Polinomização de Lógicas: Problemas e Perspectivas

Pietro Kreitlon Carolino

Dissertação apresentada ao
Departamento de Filosofia do
Instituto de Filosofia e Ciências
Humanas da Universidade
Estadual de Campinas para
obtenção do grau de Mestre em
Filosofia (Área de Lógica).

Orientador: **Prof. Dr. Walter Alexandre Carnielli**

CLE-UNICAMP, Campinas, Brasil

Durante este projeto, o autor recebeu apoio financeiro da FAPESP.

Setembro/2009

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IFCH-UNICAMP

Carolino, Pietro Kreitlon
Polinomização de Lógicas: Problemas e Perspectivas / Pietro Kreitlon Carolino – Campinas, SP: [s.n.], 2009.

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

1. Lógica. 2. Álgebra. 3. Polinômios. I. Carnielli, Walter Alexandre. II. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas III. Título.

(cn/ifch)

Título em inglês: Polinomization of Logics: Problems and Perspectives

Palavras-chave em inglês: Logic, algebra, polynomials

Área de concentração: Filosofia

Titulação: Mestre em Filosofia

Banca examinadora: Walter Alexandre Carnielli, Marcelo Esteban Coniglio, Rodrigo de Alvarenga Freire

Dia da defesa: 04-09-2009

Programa de Pós-Graduação: Filosofia

PIETRO KREITLON CAROLINO

POLINOMIZAÇÃO DE LÓGICAS: PROBLEMAS E PERSPECTIVAS

Dissertação de Mestrado apresentada ao Departamento de Filosofia do Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas, sob a orientação do Prof. Dr. Walter Alexandre Carnielli.

Este exemplar corresponde à redação final da Dissertação defendida e aprovada pela Comissão Julgadora em 04 de setembro de 2009.

BANCA

Prof. Dr. Walter Alexandre Carnielli (orientador)

Prof. Dr. Marcelo Esteban Coniglio

Prof. Dr. Rodrigo de Alvarenga Freire

Prof. Dr. Alexandre Costa-Leite (suplente)

Prof. Dr. Hugo Luiz Mariano (suplente)

Dedicado à minha esposa, Luciana.

Agradecimentos

Agradeço aos professores do Centro de Lógica, Ítala D'Ottaviano e Marcelo Coniglio, e especialmente ao meu orientador, Walter Carnielli, pelo que me ensinaram, por sustentarem um programa de Lógica de alto nível, e pela sua paciência com minhas idiossincrasias.

Agradeço aos colegas estudantes do Centro de Lógica, em especial Anderson de Araújo, Rodrigo Freire e Teófilo Reis, por inúmeras conversas enriquecedoras e por sua amizade, que fez do CLE um lar para mim.

Agradeço aos meus pais, Vicente e Priscilla, e irmãs, Ariella, Stefania e Maimbê, por sempre terem acreditado em mim, às vezes mesmo contra as evidências.

Por fim, agradeço especialmente à minha mulher, Luciana, cujo apoio e carinho deram o tom de toda esta fase da minha vida, e que me ensinou mais do que qualquer universidade.

Pro Xuxu

Resumo

A obra de George Boole, pedra fundamental da lógica contemporânea, não separa métodos de análise matemática, de métodos lógicos propriamente ditos. Se, por um lado, a falta de fronteiras metodológicas nítidas tem-lhe valido azedas críticas, por outro lado fazem da obra de Boole uma verdadeira síntese do pensamento formal, herdada de Aristóteles, Leibniz, Newton e dos analistas a partir do século XVII, como Taylor, MacLaurin e Lagrange. O que foi chamado em [12] de *polinomizar* é precisamente a tentativa de reavaliar os métodos oriundos de Boole e Leibniz, que permitem representar a semântica e a sintaxe de diversos sistemas lógicos pela manipulação algébrica. Tirando partido de resultados combinatórios elementares, é possível tratar todas as lógicas multivalentes verofuncionais com base em manipulação polinomial; não somente estas, mas também lógicas não-verofuncionais, e ainda fragmentos da lógica de primeira ordem, que formalizam a teoria clássica de silogismos de Aristóteles.

Este trabalho pretende esclarecer tais questões de forma mais abrangente, e investigar a possibilidade de estender o tratamento da *polinomização* a outras lógicas. São obtidos alguns resultados positivos, como novas demonstrações de teoremas conhecidos, mas também negativos, que mostram as limitações do método. Investiga-se também a relação da polinomização de lógicas com outros tratamentos conhecidos, como paraconsistentização, temporalização, algebrização etc.

Abstract

The work of George Boole, cornerstone of contemporary logic, does not draw a clear distinction between the methods of mathematical analysis, and those of logic proper. If, on the one hand, this lack of well-defined borders has earned it harsh criticism, on the other hand it makes Boole's work a true synthesis of formal thought, inherited from Aristotle, Leibniz, Newton and the 17th-century analysts, such as Taylor, MacLaurin and Lagrange. What was called *polynomizing* in [12] is precisely the attempty to re-evaluate the methods originiating in Boole and Leibniz, which allow one to represent the semantics and syntax of varioius logical systems through algebraic manipulation. Using elementary combinatorial results, it is possible to treat all multivalent truth-functional logics by polynomial manipulation; not only these, but some non-truth-functional logics, and also fragments of first-order logic, which formalize Aristotle's classical theory of syllogisms.

The present work intends to throw light upon such questions in a broader way, and to investigate the possibility of extending the method of *polynomization* to other logics. Some positive results are obtained, such as new proofs of known theorems, but also some negative ones, which show the inherent limitations of the method. We further investigate the relationship between polynomization of logics and other known treatments, such as paraconsistentization, temporalization, algebrization etc.

Sumário

Introdução

Um dos episódios mais fascinantes da História da Matemática, que até hoje rende frutos e é tópico de pesquisa, foi a descoberta da representação polinomial (através de séries infinitas) de funções numéricas transcendentes. Apesar de variantes elementares destes métodos terem sido desenvolvidos previamente na Europa, na China e na Índia, é possível situar este acontecimento na historiografia ocidental em torno do matemático inglês Brook Taylor, e seu livro *Methodus Incrementorum Directa et Inversa* de 1715, que levou ao desenvolvimento das conhecidas expansões de Taylor.

Em linguagem atual, e sob risco de certo anacronismo, podemos dizer que Taylor descobriu que qualquer função f, sob certas hipóteses, pode ser expressa por uma série polinomial infinita, na vizinhança de um ponto-base a:

$$f(x) = \alpha_0(a) + \alpha_1(a) \cdot (x-a) + \alpha_2(a)(x-a)^2 + \dots$$

para certos coeficientes $\alpha_k(a)$.

Ou seja, a fórmula de Taylor permite que funções transcendentais sejam representadas por funções algébricas, polinomiais — em troca, porém, de aceitar expansões infinitas. Ainda mais notável é que os coeficientes sejam múltiplos das derivadas de f em a: ou seja, o comportamento global de uma função pode ser inferido a partir de informação puramente local.

É prática comum afirmar que nem Taylor, nem os matemáticos dos cem anos subseqüentes, estavam realmente cientes de que era necessário "impor hipóteses" sobre a função f para assegurar a validade da fórmula. Porém, esta posição sofre de certo anacronismo: hoje é necessário impor hipóteses porque o conceito de função foi progressivamente alargado para incluir exemplares não-diferenciáveis, não-contínuas etc. Porém, o conceito de função e curva do século XVIII tornava intuitiva a validade desta aproximação, ao menos para x suficientemente próximo de a.

Surpreendentemente, a importância da descoberta de Taylor não obteve qualquer reconhecimento até 1755, quando Euler a utilizou em seu cálculo diferencial; e teve que esperar até 1772 para que Lagrange percebesse sua relevância e a proclamasse "a princípal fundamentação do cálculo diferencial" [20].

