$

e reciprocamente, desde que $M \models \Gamma$.

9. L_i é obtida por meio da regra 2.3.1.

Nesse caso $M \models v_m \in v_n[\alpha]$, onde α é uma avaliação de variáveis, tem a definição usual. O mesmo aplica-se para a fórmula $= v_m v_n$.

10. L_i é obtida por meio da regra 2.3.2.

Quando

$$M \models \Gamma$$

e

$$M \models \exists v_k (B(v_k) \wedge \text{Form}(v_k))[f],$$

seja A a fórmula cujo número de Gödel é

$$\delta(\varepsilon v_k (B(v_k) \wedge \text{Form}(v_k))[f]).$$

Diremos que A corresponde a

$$\varepsilon v_k (B(v_k) \wedge \text{Form}(v_k))[f]$$

e colocaremos

$$\mathbf{M} \models \varepsilon v_k (B(v_k) \wedge \text{Form}(v_k))[f][\alpha]$$

se, e somente se,

$$\mathbf{M} \models \mathcal{A}[\alpha],$$

onde α é uma avaliação de variáveis.

11. L_i é obtida por meio da regra 2.3.3.

Nesse caso $\mathbf{M} \models A \wedge B[\alpha]$, se, e somente se, $\mathbf{M} \models A[\alpha]$ e $\mathbf{M} \models B[\alpha]$, onde α é uma avaliação de variáveis. Dá-se para $\neg$ a definição correspondente.

12. L_i é obtida por meio da regra 2.3.4.

Nesse caso $\mathbf{M} \models \forall v_m A[\alpha]$, se, e somente se,

$$\forall d \in D, \mathbf{M} \models A[v_m/d; \alpha'],$$

onde α é uma avaliação de variáveis e α' é α a menos de v_m .

13. L_i é obtida por meio da regra 2.3.5.

Caso $\mathbf{M} \models \Gamma$, diremos que $\mathbf{M} \models A[f]$ quando

$$\mathbf{M} \models A[\delta \circ f].$$

14. Se L_i é obtida por meio da regra 2.3.6, $\mathbf{M} \models 0$.

15. L_i é obtida por meio da regra 2.3.7.

Nesse caso, $\mathbf{M} \models \Gamma \cup \{A[f]\}$, se, e somente se,

$$\mathbf{M} \models \Gamma \text{ e } \mathbf{M} \models A[f];$$

caso contrário, $\mathbf{M} \models \neg \Gamma$.

Tomemos agora uma fórmula A do sistema, de sorte que todas as ε -fórmulas que ocorram em A sejam ε_Γ -fórmulas e suponhamos que

$$M \models \Gamma,$$

bem como

$$M \models \exists v_p (B(v_p) \wedge Form(v_p))[f],$$

para todas as ε -fórmulas de A . Seja A' o que resulta de A após a substituição de todas as ε -fórmulas que nela ocorram, por suas fórmulas *correspondentes* (ver a definição anterior). Com essas notações vale o

LEMA 2.10.1 $M \models A[\alpha]$ se, e somente se, $M \models A'[\alpha]$.

PROVA(por indução):

Quando A é $\in v_m v_n$ ou $= v_m v_n$, A' é A .

Quando A é $B \wedge C$, A' é $B' \wedge C'$ e a indução é trivial; vale o mesmo para o caso em que A é $\neg B$.

Quando A é $\forall v_m B$, teremos A' na forma $\forall v_m B'$; daí

$$\begin{aligned} M \models A[\alpha] &\iff \forall d \in D, M \models B[v_m/d; \alpha'] \\ M \models A[\alpha] &\iff \forall d \in D, M \models B'[v_m/d; \alpha'] \\ M \models A[\alpha] &\iff M \models \forall v_m B'[\alpha] \\ M \models A[\alpha] &\iff M \models A'[\alpha]. \end{aligned}$$

Quando A é uma ε -fórmula, $M \models A[\alpha] \iff M \models A'[\alpha]$ por definição. $\square$

Não é difícil também de verificar-se que $\delta(A) = \delta(A')$. Utilizando-se ainda das notações precedentes, estenderemos a última definição para as seguintes expressões:

16. Diremos que $\mathbf{M} \models \Gamma : \models A[f]$ quando

$$\mathbf{M} \models \neg \Gamma \text{ ou } \mathbf{M} \models \Gamma \text{ e } \mathbf{M} \models A[f].$$

17. Diremos que $\mathbf{M} \models \Gamma : \vdash A$ quando sempre que

$$cl(B'_i) \in \mathbf{P}^{\mathbf{M}}, \forall B_i \in \Gamma,$$

acarretar $cl(A') \in \mathbf{P}^{\mathbf{M}}$, onde $cl(\varphi)$ é o fecho universal de φ .