Esta declaração deveria soar especialmente poderosa numa época em que não se tinha notícia de funções não-diferenciáveis; as séries de Taylor pareceriam então abarcar todos os possíveis fenômenos matemáticos, como de fato indica

o subtítulo de sua obra posterior, em que refina e esclarece os resultados nesta área: *Linear Perspective: Or, a New Method of Representing Justly All Manner of Objects as They Appear to the Eye in All Situations.*

A representação de funções em formato polinomial, apesar de freqüentemente exigir expansões infinitas, mostrou-se muito fértil enquanto ferramenta de cálculo, heurística e mesmo demonstração. O próprio Taylor, e seus predecessores e sucessores, por exemplo Newton e Euler, repetidas vezes se valeram da relativa simplicidade das operações elementares sobre tais séries para perceber padrões nos coeficientes, e assim descobrir e provar novos teoremas. Ecos desta tradição são ouvidos até hoje em Combinatória, sob a forma das chamadas Funções Geradoras.

No cerne da idéia de Taylor está um dos mais importantes fios condutores de toda a Matemática: a extensão de construções elementares a contextos infinitos. Esta instância particular, a de estender polinômios a séries (convergentes ou formais), é um método poderoso, ainda a ser completamente esclarecido, que chamamos "polinomização". A idéia de que objetos complexos podem ser tratados por representações de tipo polinomial, possivelmente infinitos, encontra hoje expressão por toda a Matemática, da Geometria Analítica à Combinatória.

Um dos pressupostos do presente trabalho é que esta intuição não impactou somente a Matemática, mas também influenciou Boole e outros lógicos. No caso de Boole, os desenvolvimentos posteriores acabaram relegando tal enfoque, por várias razões. Contudo, os trabalhos de Carnielli [11, 12] mostram que é possível resgatá-lo em vários sistemas lógicos, tais como lógicas multivalentes, paraconsistentes, fragmentos da lógica de primeira ordem, entre outros. Deste modo, esclarecem-se alguns aspectos dos métodos de Boole, e dá-se continuidade à tradição de polinomizar.

Polinômios e Lógicas

O emprego de métodos polinomiais no estudo de lógicas encontra sua primeira expressão clara no trabalho de George Boole. Porém, algumas questões ponderadas por ele podem ser entrevistas já no pensamento de Leibniz, que imagina ser possível reduzir o pensamento a uma espécie de cálculo, e se refere à mecanização do raciocínio, nos conhecidos trabalhos reunidos em [29]. De fato, com certa dose de anacronismo, pode-se ver um germe desta idéia nos *Primeiros Analíticos* de Aristóteles, como tentamos esclarecer a seguir.

Neste, que pode ser considerado o primeiro tratado sobre lógica, Aristóteles se propõe a dar uma classificação abrangente de todos os tipos de raciocínio correto. Para este fim, constrói sua teoria dos silogismos, em que delimita as formas lingüísticas que considera relevantes para o pensamento lógico, e lista explicitamente todas as deduções válidas que se pode obter com elas (cf. [3], *Prior Analytics I*). Como resultado, tem-se uma espécie de tabela de raciocínios válidos, que qualquer pessoa, em princípio, pode usar para determinar se um dado argumento está correto de acordo com os ideiais aristotélicos. Este aspecto de *verificabilidade* da argumentação, no sentido em que sempre se poderia determi-

nar qual dos lados de uma discussão está correto, encontra ecos até vinte e dois séculos mais tarde, durante a chamada crise dos fundamentos da matemática, nas preocupações de Hilbert, Gödel e Turing, entre outros.

Porém, as investigações lógicas de Aristóteles eram principalmente motivadas por problemas de fundo ontológico, e é apenas o olhar contemporâneo que pretende nelas enxergar um prenúncio das idéias booleanas. O primeiro filósofo que de fato formulou, claramente, a possibilidade de mecanizar alguns tipos de raciocínio, foi Gottfried Wilhelm Leibniz, em fins do século XVII. Encontram-se manifestações desta idéia na importância que atribuía à invenção de boa notação matemática: um exemplo é sua notação para o cálculo infinitesimal, em que regras familiares da manipulação de frações podem ser aplicadas às derivadas dy/dx, mesmo estas últimas não sendo frações no sentido usual. Esta simples comodidade faz enorme diferença, até funcionanado como heurística para a descoberta de teoremas, que depois podem ser demonstrados de modo mais rigoroso. Por exemplo, no formalismo leibniziano, a regra da cadeia é sugerida por um simples "cancelamento": $dy/dx = (dy/du) \cdot (du/dx)$

Mas Leibniz não se contentava apenas com a matemática, acreditando ser possível tornar mecânico qualquer tipo de raciocínio. Esta linha de pensamento está ligada ao seu projeto da *characteristica universalis*, uma linguagem simbólica e pictográfica hipotética que tornaria possível expressar qualquer conceito precisamente, e cuja compreensão independeria do idioma nativo dos leitores (cf. [30], *On the General Characteristic*, pp. 221-225). Ao contrário do que acontece nas línguas naturais, Leibniz pretende que os símbolos de sua *characteristica* tenham relação direta com aquilo que representam (cf. [29], *On the Art of Combination*, pp. 10–11), de modo que não só a linguagem seja intuitiva, mas que a manipulação de seus símbolos conduza a conhecimento sobre o representado: a estas manipulações dá o nome de *calculus ratiocinator*. Em [30], *Carta a Nicolas Remond, 10/01/1674*, Leibniz escreve sobre uma "álgebra geral" cujo domínio seriam as "verdades da razão". Em carta à princesa Sophia de Hanover (reproduzida em [15], p. 118), Leibniz volta a este assunto, e argumenta que um tal sistema "permitiria a intelectos modestos, com diligência e boa vontade, se não acompanhar, pelo menos seguir os intelectos maiores" — antecipando parcialmente um dos problemas filosóficos nas bases da teoria da computabilidade: a noção do computador como pessoa que faz cálculos de acordo com regras pré-estabelecidas.

Similarmente à capacidade do "intelecto menor" de seguir o "intelecto maior" mecanicamente, duas pessoas que discordassem sobre qualquer assunto poderiam, graças ao *calculus*, resolver sua disputa de maneira completamente objetiva, apenas por meio da manipulação de símbolos. Este é o significado do aforisma leibniziano, revelado na mesma missiva a Sophia: "calculemos".

Em ainda outras passagens, Leibniz leva às últimas conseqüências a idéia do conhecimento puramente simbólico, e da mecanização do pensamento. Conforme [33], admitia equações sem conteúdo aritmético, como $x + x = x$, e usava o termo "pensamento cego" para se referir ao raciocínio puro reduzido ao cálculo simbólico. Em seu *Elementa Calculi* de 1679, Leibniz atribui números a conceitos, de modo a obter uma representação completa da silogística aristotélica, o

que pode ser visto como uma concretização parcial de suas idéias. Este cálculo é estudado, e sua completude demonstrada, em [7].

Contudo, Leibniz nunca conseguiu realizar a contento sua proposta da *characteristica universalis*, nem do seu *calculus ratiocinator*, para os quais tinha expectativas altamente ambiciosas (cf. [30] p. 225), e no fim da vida veio a reconsiderar seu otimismo (cf. [15] p. 118 e [30] p. 656).

Coube a George Boole continuar o projeto de codificar as "leis do pensamento", apesar de aparentemente não conhecer o trabalho de Leibniz. Em 1847, publicou o *Mathematical Analysis of Logic* [5], em que expõe suas idéias sobre um cálculo puramente algébrico do raciocínio. Mais tarde, porém, declarou o *Analysis* um trabalho imaturo, e em 1854 publicou o influente *An Investigation of the Laws of Thought* [6], que afirmou ser o resultado final de sua pesquisa. Nesta obra, Boole explora profundamente as relações entre álgebra e lógica; mais precisamente, entre as regras de manipulação de símbolos algébricos, e as regras de certos símbolos que Boole considerava adequados para a representação do pensamento lógico.

Em particular, Boole associa símbolos, como x, y e z, a conceitos comumente expressos por adjetivos ou substantivos, por exemplo "x =cachorro" e "y =pequeno". Em seguida, define operações que constroem novos conceitos a partir dos anteriores. Uma delas é a *escolha*, ou *delimitação*: por exemplo, escolher tudo o que é pequeno dentre os cachorros define o conceito de "cachorro pequeno". Boole simboliza esta operação por xy. A outra operação é a *aglomeração*, ou *disjunção*, simbolizada por $x + y$, que define o conceito de "estar em alguma das classes x ou y". No exemplo, $x + y$ representa "ser cachorro ou ser pequeno". Boole demonstra que estas operações possuem uma série de propriedades familiares da aritmética, de modo que expressões envolvendo-as podem ser somadas, subtraídas, multiplicadas e até divididas, quase identicamente ao que ocorre com expressões numéricas. Para reforçar a analogia algébrica, Boole associa ao símbolo 1 a classe de todas as coisas, e ao 0 a classe vazia, de modo que, para qualquer x, $1x = x1 = x$, $0x = x0 = 0$, $x + 0 = 0 + x = x$ e $x + 1 = 1 + x = 1$.

O sistema exposto no *Laws of Thought* mostrou-se extremamente frutífero: hoje, Boole é considerado o pai da lógica algébrica (cf. [1]), tendo seu trabalho reconhecido por lógicos como Augustus de Morgan e Bertrand Russell. Contudo, a clareza e o rigor de seus métodos foram criticados por Dummett em [19], e, de fato, algumas de suas manipulações parecem justificar-se apenas por resolverem os problemas para os quais foram desenvolvidas. Até Corcoran, em seu [14], p. 285, julga que "Aristóteles parece superior a Boole, e mais próximo do pensamento contemporâneo".

Os métodos um pouco *ad hoc* de Boole se devem, pelo menos em parte, a certas limitações técnicas do sistema algébrico empregado. Do ponto de vista contemporâneo, estas limitações podem ser explicadas em termos da deficiência de estrutura da lógica booleana: ela não comporta inversos aditivos nem multiplicativos, e não obedece nenhuma lei do cancelamento satisfatoriamente geral. Em suma, porque a álgebra formulada por Boole não constitui um *corpo*.

Um possível caminho para a solução destes problemas, que traz consigo a

possibilidade de generalização dos métodos para lógicas multivalentes, é sugerido por um teorema de Boole, a chamada *lei indicial* [6], que pode ser enunciada como segue: escolher todos os objetos que são x, dentre aqueles que são x, define a própria classe x. Ilustrando o poder do *calculus ratiocinator* vislumbrado por Leibniz, esta afirmação é muito mais clara em símbolos: $xx = x^{\mathbf{2}} = x$. Boole considerava esta equação absolutamente central, chegando a declarar, ainda no *Laws of Thought*, que "uma lei fundamental da Metafísica é mera conseqüência de uma lei do pensamento".

Se é estipulado que as operações $+, \cdot$ de uma lógica formam um corpo, e que vale $x^2 = x$, demonstra-se que o conjunto dos valores-verdade desta lógica é o corpo de dois elementos: em linguagem algébrica contemporânea, o corpo de Galois $GF(2)$. A partir daí, uma generalização óbvia da lei indicial para $x^n = x$ conduz aos corpos $GF(p^\alpha)$, cujos elementos podem ser vistos como os valores-verdade das lógicas multivalentes, e proposições passam a ser representadas por polinômios sobre corpos finitos. É precisamente isto que os trabalhos [11, 12] de Carnielli vêm desenvolvendo nos últimos anos.

Pretendendo investigar leis combinatórias generalizadas para lógicas não-clássicas, [9] examinou estas questões para lógicas finitamente valoradas. Mais recentemente, trabalhos relacionados exploraram técnicas algébricas em conexão com complexidade de provas [13] e automatização de provas [36]. Porém, não há notícia do uso extenso de polinômios sobre corpos finitos, nem da extensão do método para todas as lógicas finitamente valoradas, tampouco infinitamente valoradas ou de primeira ordem.

Os artigos de Carnielli [11, 12] tentam abrir caminho nesta direção, apresentando um método geral de construir "cálculos polinomiais" para lógicas finitamente valoradas, e demonstrando resultados fundamentais como a correção e completude destes cálculos. A título de ilustração, diversas lógicas bem conhecidas são tratadas desta forma, incluindo as lógicas multivalentes de Post, Lukasiewicz e Sette. Os artigos ainda mostram como tratar certas lógicas não-finitamente valoradas pela polinomização: atribuindo-lhes semânticas não-verofuncionais bivaloradas (cf. [8]), que, por sua vez, são expressas por polinômios com "variáveis escondidas", isto é, com variáveis novas que não aparecem explicitamente nas fórmulas. Como exemplo, uma das principais lógicas da inconsistência formal (estudadas em [10]), **mbC**, é convenientemente polinomizada, e diversas de suas propriedades são deduzidas com o novo ferramental. Em [12] é explorada a possibilidade de se utilizar expansões polinomiais infinitas para tratar a quantificação em lógica de primeira ordem, mas apenas o fragmento monádico é considerado.

É nesse contexto que se insere este projeto. Pretendemos dar continuidade ao trabalho de Carnielli, estendendo nossa compreensão do método de polinomização a outras lógicas, descrevendo avanços no problema de polinomizar lógicas não-clássicas, e utilizando métodos algébricos para lançar novos olhares sobre resultados conhecidos da Lógica.

Caveat Emptor

Devido ao seu nome e contexto, seria razoável supor que o método de polinomização possuísse certas características e relações com outros conceitos e construções da Lógica. Cabe, talvez, citar algumas características e relações que o método de polinomização, ao menos em sua acepção atual, *não* possui, com o fim de esclarecer o escopo do trabalho e indicar possíveis direções de investigação.

Lógica Algébrica

O termo "polinomização" pode levantar questões sobre a relação do presente trabalho com a Lógica Algébrica. Esta disciplina tem origens no trabalho de Boole, DeMorgan, Peirce e Schröder, e atingiu sua expressão moderna com Łoś, Susko, Blok, e Tarski e seus estudantes, notoriamente Pigozzi. Seu ponto de partida foi a álgebra booleana, que em certo sentido preciso expressa correta e completamente o Cálculo Proposicional Clássico; o grande marco da modernização foi a descoberta das álgebras cilíndricas, que desempenham papel análogo para a Lógica de Primeira Ordem.

Grosso modo, a Lógica Algébrica busca traduções de sistemas lógicos para sistemas algébricos em que está definida uma relação de ordem; esta relação é a principal ferramenta técnica para a expressão algébrica das relações lógicas. Uma instância típica deste fenômeno seria associar sentenças α, β a elementos $\hat{\alpha}, \hat{\beta}$ de um conjunto parcialmente ordenado $(P, \leq)$ de tal modo que $\hat{\alpha} \leq \hat{\beta}$ se, e somente se, α é conseqüência de β (no sentido do sistema lógica original).

O método de polinomização também opera certas traduções de proposições em objetos algébricos. Porém, conforme veremos no Capítulo 1, é essencial ao método que estes objetos sejam elementos de certas estruturas algébricas (corpos finitos) em que não é possível definir ordens parciais que respeitem as demais operações já presentes. Por isso, a polinomização de lógicas está, ao menos no escopo deste trabalho, em um âmbito distinto da Lógica Algébrica. Seria interessante saber se há alguma maneira de superar este obstáculo técnico, e aproximar as duas.

Operações Sobre Lógicas

O termo "polinomização", por expressar uma operação que se pode realizar sobre lógicas, pode levantar questões sobre sua relação com as diversas transformações de sistemas lógicos que contempla a literatura, como fusão de lógicas, produtos, fibring, temporalização, paraconsistentização, fuzzificação etc.

A característica fundamental comum a todos estes é o mesmo: dada uma lógica, pretende-se mudá-la de alguma forma; às vezes por combinação com outras lógicas, às vezes por adição de novos recursos expressivos. Igualmente fundamental ao entendimento do método de polinomização, ao menos em sua primeira acepção, formulada por Carnielli e elaborada neste trabalho, é a percepção de que se trata de um método para expressar de modo diferente *a mesma* lógica. A Álgebra de Boole expressa exatamente o Cálculo Proposicional Clássico, e

esta é a sua força; ela não vem para adicionar modalidades, ou qualquer outro recurso expressivo adicional. Analogamente, a polinomização pretende facilitar o trabalho com exemplos particulares de lógicas, e trazer ferramentas algébricas ao estudo geral das lógicas "tais como são".

Naturalmente, seria interessante investigar como muda a polinomização de uma lógica, quando esta é paraconsistentizada, temporalizada etc. Este apenas não é o escopo do presente trabalho.