De posse dessas noções vamos mostrar que vale o

TEOREMA 2.10.1 *Seja Π um programa cujas linhas estejam satisfeitas em $\mathbf{M}$, ie,*

$$(\Gamma : \models A[f]) \in \Pi \implies \mathbf{M} \models \Gamma : \models A[f].$$

Seja $L_1, \dots, L_n$ uma Π -derivação, daí

1. *Se L_i é $IO(\Delta; \alpha)$ e $\mathbf{M} \models \Delta$, então $\alpha \in \text{dom } \delta$.*
2. *Se L_i é $F_{\mathcal{L}}(A)$, todas as ε -fórmulas de A são ε_{Δ} -fórmulas e $\mathbf{M} \models \Delta$, então $A \in \text{dom } \delta$ e $\delta(A)$ é o número de Gödel de A na estrutura $\mathbf{M}$.*
3. *Se L_i é $F_{\Lambda}(\Delta; A[f])$ e $\mathbf{M} \models \Delta$, então $f \in \text{dom } \delta$.*
4. *Se L_i é $SH(\Delta)$, então $\mathbf{M} \models \Delta$ ou $\mathbf{M} \models \neg \Delta$.*
5. *Se L_i é $\Delta : \vdash A$, todas as ε -fórmulas de A são ε_{Δ} -fórmulas, então $\mathbf{M} \models \Delta : \vdash A$.*
6. *Se L_i é $\Delta : \models A[f]$, todas as ε -fórmulas de A são ε_{Δ} -fórmulas então $\mathbf{M} \models \Delta : \models A[f]$.*

PROVA:

Examinemos, para cada linha L_i , a regra envolvida na sua obtenção.

1. O caso das regras 2.2.1, 2.2.2, 2.2.5, 2.2.6 e 2.2.7 é evidente a partir da definição anterior.
No caso da regra 2.2.3, utilizando-se a indução no último item deste teorema junto com $M \models \Delta$, chegaremos à conclusão que $\alpha \in \text{dom } \delta$.
2. O caso da regra 2.3.1 é evidente pela definição anterior, o mesmo aplica-se para as regras 2.3.3 e 2.3.4.
O caso da regra 2.3.2 envolve a indução sobre o último item deste teorema junto com a hipótese $M \models \Delta$.
3. Foi utilizada a regra 2.3.5; aplicamos então a indução sobre a linha $IO(\Delta; f)$ (que ocorreu anteriormente) junto com o primeiro item deste teorema.
4. No caso da regra 2.3.6, $M \models \Delta$ por definição.
No caso da regra 2.3.7, aplicamos a indução sobre as linhas

$$F_{\Lambda}(\Xi; A[f]) \text{ e} \\ SH(\Xi);$$

caso tenhamos $M \models \neg \Xi$, teremos $M \models \neg \Delta$; senão, $f \in \text{dom } \delta$, em virtude do terceiro item deste teorema. Assim faz sentido perguntar-se se $M \models A[f]$; dependendo da resposta obteremos ou $M \models \Delta$ ou $M \models \neg \Delta$.

5. Vamos supor que $M \models \Delta$ (pois $SH(\Delta)$ é uma linha que já ocorreu).

Se a regra utilizada foi a 2.4.1, teremos certamente $M \models cl(A')$.

Se foi utilizada a regra 2.4.2, basta observar que a linha $\Delta \vdash A$ precedeu $\Delta \vdash \forall x A$, ou seja, sabemos que $cl(A') \in P^M$. Como

$$\vdash cl(A') \rightarrow cl(\forall x A'),$$

certamente temos $cl(\forall x A') \in P^M$ e portanto

$$M \models \Delta \vdash \forall x A.$$

O restante dos casos, correspondentes às regras 2.4.3, 2.4.4, 2.4.5, 2.4.6, 2.4.6 e 2.4.7 tem provas análogas.

Vejamos agora a regra 2.6.5. Pelo último item deste teorema, temos

$$M \models \Delta \vdash A \in P,$$

ou seja, $M \models A \in P$, o que acarreta $A \in P^M$, por esta razão $cl(A') \in P^M$, ou seja

$$M \models \Delta \vdash A.$$

6. Semelhantemente ao que viemos fazendo, vamos exemplificar com alguns casos típicos. Suponhamos que $M \models \Delta$.

Caso da regra 2.5.2: aplicando a indução sobre todas as linhas que antecederem L_i , sabemos, entre

outras coisas, que “ v_m não é livre em A ”; devemos, conforme a situação, interpretar isso como

$$:\models \neg \text{Livre}(v_0, v_1)[v_0/v_m; v_1/A];$$

por indução, teremos

$$M \models \neg \text{Livre}(v_0, v_1)[v_0/v_m; v_1/A];$$

o que significa, no modelo natural, que v_m não é livre em A' ; como

$$M \models A[v_m/a; f],$$

temos

$$M \models A'[v_m/\delta(a); \delta \circ f],$$

donde $M \models A'[\delta \circ f]$, ou seja, $M \models A[f]$, como desejado.

O restante dos casos, pelas regras 2.5.1, 2.5.3, 2.5.4, 2.5.5, 2.5.6, 2.5.7, 2.5.8, 2.5.9, 2.6.1, 2.6.2, 2.6.3, 2.6.4, 2.6.6, 2.6.7, 2.6.8, 2.6.9, decorrem ou por definição, ou por ser M o modelo natural (caso, por exemplo, da regra 2.6.7), ou ainda por argumentos simples da lógica.

Um último caso é quando $\Delta \models A[f]$ está em Π ; a hipótese do teorema, porém, cuida deste caso. $\square$

LEMA 2.10.2 *Se a expressão $\Gamma : \models A[f]$ tem prova auto-poiética, então $M \models \Gamma : \models A[f]$.*

PROVA:

Seja $\langle L_i, D_i, \Pi_i : 1 \leq i \leq n \rangle$ uma dedução de $\Gamma : \models A[f]$. Da definição 2.7.4 segue-se que $L_1, \dots, L_n$

é uma Π_n -derivação, assim, em vista do teorema anterior, basta mostrar que toda linha de Π_n está satisfeita na estrutura M . Por indução, suponha que L_i é $\Gamma : \models A[f]$ e que

- Toda linha de Π_{i-1} está satisfeita em M .
- $\Gamma, \Sigma, D_i, \Pi_i$ e Π_{i-1} satisfazem as condições da definição 2.7.4.

Se $M \models \Gamma$, então toda linha de Σ está satisfeita em M ; como D_i é uma Σ derivação, segue-se que $M \models A[f]$; assim $M \models \Gamma : \models A[f]$.

Se não ocorrer $M \models \Gamma$, seja $B_j[g_j]$ o menor elemento de Γ , numa ordem conveniente, que não é satisfeito em M . Seja D'_i a parte de D_i que precede $B_j[g_j]$. D'_i é uma Σ' derivação, onde

$$\Sigma' = \Pi_{i-1} \cup \{ : \models B_1[g_1], \dots, : \models B_{j-1}[g_{j-1}] \}.$$

Do fato de $: \models B_j[g_j]$ ser a primeira linha de D_i que não aparece em D'_i , decorre que há uma linha formada por $IO(0; g_j)$ em D'_i . Como toda linha de D'_i está satisfeita em M , segue-se que $g_j \in \text{dom } \delta$ e que, portanto, a pergunta $M \models B_j[g_j]$ faz sentido em M . Obviamente a resposta é não, e por isso $M \models \neg \Gamma$, o que acarreta $M \models \Gamma : \models A[f]$. $\square$

Como corolário dos resultados anteriores, temos o

TEOREMA 2.10.2 *A teoria $T(P)$ é consistente.*

PROVA:

Senão, poder-se-ia autopoieticamente obter a expressão $: \models \perp \in P$, uma cuja consequência seria $: \models \perp$, o que conduz a $M \models \perp$. $\square$

O sistema formal de provas autopoiéticas é forte o suficiente para permitir a internalização da sua metamatemática; seja então $PrAuPo(v_0)$ a fórmula da linguagem de ZF cuja interpretação no modelo natural $\mathbf{M}$ de $ZF + V = OD$ é o subconjunto dos números de Gödel das sentenças S da linguagem de ZF para as quais $:\models S \in \mathbf{P}$ tenha prova autopoiética; utilizando-se do teorema do ponto fixo, é fácil ver que $:\models \sigma \leftrightarrow \neg PrAuPo(\ulcorner \sigma \urcorner) \in \mathbf{P}$ tem prova autopoiética; consequentemente, de acordo com o lema 2.10.2 devemos ter $\mathbf{M} \models \sigma$ sem que $:\models \sigma \in \mathbf{P}$ tenha prova autopoiética, ou seja, a teoria não é completa. Isso explica a introdução das ε -fórmulas, sem o concurso das quais não seria possível o teorema 2.8.3, essencial para a internalização no sistema das estruturas autorreferentes, objeto do teorema 2.9.1. Essa discussão também esclarece porque se deve interpretar, consoante o item 1 da definição 2.10.1, a constante $\mathbf{P}$ no modelo natural como o conjunto dos números de Gödel das sentenças verdadeiras e não como o das que sejam autopoieticamente dedutíveis: com essa última interpretação não poderíamos obter nada equivalente ao teorema 2.10.1; para ver como, suponha que $\mathbf{P}$ pudesse ser interpretado como o conjunto dos números de Gödel das sentenças que admitam uma prova autopoiética; tomando-se a mesma σ acima que é verdadeira em $\mathbf{M}$, mas para a qual $:\models \sigma \in \mathbf{P}$ não tenha prova autopoiética, considere-se o programa $\Pi = \{\neg\sigma : \models \perp \in \mathbf{P}\}$ e a Π -derivação abaixo

1. $\neg\sigma : \models \perp \in \mathbf{P}$.
2. $\neg\sigma : \vdash \perp$.
3. $:\vdash \neg\sigma \rightarrow \perp$.

4. $:\vdash \sigma$.

5. $:\models \sigma \in \mathbf{P}$.

Daí $\mathbf{M} \models :\models \sigma \in \mathbf{P}$ teria que ser verdadeiro, posto que toda linha de Π o é; obteríamos então uma contradição.

Vamos mostrar que a noção de prova autopoiética automodificável é estritamente mais forte que a sem automodificações. Com isso estará provado que as modificações alteram genuinamente o curso das demonstrações conferindo ao sistema de provas um *caráter dinâmico* muito apreciado nos sistemas de inferência elaborados com o propósito de imitarem o raciocínio natural. O resultado abaixo mostra que a propriedade de ser $\mathbf{P}$ um predicado verdade depende, de modo essencial, da noção de prova adotada. De fato, será mostrado que com a noção mais restrita de prova, a que não inclui modificações, não é possível provar-se que, para toda sentença σ da linguagem, tem-se $:\models \sigma \leftrightarrow \sigma \in \mathbf{P}$; fica, portanto, estabelecido que o conceito de prova com modificações é estritamente mais forte que o original (ver def 2.7.1), bem como o fato de ser $\mathbf{P}$ um predicado verdade depender da noção de prova empregada.