Polinômios e Matemática

Outra linha de justificativa para o uso de polinômios advém do quanto têm se mostrado frutíferos em Matemática. A seguir, ilustramos este fenômeno com três exemplos, listados em ordem cronológica. O fio condutor que passa por todos os três é o mesmo que dá ao conceito de polinômio sua força e utilidade: o fato de que uma expressão polinomial, especialmente sobre um corpo adequado, contém uma quantidade "finita" de informação.

No primeiro exemplo, esta heurística se manifesta na facilidade de calcular raízes de um polinômio. No segundo, usa-se que, sobre um corpo algebricamente fechado, a informação "finita" de um polinômio vem dada de forma explícita como um conjunto finito de raízes. Por fim, no terceiro exemplo, a expressão de uma série infinita em termos de um polinômio é um reflexo do fato de que a série é determinada por uma "quantidade finita" de informação — seus dois primeiros termos.

Teorema de Taylor

O primeiro emprego do Teorema de Taylor, feito por seu próprio autor, foi como mecanismo calculacional para encontrar raízes aproximadas de equações [20].

Seu método era mais ou menos o seguinte. Seja f uma função (evitamos escrever $f : \mathbb{R} \to \mathbb{R}$ porque o conceito de "conjunto dos números reais" ainda não estava formado), e a uma raiz aproximada da equação $f(x) = 0$. Considere o valor de f em a, bem como de suas derivadas, denotados $f(a), f'(a), f''(a)$ etc. Defina h por $x = a + h$, e escreva a série

$$f(x) = f(a+h) = f(a) + \frac{f'(a)}{1!}h + \frac{f''(a)}{2!}h^2 + \ldots$$

Agora despreze os termos de grau maior que 2 em h, obtendo uma equação aproximada

$$0 = f(a) + \frac{f'(a)}{1!}h + \frac{f''(a)}{2!}h^2$$

Resolva a equação quadrática para encontrar h. Obtém-se, assim, um novo valor $a' = a + h$ para a raiz aproximada, e pode-se aplicar o método novamente. É interessante notar que este método, apesar de simples, está resguardado de parar abruptamente (obtendo $h = 0$ em alguma iteração), pois isto só pode

ocorrer se o termo constante da expressão quadrática acima, $f(a)$, for nulo — isto é, se a raiz aproximada a for uma raiz de fato.

Álgebra Linear

Um emprego particularmente elegante de polinômios na vertente "conceitual" é aquele iniciado por Euler na Álgebra Linear.

Euler estava interessado em transformações lineares $\mathbb{R}^n \to \mathbb{R}^n$, ou, na linguagem da época, em matrizes quadradas com entradas complexas; em particular, no problema de encontrar *autovalores* de uma matriz A, isto é, um número λ tal que exista algum $x \in \mathbb{R}^n$ (um "vetor-coluna") não-nulo com a propriedade $Ax = \lambda x$. Este problema surgiu durante investigações sobre a rotação de corpos rígidos, e tem encontrado as mais diversas aplicações desde então: do algoritmo de busca do Google às equações da mecânica quântica.

Se denotamos por I a transformação identidade (i.e. a matriz com todos os elementos da diagonal iguais a 1, e os demais iguais a 0), segue das propriedades básicas da álgebra matricial que a equação $Ax = \lambda x$ é equivalente a $(A - \lambda I)x = 0$. Se desejamos que exista uma solução $x \neq 0$, então a matriz $A - \lambda I$ não deve ser inversível, e da Álgebra Linear sabemos que isto equivale a $\det(A - \lambda I) = 0$. Deste modo, uma pergunta aparentemente difícil sobre objetos complexos foi reduzida a uma questão polinomial, pois, pela conhecida expressão do determinante, $\det(A - \lambda I)$ é um polinômio em λ, chamado *polinômio característico de* A e denotado $p_A(\lambda)$.

Através desta redução, várias técnicas, como a de Taylor, podem ser de ajuda. Mais ainda, obtemos resultados puramente teóricos, nada triviais, como:

Teorema 1 *Qualquer matriz possui apenas um número finito de autovalores.*

Teorema 2 *Se permitimos autovalores complexos, então toda matriz possui ao menos um autovalor (pelo Teorema Fundamental da Álgebra).*

Matemáticos do século XIX foram ainda mais longe, e consideraram polinômios em que as variáveis são não apenas números, mas também matrizes. O produto e soma matriciais levam a uma definição natural para este caso.

Um dos teoremas centrais da Álgebra Linear, o chamado Teorema de Cayley-Hamilton, lança mão do segundo fato acima, potencializado por um processo de indução na dimensão do espaço, para mostrar que qualquer matriz A é raiz de seu polinômio característico, se estendemos a noção de polinômio conforme descrito acima: $p_A(A) = 0$. Daí obtemos mais um resultado teórico:

Teorema 3 *Se o único autovalor complexo de A é 0, então A é nilpotente; em particular, $A^n = 0$.*

Combinatória

Ilustramos brevemente o uso de séries infinitas em Combinatória, para que o leitor esteja a par de suas características gerais e possa admirar a variedade de aplicações dos conceitos de polinômio e série.

A bem conhecida seqüência de Fibonacci, $F_0, F_1, F_2, \ldots$, é definida do seguinte modo: $F_0 = F_1 = 1$ e, para $n \geq 2$, $F_n = F_{n-1} + F_{n-2}$. Os primeiros termos são $1, 1, 2, 3, 5, 8, \ldots$. Um problema clássico é encontrar uma fórmula fechada para o n-ésimo termo, F_n; entendemos "fórmula fechada" como uma expressão envolvendo apenas somas, produtos, exponenciação e constantes, como 2^n ou $3 + n^2$.

Uma técnica introduzida por Lagrange no século XVIII, à primeira vista surpreendente, é considerar a série infinita

$$F(x) = \sum_{n \geq 0} F_n x^n$$

Primeiro, notamos que uma prova indutiva simples mostra que $F_n \leq 2^n$, de modo que a série acima converge para $0 \leq x < 1/2$. Aplicando a relação definidora da seqüência de Fibonacci, temos

$$\begin{aligned}
F(x) &= \sum_{n \geq 0} F_n x^n = 1 + x + \sum_{n \geq 2} (F_{n-1} + F_{n-2}) x^n \\
&= 1 + x + \sum_{n \geq 2} F_{n-1} x^n + \sum_{n \geq 2} F_{n-2} x^n \\
&= 1 + x + x \cdot \sum_{n \geq 2} F_{n-1} x^{n-1} + x^2 \cdot \sum_{n \geq 2} F_{n-2} x^{n-2} \\
&= 1 + x + x \cdot \sum_{n \geq 1} F_n x^n + x^2 \cdot \sum_{n \geq 0} F_n x^n \\
&= 1 + x \cdot \sum_{n \geq 0} F_n x^n + x^2 \cdot \sum_{n \geq 0} F_n x^n \\
&= 1 + x \cdot F(x) + x^2 \cdot F(x)
\end{aligned}$$

donde deduzimos que

$$F(x) = \frac{1}{1 - x - x^2}$$

Esta curiosa expressão pode parecer tão enigmática quanto a seqüência original, mas estamos na posição vantajosa de ter trocado um objeto discreto por outro, analítico, sobre o qual talvez possamos aplicar técnicas do Cálculo. Porém, neste caso em particular algo ainda mais simples é suficiente. Pela fórmula quadrática, sabemos que o polinômio $1 - x - x^2$ possui duas raízes reais, que denotaremos $\alpha = \frac{-1+\sqrt{5}}{2}$ e $-\frac{1}{\alpha} == \frac{-1-\sqrt{5}}{2}$. Escrevendo $1 - x - x^2 = (x - \alpha)(x + \frac{1}{\alpha})$, temos

$$F(x) = \frac{1}{\sqrt{5}} \left(\frac{1}{\alpha - x} + \frac{1}{x + \frac{1}{\alpha}} \right) = \frac{1}{\sqrt{5}} \left(\frac{1}{\alpha} \frac{1}{1 - \frac{x}{\alpha}} + \alpha \frac{1}{1 + \alpha x} \right)$$