TEOREMA 2.10.3 *A noção de prova auto-modificável é estritamente mais forte que a sem auto-modificações, mais propriamente, há uma sentença σ tal que $:\models \sigma \leftrightarrow \sigma \in \mathbf{P}$ não é dedutível no sistema (sem auto-modificações).*

PROVA:

Para tanto, considere o objeto interno

$$\varepsilon y \forall x (x \in y \leftrightarrow DSM(x)),$$

onde DSM é uma fórmula da linguagem da teoria dos conjuntos cuja interpretação standard é “ x é número de Gödel de uma sentença σ da linguagem da teoria dos conjuntos tal que $:\vdash\sigma$ possui uma prova sem modificações”. Para simplificar a escrita, seja T tal ε -termo.

A seguir, substituamos em toda prova sem modificações de sentenças da linguagem da teoria dos conjuntos, a constante P por T . Tal operação produzirá uma sequência de linhas que afirmamos serem satisfeitas no modelo natural. Basta verificar a satisfatibilidade das regras onde ocorre a constante P , sendo a 2.6.6 a única que mereça um pequeno comentário. De acordo com ela, devemos mostrar que se $:\models \sigma \in P$ é satisfeita em M substituindo-se a interpretação de P pela de T , quando a mesma ocorre após uma linha (de uma prova sem modificação) da forma $:\vdash\sigma$; ora, tal fato é evidente pela própria forma pela qual se obteve $:\models \sigma \in P$, ie, através de uma prova sem auto-modificações.

Se as noções de provas fossem realmente coincidentes, obteríamos, sem modificações, uma prova de $:\models \sigma \leftrightarrow \sigma \in T$, para toda sentença σ da linguagem da teoria dos conjuntos, satisfatível no modelo natural; tal fato leva à existência dum conjunto verdade definível por uma fórmula da linguagem da teoria dos conjuntos, contradizendo o teorema da indefinibilidade da verdade de Tarski. $\square$

Capítulo 3

Sistemas de Ordem Superior

3.1 Introdução

Neste capítulo ampliaremos nosso sistema de forma a incorporá-lo dentro de um outro que é praticamente uma réplica do original. Esse sistema “de nível superior” será, contudo, estritamente mais forte, quer em poder expressivo quer em recursos de prova; nele provaremos a existência de um objeto interno M que satisfaz a fórmula $EstAR(x)$ (ver a seção 2.9), ou seja, a existência de um objeto cuja teoria é P , afirmação que sequer podia ser expressada na sintaxe do sistema anterior. Tal M é uma imagem que o sistema com dois níveis faz de si próprio; nesse sentido o sistema pode ser considerado como possuindo um nível de consciência. Como um exemplo de uma afirmação que não poderia ser provada no sistema original e expressável em sua sintaxe temos a sentença $Con(P)$ (ver teorema 3.2.1), ou seja, o nível superior tem recursos suficientes para provar a consistência do de nível mais baixo.

Os fatos aludidos no parágrafo anterior formam o corpo da próxima seção. Seguindo-a, temos uma seção dedicada à

tarefa de mostrar que as partes hereditariamente finitas dos dois níveis são isomorfas, o que nos permitirá passar livremente objetos finitos de um nível para o outro. Finalmente a última seção é consagrada a uma prova de consistência desses sistemas relativamente à uma teoria do tipo $ZF + V = OD + \textit{Existe um grande cardinal conveniente}$.

3.2 Extensão de mais um nível

Vamos ampliar o sistema formal de forma a produzir uma sua réplica num nível superior que seja, contudo, estritamente mais forte que o original; além disso, o sistema original estará embutido no de nível mais alto. Para tanto, ampliaremos a linguagem, nela incluindo o antigo símbolo $\mathbf{P}$ grafando-o, porém, como $\mathbf{P}_0$; dessa forma, a antiga regra 2.3.1 passa a incluir as cláusulas $F_{\mathcal{L}}(\in v_m \mathbf{P}_0)$, $F_{\mathcal{L}}(\in \mathbf{P}_0 v_m)$, $F_{\mathcal{L}}(= v_m \mathbf{P}_0)$ e $F_{\mathcal{L}}(= \mathbf{P}_0 v_m)$.

O lugar de $\mathbf{P}$ passará a ser ocupado por $\mathbf{P}_1$. Além disso, enriquecemos o sistema com as seguintes duas novas regras:

REGRA 3.2.1 *Se Γ é um conjunto de sentenças, tivermos $\Gamma : \models A[v_m/\mathbf{P}_0]$, a única variável livre de A é v_m e todas as ε -fórmulas de A são ε_{Γ} -fórmulas, então $\Gamma : \vdash \hat{A}$, onde $\hat{A}$ resulta de A pela substituição de v_m por $\mathbf{P}_0$.*

REGRA 3.2.2 $: \models \forall x (PrAuPo(x) \rightarrow x \in y)[y/\mathbf{P}_1]$.

E assim, procedendo indutivamente, podemos repetir o que acabamos de fazer com $\mathbf{P}_0$ e $\mathbf{P}_1$ em um outro nível ainda mais alto, ie, refazer tudo isso com $\mathbf{P}_1$ e $\mathbf{P}_2$, numa cadeia infinita.

Ficará evidente a seguir como reobter os resultados análogos em níveis superiores.

Para mostrarmos que esse novo sistema é estritamente mais forte que o anterior, vamos mostrar que seus recursos são suficientes para nele demonstrar a consistência do sistema que o precedeu, assim teremos exibido explicitamente uma sentença que jamais teríamos podido provar anteriormente. Expressaremos a consistência do sistema anterior por $Con(\mathbf{P}_0)$.

Podemos exprimir a consistência do sistema antigo por meio da seguinte sentença:

$$\forall x \neg (PrAuPo(x) \wedge D_{\perp}(x)).$$

Ela, por sua vez, é logicamente equivalente a

$$\exists x (PrAuPo(x) \wedge D_{\perp}(x)) \rightarrow \perp,$$

onde $\perp$ é $\forall x (x \neq x)$. Com o fito de simplificar a escrita, escreveremos a sentença $\exists x (PrAuPo(x) \wedge D_{\perp}(x))$ como $\neg con$.

TEOREMA 3.2.1 : $\vdash Con(\mathbf{P}_0)$ tem prova autopoietica.

PROVA:

Vamos mostrar que $\neg con : \vdash \perp$, o que é suficiente.

Assumindo $: \models \neg con$, podemos escrever

$$: \models PrAuPo(x) \wedge D_{\perp}(x)[x/\sigma],$$

onde σ é $\varepsilon x (PrAuPo(x) \wedge D_{\perp}(x))$.

Consequentemente, vale $: \models D_{\perp}(x)[x/\sigma]$.

Mas, como $: \models \forall x \forall y (D_{\perp}(x) \wedge D_{\perp}(y) \rightarrow x = y)$ tem prova autopoietica e como $: \models D_{\perp}(x)[x/\perp]$ é uma regra, não

é difícil concluir-se que $:\models \perp = \sigma$; daí pode-se escrever $:\models \text{PrAuPo}(x)[x/\perp]$ e, utilizando-se agora da regra 3.2.2, $:\models \perp \in P_1$, o que leva a $\neg \text{con} :\models \perp \in P_1$, de acordo com a definição de prova autopoiética. Neste momento, aplicando-se a regra 2.6.5 adequadamente modificada, obtemos $\neg \text{con} :\vdash \perp$. $\square$

A mesma demonstração que acabamos de apresentar, com as adaptações convenientes, poderia ter sido utilizada no sistema mais fraco para se obter $\text{Con}(ZF + V = OD)$, isto é, poderíamos, no capítulo anterior, ter mostrado que $:\vdash \text{Con}(ZF + V = OD)$ é autopoiéticamente derivável. Isso explica também porque tivemos que assumir uma teoria mais forte que $ZF + V = OD$, em particular que ZF , para provar a consistência do sistema: caso contrário estaríamos nos conflitando com o segundo teorema de Gödel.

Vamos abordar agora a questão do observador. As sentenças σ tais que $\sigma \in P_0$ formam uma teoria consistente (teorema 3.2.1) e completa (teorema 2.8.3). Seja então $M = \langle D, E \rangle$ um modelo para a teoria P_0 . Devido à hipótese $V = OD$, podemos supor que todo elemento de D seja definível por meio de uma fórmula (ver comentários do parágrafo subsequente ao teorema 2.9.1), ou seja, que M é auto-referente; desta forma M é uma estrutura auto-referente cuja teoria é P_0 , vale, portanto, o seguinte

TEOREMA 3.2.2 *Dentro do sistema ampliado é possível deduzir a expressão abaixo, isto é, pode-se provar, dentro do sistema, a existência do observador (ou seja, O OBSERVADOR COMO PARTE DA DESCRIÇÃO).*

$$:\vdash \exists M EstAR(M).$$

A escolha do nome do teorema acima justifica-se se estivermos de acordo com as seguintes considerações: (1) o sistema é capaz, com seus próprios recursos, de provar a existência de alguma coisa muito semelhante a si mesmo, a saber, de um modelo sobre a “realidade”; (2) esse mesmo modelo, por sua vez, é obtido totalmente a partir da atividade linguística do sistema na medida em que pode ser pensado como formado por predicados linguísticos, no sentido da definição 1.3.2. Essas ponderações estão muito próximas das de Maturana e Varela¹, cujos trechos essenciais passamos a reproduzir:

(iii) *“An autopoietic system capable of interacting with its own states (as an organism with a nervous system can do), and capable of developing with others a linguistic domain, can treat its own linguistic states as a source of deformations² and thus interact linguistically in a closed linguistic domain. Such a system has two remarkable properties:*

1. *Trough recursive interactions with linguistically generated states it can treat some of these states as objects of further interactions, giving rise to a metadomain of consensual distinctions that appears to an observer as a domain of interactions with representations of interactions. When*

¹H. R. Maturana and F. J. Varela, *Autopoiesis and Cognition—The Realization of the Living*, pp 120–121, Boston Studies in the Philosophy of Science, Vol 42, 1980, D. Reidel Publishing Company.

²Ou, dito de outra forma, como se eles fossem um agente externo.

this happens the system operates as an observer. The domain of such recursive interactions is, in principle, infinite because once the system has attained the mechanism for doing so there is no moment in which it will not be in the position of recursively interacting with its own states, unless autopoiesis is lost...