Agora, tendo em mente que, para $0 \leq x < 1/2$, tanto $\frac{x}{\alpha}$ quanto αx são menores que 1, usamos a fórmula da série geométrica para escrever

$$\begin{aligned}\frac{1}{1-\frac{x}{\alpha}} &= 1+\frac{x}{\alpha}+\left(\frac{x}{\alpha}\right)^2+\left(\frac{x}{\alpha}\right)^3+\ldots \\ \frac{1}{1+\alpha x} &= 1+\alpha x+(\alpha x)^2+(\alpha x)^3+\ldots\end{aligned}$$

Agrupando os termos de mesmo grau, obtemos

$$F(x)=\sum_{n\geq 0}\frac{1}{\sqrt{5}}\left(\frac{1}{\alpha^{n+1}}+\alpha^{n+1}\right)x^n$$

Comparando com a definição inicial, obtemos a seguinte fórmula, que dificilmente seria evidente à primeira vista:

$$F_n=\frac{1}{\sqrt{5}}\left(\frac{1}{\alpha^{n+1}}+\alpha^{n+1}\right).$$

Capítulo 1

Lógicas e Polinômios

Neste capítulo, iniciamos nossa investigação sobre a possibilidade de expressar vários objetos de estudo da Lógica Proposicional em termos de polinômios. As primeiras tarefas incluem selecionar uma classe de lógicas com a qual trabalhar, e explorar os vários domínios sobre os quais se pode definir polinômios.

A finalidade é guiar os esforços do próximo capítulo, em que efetivamente polinomizamos lógicas e tentamos determinar se a polinomização traz benefícios: quer calculacionais, permitindo computações rápidas em várias lógicas; quer conceituais, proporcionando demonstrações mais simples, ou ao menos novas, de fatos já conhecidos da Lógica.

1.1 Lógicas Proposicionais

Iniciamos este capítulo com uma exploração dos possíveis tipos de polinômio que podem se adequar a esta tarefa, estendendo as investigações de Carnielli [11]. Para tanto, delimitamos uma classe de lógicas proposicionais, ditas *tarskianas*, com as seguintes definições [10].

Definição 1 *Seja $\mathcal{V}$ um conjunto enumerável e $\mathcal{C}$ um conjunto de símbolos de função. Para cada n natural, denotamos por $\mathcal{C}_n$ o subconjunto dos símbolos de aridade n. A* álgebra de fórmulas *sobre $(\mathcal{V}, \mathcal{C})$ é a álgebra gerada livremente pelos símbolos de $\mathcal{C}$, usando os elementos de $\mathcal{V}$ como variáveis. Nos referiremos aos elementos de $\mathcal{C}$ como "conectivos".*

Usaremos o termo *álgebra de fórmulas sobre $\mathcal{C}$* quando não for importante qual é o conjunto de variáveis, ou apenas *álgebra de fórmulas* quando não for necessário especificar sequer os símbolos de função. De fato, será bastante raro atribuirmos alguma importância a um conjunto específico de variáveis: dados dois conjuntos enumeráveis $\mathcal{V}, \mathcal{V}'$, qualquer bijeção $\mathcal{V} \to \mathcal{V}'$ pode ser estendida, de modo canônico, a um isomorfismo entre as álgebras de fórmulas sobre $(\mathcal{V}, \mathcal{C})$ e sobre $(\mathcal{V}', \mathcal{C})$.

Definição 2 *Uma* lógica proposicional $\mathcal{L}$ *é composta por dois objetos:*

- *Uma álgebra de fórmulas sobre um conjunto de conectivos* $\mathcal{C}_\mathcal{L}$*. Denotamos esta álgebra por* $\mathrm{For}(\mathcal{L})$*, e o conjunto subjacente de variáveis por* $\mathcal{V}_\mathcal{L}$*;*
- *Uma relação de conseqüência* $\vdash_\mathcal{L} \subseteq 2^{\mathrm{For}(\mathcal{L})} \times \mathrm{For}(\mathcal{L})$*, denominada "de conseqüência", ou "deduz".*

A lógica é tarskiana *se a relação* $\vdash_\mathcal{L}$ *satisfaz às seguintes propriedades, para quaisquer* $\alpha, \beta \in \mathrm{For}(\mathcal{L})$ *e* $\Gamma, \Delta \subseteq \mathrm{For}(\mathcal{L})$*:*

1. *Se* $\alpha \in \Gamma$*, então* $\Gamma \vdash_\mathcal{L} \alpha$*. (Reflexividade)*
2. *Se* $\Gamma \vdash_\mathcal{L} \alpha$ *e* $\Gamma \subseteq \Delta$*, então* $\Delta \vdash_\mathcal{L} \alpha$*. (Monotonicidade)*
3. *Se* $\Gamma, \alpha \vdash_\mathcal{L} \beta$ *e* $\Delta \vdash_\mathcal{L} \alpha$*, então* $\Gamma, \Delta \vdash_\mathcal{L} \beta$*. (Corte)*

Estaremos interessados principalmente nas lógicas proposicionais tarskianas. Nesta classe de lógicas, a relação de conseqüência goza de propriedades que facilitam seu manejo, e que comumente se associam ao processo de deduzir certas fórmulas tendo outras como hipótese. Ao mesmo tempo, a classe é suficientemente geral para incluir a maior parte das lógicas geralmente estudadas.

Primeiros exemplos incluem a Lógica Proposicional Clássica, com conjuntos de conectivos unários $\{\neg\}$ e de conectivos binários $\{\vee, \wedge, \rightarrow\}$, assim como a Lógica Proposicional Intuicionista. De fato, as lógicas tarskianas incluem duas grandes famílias de lógicas proposicionais de interesse: aquelas em que a relação de conseqüência é definida sintaticamente, por meio de um cálculo axiomático hilbertiano, e aquelas em que a relação de conseqüência é determinada semanticamente, por meio de interpretações das variáveis e valores de verdade distinguidos.

Tornamos esta afirmação precisa em dois passos: com a Definição 5 e o Teorema 3, e mais tarde com a Definição 8 e o Teorema 4.

1.1.1 Relações de Conseqüência Sintáticas

A seguir, tentaremos dar uma definição algébrico-combinatória da bem conhecida idéia de um sistema axiomático hilbertiano. Ao formulá-la, priorizamos maior generalidade sobre perspicuidade, pois trata-se de um conceito suficientemente familiar à Lógica para que seu contexto e intenção sejam claros. Além disso, com esta linha de raciocínio visamos estabelecer a adequação da Definição 2, pelo que se faz importante o maior nível de generalidade.

Definição 3 *Seja* $\mathcal{F}$ *uma álgebra de fórmulas. Chamamos uma relação* $\vdash^0 \subseteq 2^\mathcal{F} \times \mathcal{F}$ *de* regras de inferência *para* $\mathcal{F}$ *se o domínio de* $\vdash^0$ *consiste apenas de subconjuntos finitos de* $\mathcal{F}$*. Isto é, para cada* $\Gamma \subseteq \mathcal{F}$*, existe* $\alpha \in \mathcal{F}$ *tal que* $\Gamma \vdash^0 \alpha$ *somente se* Γ *é finito.*

A definição acima tem em vista o conceito de demonstração sintática, que formulamos em grande generalidade a seguir. Vale notar que a maior parte das lógicas tradicionalmente estudadas conta com vários mecanismos sintáticos denominados "regras de inferência", e não apenas um. Por exemplo, em apresentações usuais, a Lógica de Primeira Ordem conta com as regras de Modus Ponens e Generalização.

A separação dos vários modos de inferência é muito importante em um contexto filosófico, ou quando as propriedades de uma lógica em particular estão em questão. Porém, para o presente fim, que é estabelecer uma propriedade de *todas* as lógicas com regras de inferência, essa distinção não é importante. Se são dadas várias regras de inferência $\vdash^0_1, \ldots, \vdash^0_N \subseteq 2^{\mathcal{F}} \times \mathcal{F}$, definimos $\vdash^0 = \bigcup_{1 \leq i \leq N} \vdash^0_i$ e trabalhamos com esta única regra de inferência. Está claro que $\Gamma \vdash^0 \alpha$ se, e somente se, ocorre ao menos um dos $\Gamma \vdash^0_i \alpha$. Isto será suficiente para as aplicações a seguir. (Note que a mesma construção vale para qualquer número infinito de regras de inferência.)