2. *A living system capable of being an observer can interact with those of its own descriptive states which are linguistic descriptions of itself³. By doing so it generates the domain of self-linguistic descriptions within which it is an observer of itself as an observer, a process which can be necessarily repeated in an endless manner⁴...*

3.3 Invariância do Finito

Mostraremos nesta seção que as partes finitas do Observador e do sistema são isomorfas. Com isso será possível intercambiarmos as noções de finitude. Vamos começar provando que há um isomorfismo entre os números naturais do Observador e os do sistema, para tanto considere a seguinte fórmula $\Pi(P, n)$:

$$\begin{aligned} \Pi(P, n) \leftrightarrow & \exists \varphi (M \models \exists! x \varphi(x) \wedge \\ & M \models \varphi(x)[x/P] \wedge M \models \forall x (\varphi(x) \rightarrow x \in \omega) \wedge \\ & Form(\varphi, \omega) \wedge \omega \models \varphi(x)[x/n]), \end{aligned}$$

onde $\exists \varphi$ está relativizado à $Form(x)$ e $Form(\varphi, \omega)$ significa que a fórmula φ está relativizada à ω .

³Como é o caso da M que satisfaça $EstAR(x)$.

⁴O que corresponde à sucessão de sistemas $P_0 \hookrightarrow P_1 \hookrightarrow P_2 \hookrightarrow \dots$

O passo importante é mostrar que $\forall P \exists n \Pi(P, n)$, sendo relativamente corriqueiro mostrar que

$$\begin{aligned} & \forall P \forall m \forall n (\Pi(P, m) \wedge \Pi(P, n) \rightarrow m = n) \wedge \\ & \Pi(0, 0) \wedge \forall P \forall n (\Pi(P, n) \rightarrow \Pi(S(P), S(n))) \wedge \\ & \forall n \exists P \Pi(P, n). \end{aligned}$$

Passemos então à prova de $\forall P \exists n \Pi(P, n)$, ou seja, de que não há, no Observador, números naturais não standard. Nosso objetivo é claramente equivalente à provar que

$$\exists P \forall n \neg \Pi(P, n) \rightarrow \perp.$$

Assumindo o antecedente, podemos concluir

$$\forall n \neg \Pi(\hat{P}, n),$$

onde $\hat{P}$ é o ε -termo correspondente e a última expressão acima abrevia

$$\forall n \neg \Pi(P, n) [P / \varepsilon P \forall n \neg \Pi(P, n)].$$

Contudo, tendo em vista o comentário que precede o teorema 1.3.1, podemos afirmar que

$$\begin{aligned} & \forall P (P \in D \wedge \mathbf{M} \models x \in \omega[x/P] \rightarrow \\ & \exists \varphi (Form(\varphi, \omega) \wedge \mathbf{M} \models \exists! x \varphi(x) \wedge \mathbf{M} \models \varphi(x)[x/P] \wedge \\ & \mathbf{M} \models \forall x (\varphi(x) \rightarrow x \in \omega), \end{aligned}$$

onde D é o domínio da estrutura $\mathbf{M}$. Consequentemente podemos escrever

$$\begin{aligned} & \exists \varphi (Form(\varphi, \omega) \wedge \mathbf{M} \models \exists! x \varphi(x) \wedge \\ & \mathbf{M} \models \varphi(x)[x/\hat{P}] \wedge \mathbf{M} \models \forall x (\varphi(x) \rightarrow x \in \omega). \end{aligned}$$

Tomando $\hat{\varphi}$ o ε -termo correspondente, teremos

$$\begin{aligned} & \text{Form}(\hat{\varphi}, \omega) \wedge \mathbf{M} \models \exists! x \hat{\varphi}(x) \wedge \\ & \mathbf{M} \models \hat{\varphi}(x)[x/\hat{P}] \wedge \mathbf{M} \models \forall x (\hat{\varphi}(x) \rightarrow x \in \omega). \end{aligned}$$

Conjugando essa última expressão com $\forall n \neg \Pi(\hat{P}, n)$, ficamos com $\forall n \neg (\omega \models \hat{\varphi}(x)[x/n])$, o que é equivalente a $\omega \models \neg \exists x \hat{\varphi}(x)$.

Contudo, $\mathbf{M} \models \exists x \hat{\varphi}(x) \rightarrow \exists x \hat{\varphi}(x) \in \mathbf{P}_0$, já que $\mathbf{M}$ é uma estrutura auto-referente cuja teoria é $\mathbf{P}_0$. Se pudermos mostrar que

$$\forall \sigma (\text{Sent}(\sigma, \omega) \rightarrow (\sigma \in \mathbf{P}_0 \leftrightarrow \omega \models \sigma)),$$

teremos concluído que $\hat{\sigma} \wedge \neg \hat{\sigma} \in \mathbf{P}_0$, onde $\hat{\sigma}$ é $\exists x \hat{\varphi}(x)$; daí então $\perp \in \mathbf{P}_0$ segue facilmente.

Para mostrar-se a validade da afirmação que suposemos, notemos que se T é uma teoria consistente que estende ZF então para toda sentença σ da linguagem da teoria dos conjuntos e que esteja relativizada à ω vale

$$T \vdash \sigma \leftrightarrow \forall x (\Delta_\sigma(x) \rightarrow \omega \models x),$$

onde Δ_σ é uma fórmula definidora de σ através de uma aritmetização conveniente. Notando-se que o processo de obtenção de Δ_σ a partir de σ é recursivo, podemos internalizar o esquema acima pela seguinte fórmula

$$\begin{aligned} & \forall T \forall \Delta \forall \sigma (\text{Sent}(\sigma, \omega) \wedge \nabla(\sigma, \Delta) \wedge \\ & \text{Con}(T) \wedge ZF \subseteq T \rightarrow [T \vdash \sigma \leftrightarrow \forall x (\Delta(x) \rightarrow \omega \models x)]), \end{aligned}$$

onde $\nabla(\sigma, \Delta)$ internaliza o processo de obtenção de Δ_σ a partir de σ . Assim, tomando-se $T = \mathbf{P}_0$ obtemos sem dificuldade

$$\begin{aligned} & \forall \Delta \forall \sigma (\text{Sent}(\sigma, \omega) \wedge \nabla(\sigma, \Delta) \rightarrow \\ & (\sigma \in \mathbf{P}_0 \leftrightarrow (\forall x (\Delta(x) \rightarrow \omega \models x)) \in \mathbf{P}_0)), \end{aligned}$$

ou ainda

$$\forall \sigma (Sent(\sigma, \omega) \rightarrow (\sigma \in P_0 \leftrightarrow \omega \models \sigma)),$$

já que viemos cometendo o abuso de escrever-se $\omega \models \varphi$ ao invés de $\forall x (\Delta_\varphi(x) \rightarrow \omega \models x)$. O restante é agora imediato.

Sabendo-se que um isomorfismo entre os naturais estende-se para os hereditariamente finitos, podemos enunciar o

TEOREMA 3.3.1 *Há um isomorfismo entre as partes finitas do Observador e o sistema.*

3.4 Um modelo conveniente com sua interpretação

Antes de mais nada vamos admitir a existência dum modelo natural de $ZF + V = OD$ que possua uma expansão elementar, isto é, de um M_0 e de um M_1 , ambos modelos de $ZF + V = OD$, com a propriedade de M_0 seja subestrutura elementar de M_1 . No momento oportuno, será provada a existência de tais objetos a partir da de $0^\sharp$.

Mantendo o espírito da definição 2.10.1, podemos definir nesse novo contexto uma nova função de denotação δ que associa a P_1 , em M_1 , o conjunto dos números de Gödel das sentenças verdadeiras em M_1 , com uma diferença contudo: agora P_0 faz parte da linguagem. Fica evidente então que é primeiro preciso definir a semântica para essa linguagem ampliada com P_0 . Isso é feito da maneira usual; define-se a interpretação da constante P_0 , após o que o sentido da frase " $M_1 \models v_0 \in P_0[v_0/d_1; \delta(P_0)]$ " é o de ' $d_1 \in^{M_1} \delta(P_0)$ ', etc... (não escreveremos mais $\in^{M_1}$), o restante processando-se como na

definição 2.10.1. Resta atribuir a $\delta(\mathbf{P}_0)$ um elemento conveniente de D_1 e a maneira natural de fazê-lo é escolher o conjunto dos números de Gödel das sentenças de $\mathcal{L}$, portanto sem $\mathbf{P}_0$, verdadeiras em $\mathbf{M}_0$ (ou $\mathbf{M}_1$, tanto faz), ou seja $\delta(\mathbf{P}_0)$ é cópia de $\delta(\mathbf{P})$.

De modo totalmente análogo pode-se provar uma cópia do lema 2.10.1, dar-se uma definição semelhante à 2.10.1, bem como obter-se nesse novo contexto os resultados correspondentes ao teorema 2.10.1, ao lema 2.10.2 e ao teorema 2.10.2, com um pequeno cuidado apenas: o de se provar que as novas regras são verificáveis no modelo que estamos considerando. Vamos comprovar apenas a validade do emprego da regra 3.2.1, o da 3.2.2 não oferece dificuldades. Ela, por sua vez, decorre imediatamente do seguinte

LEMA 3.4.1 *Seja A uma fórmula de $\mathcal{L}$, então*

$$\mathbf{M}_1 \models A[\alpha; x/\mathbf{P}_0] \text{ se, e somente se, } \mathbf{M}_1 \models \hat{A}[\alpha],$$

onde α é uma avaliação de variáveis.

PROVA:

- A é atômica, por exemplo, $y \in x$; daí $\hat{A}$ é $y \in \mathbf{P}_0$. Nesse caso, claramente tem-se

$$\mathbf{M}_1 \models y \in x[y/d; x/\mathbf{P}_0]$$

se, e somente se,

$$\mathbf{M}_1 \models y \in \mathbf{P}_0[y/d].$$

- A é $B \vee C$, $\mathbf{M}_1 \models A[\alpha; x/\mathbf{P}_0]$ se, e somente se, $\mathbf{M}_1 \models B \vee C[\alpha; x/\mathbf{P}_0]$, portanto, $\mathbf{M}_1 \models A[\alpha; x/\mathbf{P}_0]$ se,

e somente se, $M_1 \models B[\alpha; x/P_0]$ ou $M_1 \models C[\alpha; x/P_0]$. Aplicando a hipótese de indução, obtemos $M_1 \models A[\alpha; x/P_0]$ se, e somente se, $M_1 \models \hat{B}[\alpha]$ ou $M_1 \models \hat{C}[\alpha]$, como antes, $M_1 \models A[\alpha; x/P_0]$ se, e somente se, $M_1 \models B \hat{\vee} C[\alpha]$; ou seja, $M_1 \models A[\alpha; x/P_0]$ se, e somente se, $M_1 \models \hat{A}[\alpha]$.