Definição 4 *Seja $\mathcal{F}$ uma álgebra de fórmulas, e $\vdash^0$ regras de inferência para $\mathcal{F}$. Se $\Gamma \subseteq \mathcal{F}$, uma seqüência $\pi = \alpha_1, \ldots, \alpha_n$ é uma* prova a partir de Γ *se, e somente se, para cada $i \in 1, \ldots, n$, ao menos uma das seguintes condições é satisfeita:*

1. *$\alpha_i \in \Gamma$; ou*
2. *Existe $S \subseteq \{\alpha_1, \ldots, \alpha_{i-1}\}$ tal que $S \vdash^0 \alpha_i$.*

No primeiro caso, dizemos que α_i é uma "hipótese", e, no segundo, que α_i "decorre de S pelas regras de inferência". (Note que estes casos não são mutuamente exclusivos.)

Se $\alpha_1, \ldots, \alpha_n$ é uma prova a partir de Γ, dizemos que é uma prova de α a partir de Γ *se $\alpha_n = \alpha$, e que $|\pi| = n$ é o* comprimento *da prova. Finalmente, denotamos por $\pi(i)$ a i-ésima fórmula de π.*

Aqui também divergimos um pouco da tradição com a terminologia "prova a partir de Γ", sem especificar de que proposição, exatamente, é a prova. Fazemos isto porque a propriedade "ser prova a partir de Γ" é muito flexível: como veremos mais adiante, podemos "interromper" uma tal prova π, isto é, considerar apenas um trecho inicial $\pi' = \pi(1), \ldots, \pi(k)$, e este ainda é uma prova a partir de Γ. Porém, fato análogo não se verifica para a propriedade mais restrita "ser uma prova de α a partir de Γ", uma vez que a última fórmula do trecho π' é, em geral, diferente da última fórmula da prova original π. Esta flexibilidade será importante nos argumentos indutivos a seguir.

Com estas preliminares, estamos prontos para definir as lógicas hilbertianas, em que a relação de conseqüência se apóia na existência de demonstrações sintáticas. O único ingrediente faltante até agora é a noção de axioma, que é facilmente incluído.

Definição 5 *Uma lógica proposicional $\mathcal{L}$ é dita* hilbertiana *se há um subconjunto $\mathcal{A} \subseteq \mathrm{For}(\mathcal{L})$ (os "axiomas") e regras de inferência $\vdash^0_{\mathcal{L}}$ para $\mathrm{For}(\mathcal{L})$, que, juntos, caracterizam inteiramente a relação de conseqüência $\vdash_{\mathcal{L}}$ de $\mathcal{L}$, da seguinte*

forma. Para quaisquer $\Gamma \subseteq \mathrm{For}(\mathcal{L})$ *e* $\alpha \in \mathrm{For}(\mathcal{L})$, *tem-se* $\Gamma \vdash_{\mathcal{L}} \alpha$ *se, e somente se, há uma prova de* α *a partir de* $\Gamma \cup \mathcal{A}$.

A relevância histórica dos cálculos axiomáticos é bem conhecida [34], e a maior parte das lógicas estudadas atualmente possui uma apresentação em termos hilbertianos [10, 21]. Mesmo as lógicas não-proposicionais, via de regra, contam com um sistema axiomático fortemente análogo ao da Definição 5. Por isso, é de interesse estabelecer o Teorema 3.

Resumimos algumas propriedades simples de provas para uso futuro com dois Lemas. O primeiro é uma expressão recursiva da propriedade "ser uma prova", e enfatiza o caráter "local" destas, mostrando que é possível encurtar ou prolongar uma prova (de modo a manter o status de prova) sem muitas dificuldades.

Lema 1 *Seja* $\mathcal{F}$ *uma álgebra de fórmulas com regras de inferência* $\vdash^0$. *Então* $\pi = \alpha_1, \ldots, \alpha_n, \alpha_{n+1}$ *é uma prova a partir de* Γ *se, e somente se,* $\pi' = \alpha_1, \ldots, \alpha_n$ *é uma prova a partir de* Γ, *e* α_{n+1} *é um elemento de* Γ *ou decorre de algum* $S \subseteq \{\alpha_1, \ldots, \alpha_n\}$ *por* $\vdash^0$.

Demonstração. Se π é uma prova a partir de Γ, então, para cada $i \in \{1, \ldots, n+1\}$, a fórmula α_i satisfaz $\alpha_i \in \Gamma$ ou há $S \subseteq \{\alpha_1, \ldots, \alpha_{i-1}\}$ tal que $S \vdash^0 \alpha_i$. Em particular, isto vale para cada $i \in \{1, \ldots, n\}$, mostrando que π' é uma prova a partir de Γ.

Reciprocamente, π é uma prova a partir de Γ se as condições acima se verificam para $i \in \{1, \ldots, n+1\}$. O fato de π' ser uma prova a partir de Γ garante que valem para $i \in \{1, \ldots, n\}$, e resta checá-las para $i = n+1$; mas é isto que diz o resto da hipótese. □

Lema 2 *Seja* $\mathcal{F}$ *uma álgebra de fórmulas com regras de inferência* $\vdash^0$ *e* $\pi_1, \ldots, \pi_n$ *provas a partir de* $\Gamma_1, \ldots, \Gamma_n$ *respectivamente. A* concatenação *destas provas,* $\pi = \pi_1 \circ \ldots \circ \pi_n = \gamma_1, \ldots, \gamma_r$, *definida por por* $\gamma_j = \pi_1(j)$ *para* $1 \leq j \leq |\pi_1|$ *e* $\gamma_{|\pi_1|+\ldots+|\pi_k|+j} = \pi_{k+1}(j)$ *para* $|\pi_1| + \ldots + |\pi_k| < j \leq |\pi_1| + \ldots + |\pi_k| + |\pi_{k+1}|$, *é uma prova a partir de* $\Gamma_1 \cup \ldots \cup \Gamma_n$.

Demonstração. Está claro que, se estabelecermos o resultado para $n = 2$, o caso geral seguirá por indução. Sejam então π_1, π_2 provas a partir de Γ_1, Γ_2 respectivamente; procederemos por indução no comprimento de π_2.

No caso em que $|\pi_2| = 1$, π_2 consiste de uma única fórmula β, que deve então ser elemento de Γ_2 ou decorrer do conjunto vazio por $\vdash^0$. Em todo caso, $\pi_1 \circ \beta$ é uma prova a partir de $\Gamma_1 \cup \Gamma_2$, pelo Lema 1. Agora, se $|\pi_2| = k+1$, temos $|\pi_2| = \beta_1, \ldots, \beta_{k+1}$; ponha $\pi_2' = \beta_1, \ldots, \beta_k$, de comprimento k. Novamente pelo Lema 1, π_2' é uma prova a partir de Γ_2. Por hipótese de indução, $\pi_1 \circ \pi_2'$ é uma prova a partir de $\Gamma_1 \cup \Gamma_2$. Finalmente, o caso-base diz que $(\pi_1 \circ \pi_2') \circ \beta_{k+1} = \pi_1 \circ (\pi_2' \circ \beta_{k+1}) = \pi_1 \circ \pi_2$ é uma prova a partir de $\Gamma_1 \cup \Gamma_2$. □

Destes Lemas segue facilmente o resultado central desta subseção.

Teorema 3 *Toda lógica hilbertiana é também tarskiana.*

Demonstração. Dada uma lógica hilbertiana $\mathcal{L}$, vamos estabelecer as propriedades 1–3 da Definição 2. A primeira é imediata: se $\alpha \in \Gamma$, então a seqüência unitária α estabelece que $\Gamma \vdash_{\mathcal{L}} \alpha$.

A segunda também é simples: se $\Gamma \vdash_{\mathcal{L}} \alpha$, então há uma prova $\alpha_1, \ldots, \alpha_n = \alpha$ onde cada α_i pertence a $\mathcal{A} \cup \Gamma$ ou decorre de algum $S \subseteq \alpha_1, \ldots, \alpha_{i-1}$ pelas regras de inferência. Agora, se $\Gamma \subseteq \Delta$, então $\mathcal{A} \cup \Gamma \subseteq \mathcal{A} \cup \Delta$, de modo que cada α_i pertence a $\mathcal{A} \cup \Delta$ ou decorre de algum $S \subseteq \alpha_1, \ldots, \alpha_{i-1}$ pelas regras de inferência. Ou seja, $\alpha_1, \ldots, \alpha_n$ também é uma prova a partir de $\mathcal{A} \cup \Delta$. Segue que $\Delta \vdash_{\mathcal{L}} \alpha$.

Quanto à terceira propriedade, suponha que $\Gamma, \alpha \vdash_{\mathcal{L}} \beta$, com prova $\beta_1, \ldots \beta_m = \beta$, e também $\Delta \vdash_{\mathcal{L}} \alpha$, com prova $\alpha_1, \ldots, \alpha_n = \alpha$. Indutivamente, construiremos uma prova de cada β_i a partir de $\mathcal{A} \cup (\Gamma \cup \Delta)$, o que nos levará a uma prova de $\beta_m = \beta$ a partir deste mesmo conjunto.

Seja $k \leq m$ e suponha que, para cada $i < k$, temos uma prova π_i de β_i a partir de $\mathcal{A} \cup (\Gamma \cup \Delta)$. Vejamos como construir uma prova de β_k com as mesmas hipóteses. Se $\beta_k = \alpha$, basta reproduzir a prova $\alpha_1, \ldots, \alpha_n$ a partir de $\mathcal{A} \cup \Delta$; se $\beta_k \in \Gamma$, a seqüência unitária β_k constitui uma prova a partir de $\mathcal{A} \cup (\Gamma \cup \Delta)$. Por fim, se β_k decorre de algum $S \subseteq \{\beta_1, \ldots, \beta_{k-1}\}$ pelas regras de inferência, consideramos a concatenação $\pi = \pi_1 \circ \ldots \circ \pi_{k-1}$, que é uma prova a partir de $\mathcal{A} \cup (\Gamma \cup \Delta)$ pelo Lema 2. Agora, os $\beta_1, \ldots, \beta_{k-1}$ são as últimas fórmulas de $\pi_1, \ldots, \pi_{k-1}$, respectivamente. Portanto, S é um subconjunto das fórmulas que aparecem em π, e, como β_k segue de S por $\vdash^0$, a seqüência $\pi \circ \beta_k$ é uma prova a partir de $\mathcal{A} \cup (\Gamma \cup \Delta)$.

Assim, estabelecemos que $\Gamma, \Delta \vdash_{\mathcal{L}} \beta$, como desejávamos. □

O Teorema 3 mostra que a definição de lógica tarskiana inclui as lógicas hilbertianas, de caráter sintático. A seguir, concentramo-nos na contrapartida destas.

1.1.2 Relações de Conseqüência Semânticas

Outro importante método de construção de lógicas proposicionais é fortemente relacionado ao ferramental da Teoria de Modelos. Este paradigma "semântico" de construção de lógicas também se vê incluído na definição de lógica tarskiana.

Para esclarecer este ponto, e em analogia com a subseção anterior, formularemos uma definição que visa englobar totalmente o método semântico, ou ao menos o que suas diversas encarnações têm em comum. Esta grande generalidade obscurece a motivação lógica do viés semântico, mas oferece a possibilidade de demonstrar, de uma só vez, que lógicas definidas segundo um mecanismo muito geral são necessariamente tarskianas.

Definição 6 *Seja $\mathcal{F}$ uma álgebra de fórmulas sobre os conectivos $\mathcal{C}$, e seja A um conjunto. Uma* interpretação de $\mathcal{F}$ sobre A *é uma seqüência de funções $\phi = (\phi_n)_{n \in \mathbb{N}}$, onde cada ϕ_n leva $\mathcal{C}_n$ em A^{A^n}. Ou seja, ϕ atribui, a cada símbolo de função n-ário, uma operação n-ária sobre A.*

Dada uma interpretação ϕ dos conectivos de uma álgebra de fórmulas em termos de operações sobre um conjunto, há uma única extensão $\overline{\phi}$ a toda $\mathcal{F}$, que

associa a cada fórmula uma função. Em geral, a uma fórmula em que aparecem n variáveis corresponde uma função $A^{A^n} \to A$. Por abuso de notação, não distinguiremos entre ϕ e $\overline{\phi}$.

Resta especificar o papel das variáveis.

Definição 7 *Seja $\mathcal{F}$ uma álgebra de fórmulas sobre as variáveis $\mathcal{V}$ e conectivos $\mathcal{C}$, A um conjunto, e ϕ uma interpretação de $\mathcal{F}$ sobre A. Uma* valoração das variáveis de $\mathcal{F}$ sobre A *é qualquer função $\nu : \mathcal{V} \to A$. Uma* valoração das fórmulas de $\mathcal{F}$ sobre A, de acordo com ϕ, *é uma extensão de uma valoração das variáveis ν a uma aplicação $\nu_\phi : \mathrm{For}(\mathcal{L}) \to A$, do seguinte tipo. Para cada $n \in \mathbb{N}$, $c_n \in \mathcal{C}_n$ e fórmulas $\alpha_1, \ldots, \alpha_n$, tem-se $\nu_\phi(c_n(\alpha_1, \ldots, \alpha_n)) = \phi_n(c_n)(\nu_\phi(\alpha_1), \ldots, \nu_\phi(\alpha_n))$.*

Esta situação é completamente análoga àquela encontrada na Teoria de Modelos, onde recursos semelhantes são empregados para definir interpretações de teorias de primeira ordem. Na linguagem desta área, tanto a álgebra de fórmulas $\mathcal{F}$ quanto o conjunto A são $\mathcal{C}$-estruturas, e a exigência aqui é que ν_ϕ seja um morfismo de $\mathcal{C}$-estruturas.

As observações logo após a Definição 1 mostram que esta construção não resulta ambígua, e que há uma única ν_ϕ satisfazendo as condições estipuladas. Conforme se vê na Álgebra Universal, é característico de estruturas livremente geradas que sempre exista um único morfismo de álgebras estendendo uma função dada nos geradores [22, 2].

Porém, o enfoque da Lógica Proposicional sobre as variáveis de $\mathcal{V}$ é enquanto valores-verdade, e uma primeira tentativa, bastante difundida na literatura [8, 21], de levar isto em conta, é separar alguns valores-verdade como "distinguidos". O emprego desta noção é definir relações de conseqüência do seguinte tipo.

Definição 8 *Uma* semântica verofuncional *para uma lógica proposicional $\mathcal{L}$ é dada por um conjunto A, uma interpretação ϕ de $\mathrm{For}(\mathcal{L})$ sobre A, e um subconjunto $D \subseteq A$ (valores "distinguidos") que caracterizam a relação de conseqüência $\vdash_{\mathcal{L}}$ da seguinte forma. Para quaisquer $\Gamma \subseteq \mathrm{For}(\mathcal{L})$ e $\alpha \in \mathrm{For}(\mathcal{L})$, vale a relação $\Gamma \vdash_{\mathcal{L}} \alpha$ se, e somente se, para toda valoração das variáveis $\nu : \mathcal{V}_{\mathcal{L}} \to A$, tem-se que $\nu_\phi(\Gamma) \subseteq D$ implica $\nu_\phi(\alpha) \in D$. Se $\mathcal{L}$ admite semântica verofuncional, dizemos que é uma* lógica verofuncional.

Em muitos casos de interesse, as lógicas verofuncionais possuem semânticas particularmente simples. Dizemos que uma lógica verofuncional é *m-valorada* se possui uma semântica verofuncional sobre um conjunto com m elementos, e *finitamente valorada* se é m-valorada, para algum $m \in \mathbb{N}$. Esta terminologia é justificada pelo fato de que o conjunto particular A "não importa": se $A \xrightarrow{f} B$ é qualquer bijeção, uma semântica verofuncional ϕ_A de $\mathcal{L}$ sobre A induz naturalmente outra, ϕ_B, sobre B, como segue: se c é um conectivo n-ário, define-se $\phi_B(c)(b_1, \ldots, b_n) = f \circ \phi_A(f^{-1}(b_1), \ldots, f^{-1}(b_n))$, e o conjunto de valores-verdade distinguidos é $f(D)$.