- A é $\neg B$. Nesse caso, utilizando-se a mesma técnica do item anterior, obtemos, sem dificuldade, o resultado.
- A é $\exists v_m B$. Como $M_1 \models \exists v_m B[\alpha; x/P_0]$ se, e somente se, $M_1 \models B[\alpha; v_m/d; x/P_0]$; por indução, obtemos $M_1 \models \hat{B}[\alpha; v_m/d]$. Desse modo, $M_1 \models A[\alpha; x/P_0]$ se, e somente se, $M_1 \models \exists v_m \hat{B}[\alpha]$, ou seja, se, e somente se, $M_1 \models \hat{A}[\alpha]$.
- A é $\varepsilon v_p (Form(v_p) \wedge B(v_p))[f]$. Temos então:

$$M_1 \models A[\alpha; x/P_0]$$

se, e somente se,

$$M_1 \models A'[\alpha; x/P_0],$$

onde a ' tem seu significado explicitado na definição 2.10.1.

Dessa forma, $M_1 \models A[\alpha; x/P_0]$ se, e somente se, $M_1 \models \hat{A}'[\alpha]$, como decorrência da hipótese de indução. Contudo, no caso de ε -fórmulas, a definição de $\hat{A}$ é precisamente $\hat{A}'$, que é o único objeto que satisfaz

$$:\models Subst(x, \varphi, y, \psi)[x/x; \varphi/A; y/P_0; \psi/\hat{A}],$$

onde $\text{Subst}(x, \varphi, y, \psi)$ codifica a operação da substituição de x na fórmula φ por y resultando na fórmula ψ . $\square$

Resta então mostrar a existência de modelos como M_0 e M_1 descritos acima. Como já dissemos no início desta seção, M_0 deverá ser uma subestrutura elementar de M_1 . Para tanto admitiremos a existência em V , o universo dos conjuntos, de $0^\sharp$, conjunto cuja existência pode ser inferida a partir de um cardinal de Erdős, $\kappa(\omega_1)$ por exemplo⁵. De fato, da existência de $\kappa = \kappa(\omega_1)$ decorre que a estrutura $\langle L_\kappa, \in \rangle$ tem um subconjunto de indiscerníveis não enumerável, sendo esta última asserção equivalente à existência de $0^\sharp$ ⁶.

Com o concurso de $0^\sharp$, pode-se provar o seguinte resultado:

TEOREMA 3.4.1 *Se $\kappa < \lambda$ são cardinais não enumeráveis, então L_κ é uma subestrutura elementar de L_λ (e também de L)⁷.*

Como L_λ é um modelo de $ZF + V = OD$, graças aos seus ordinais indiscerníveis, obtemos M_0 e M_1 tomando, por exemplo, L_{ω_ω} e $L_{\omega_\omega+1}$. Usando ainda resultados como o acima é possível provar a consistência de qualquer sucessão finita de sistemas $P_0 \hookrightarrow P_1 \hookrightarrow P_2 \hookrightarrow \dots$

⁵ $\kappa(\omega_1)$ é o menor cardinal κ para o qual vale $\kappa \rightarrow (\omega_1)_2^{<\omega}$, isto é, κ é tal que sempre que $\pi = \{P_1, P_2\}$ é uma partição do conjunto de todas as seqüências finitas estritamente crescentes de elementos de κ , existe um subconjunto C_π de κ , de cardinalidade ω_1 , de modo que para cada natural n , todas as seqüências estritamente crescentes de n elementos de C_π estão num mesmo membro P_{i_n} ($i_n = 1, 2$) da partição π .

⁶Ver K. J. Devlin, *op. cit.*, p 174 e 184.

⁷Esse teorema é uma consequência imediata de um resultado de Gaifman, obtido primeiramente a partir da existência de um cardinal mensurável. Dessa prova, Solovay extraiu $0^\sharp$. Ver K. J. Devlin, *op. cit.*, p 182 e 412.

Bibliografia

- [1] K.J. Devlin, **Constructibility**, Perspectives in Mathematical Logic, 1984, Springer Verlag.
- [2] S. Feferman, **Toward useful type-free theories. I**, The Journal of Symbolic Logic, Vol 49, Number 1, 1984.
- [3] P.C. Gilmore, **The consistency of partial set theory without extensionality**, Axiomatic Set Theory, Proceedings of Symposia in Pure Mathematics, 1974, Vol 13, Part II, AMS, Providence, RI, pp 147—153.
- [4] S. Kripke, **Outline of a theory of truth**, Journal of Philosophy, 1975, Vol 72, pp 690—716.
- [5] H.R. Maturana and F.J. Varela, **Autopoiesis and Cognition—The Realization of the Living**, Boston Studies in the Philosophy of Science, Vol 42, 1980, D. Reidel Publishing Company.
- [6] H. Maturana y F. Varela, **El árbol del conocimiento**, Editorial Universitaria, 1987.
- [7] D. Perlis, **Languages with self-reference I: Foundations**, Artificial Intelligence, 25, 1985, pp 301–322.

- [8] D. Perlis, **Languages with self-reference II: Knowledge, Belief and Modality**, Artificial Intelligence, 34, 1988, pp 179–212.
- [9] Robert Rosen, Ed. **Foundations of Mathematical Biology**, Vol 1, 1972, Academic Press.
- [10] F.J. Varela, **Principles of Biological Autonomy**, The North Holland Series in General Systems Research, Vol 2, 1979, Academic Press.