Lógicas finitamente valoradas serão objeto de maior atenção quando viermos a discutir a polinomização.

De posse das definições acima, estamos prontos para enunciar e demonstrar o resultado central desta subseção.

Teorema 4 *Toda lógica verofuncional é tarskiana.*

Demonstração. Seja $\mathcal{L}$ uma lógica verofuncional com interpretação ϕ sobre um conjunto A, com valores distinguidos $D \subseteq A$. Estabeleceremos as propriedades 1–3 da Definição 2. Para tanto, sejam $\Gamma, \Delta \subseteq \mathrm{For}(\mathcal{L})$ e $\alpha, \beta \in \mathrm{For}(\mathcal{L})$, e ainda $\nu : \mathcal{V} \to A$ uma valoração arbitrária das variáveis.

A primeira propriedade é simples: se $\alpha \in \Gamma$, então $\nu_\phi(\alpha) \in \nu_\phi(\Gamma)$. Portanto, se $\nu_\phi(\Gamma) \subseteq D$, necessariamente $\nu_\phi(\alpha) \in D$, de modo que $\Gamma \vdash_{\mathcal{L}} \alpha$.

Para a segunda, note que, se $\Gamma \subseteq \Delta$, então $\nu_\phi(\Gamma) \subseteq \nu_\phi(\Delta)$. Deste modo, se $\nu_\phi(\Delta) \subseteq D$, então $\nu_\phi(\Gamma) \subseteq D$, e daí segue, por hipótese, que $\nu_\phi(\alpha) \in D$. Assim, $\Delta \vdash_{\mathcal{L}} \alpha$.

Por fim, suponha que $\Gamma, \alpha \vdash_{\mathcal{L}} \beta$, e que $\Delta \vdash_{\mathcal{L}} \alpha$. Se $\nu_\phi(\Gamma \cup \Delta) \subseteq D$, então $\nu_\phi(\Delta) \subseteq D$. Por hipótese, isto implica que $\nu_\phi(\alpha) \in D$. Agora, novamente por hipótese, resulta que $\nu_\phi(\beta) \in D$. Deste modo, $\Gamma, \Delta \vdash_{\mathcal{L}} \beta$. □

1.2 Polinômios

Tendo circunscrito o conceito de Lógica Proposicional empregado nesta Dissertação, e argumentado em favor da definição usada, passamos a investigar a possibilidade de expressar lógicas proposicionais em termos de polinômios.

Uma primeira tentativa é inspirada no enfoque semântico da subseção 1.2: dada uma lógica proposicional $\mathcal{L}$, buscamos não apenas uma interpretação ϕ de seus conectivos sobre um conjunto A, mas também que as imagens das funções ϕ_n sejam polinômios a n variáveis. Está claro que, para esta exigência fazer sentido, o conjunto A deve vir munido de alguma estrutura que permita falar de polinômios; no mínimo, duas operações binárias $+, \cdot : A \times A \to A$ que façam a parte da soma e do produto na acepção mais comum de polinômio.

Tendo em vista os dois fins deste trabalho — a simplificação de certas computações em lógicas proposicionais, e o emprego de técnicas e teoremas da Álgebra sobre problemas da Lógica — iremos focar nossa atenção sobre possibilidades cada vez mais restritas para o domínio A e as operações binárias definidas sobre este.

As primeiras restrições que faremos serão sobre as operações, e serão brandas: deixaremos de lado de nossas considerações as operações não-associativas e não-comutativas. De fato, a expressão comum de polinômios, algo do tipo $X^3 + X^2 + X$, tacitamente assume associatividade: do contrário, seria ambíguo se o termo X^3 se refere a $X \cdot (X \cdot X)$ ou a $(X \cdot X) \cdot X$. Similarmente, não fica claro em que ordem deveríamos realizar as somas.

Mais significativamente, a Álgebra das operações não-associativas encontra-se muito menos desenvolvida que sua contraparte associativa, fato devido, sem dúvida, à maior abundância de modelos de operações não-associativas, que acarreta a pobreza do conjunto de teoremas satisfeitos por todas tais operações. Se

visamos encontrar, no palácio da Álgebra, um cômodo satisfatório para problemas da Lógica, é aconselhável começar a busca pelos salões principais, construídos primeiro e laboriosamente decorados.

Por este motivo, elegemos também considerar principalmente operações comutativas. De fato, a Álgebra Comutativa vem sendo vigorosamente desenvolvida há mais de cem anos, em resposta às necessidades da Teoria Algébrica de Números e da Geometria Algébrica [22, 32]. Ademais, certas técnicas que empregaremos no Capítulo seguinte guardam relação com conceitos da Geometria Algébrica, o que reforça nossa escolha.

Com isto em mente, chegamos, pois, a uma primeira aproximação para os domínios sobre os quais consideraremos os polinômios. Trata-se de um tipo algébrico amplamente estudado e com bons teoremas de estrutura: o anel comutativo. Referimos o leitor ao Apêndice para as definições e pequenos resultados algébricos que não incluímos no corpo do texto.

Definição 9 *Seja A um conjunto e $+, \cdot : A \times A \to A$ operações binárias. Dizemos que $(A, +, \cdot)$ é um* anel *se: o par $(A, +)$ é um grupo comutativo; $\cdot$ é associativa; existe um elemento "identidade", $1 \in A$, tal que para todo $x \in A$, $x \cdot 1 = 1 \cdot x = x$; e valem as leis distributivas, isto é, para quaisquer $x, y, z \in A$ tem-se $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ e $(y + z) \cdot x = (y \cdot x) + (z \cdot x)$. Dizemos que $(A, +, \cdot)$ é um* anel comutativo *se $\cdot$ é comutativa.*

Se $(A, +, \cdot)$ e $(A', +', \cdot')$ são anéis, uma aplicação $f : A \to A'$ é um morfismo de anéis *se é um morfismo de grupos, $f(1_A) = 1_{A'}$, e ademais vale $f(x \cdot y) = f(x) \cdot' f(y)$ para quaisquer $x, y \in A$.*

(No resto da Dissertação, seguiremos as práticas usuais de ter a operação $\cdot$ com precedência sobre $+$, omitindo parênteses de acordo, e de nos referirmos ao próprio conjunto A como anel ou grupo quando as operações $+, \cdot$ estiverem claras do contexto ou não forem importantes. Também será comum suprimirmos o operador $\cdot$, em favor da concatenação de símbolos, isto é, escrever xy em lugar de $x \cdot y$.)

A inclusão dos elementos neutro e identidade, e de opostos, é feita em respeito à tradição algébrica, e por propriedades técnicas que eles trazem. De todo modo, há condições bem conhecidas, e relativamente brandas, sob as quais é possível acrescentar, de maneira canônica, o elemento neutro e opostos a uma estrutura algébrica que careça deles [22]. Por fim, a lei distributiva é a maneira usual de fazer a conexão entre as duas operações definidas sobre A. É a praxe matemática que, quando duas operações ou estruturas são impostas sobre um mesmo domínio, entre elas deve subsistir alguma forma de relação. Se na lógica não-polinomizada já se goza deste benefício — por exemplo, $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$ na Lógica Proposicional Clássica — não seria vantagem alguma abandoná-la na passagem para polinômios.

Há uma interação não-trivial entre um anel $(A, +, \cdot)$ e seu anel de endomorfismos como grupo comutativo $(A, +)$, dada pelo seguinte Lema.

Lema 5 *Seja A um anel e $\mathrm{End}(A)$ o anel de endomorfismos do grupo subjacente. Há um morfismo injetivo de anéis $A \overset{j}{\hookrightarrow} \mathrm{End}(A)$ definido por $j(a) = a\cdot$,*

multiplicação à esquerda por a. Em outras palavras, $j(a)$ age sobre A por $x \mapsto a \cdot x$.

Demonstração. Segue das leis distributivas para anéis que, para cada $a \in A$, o mapa $A \to A$ dado por $x \mapsto a \cdot x$ é um endomorfismo do grupo $(A, +)$, de modo que $j(a) \in \mathrm{End}(A)$. Ademais, para quaisquer $a, b, x \in A$ tem-se $j(a+b)(x) = (a+b) \cdot x = a \cdot x + b \cdot x = j(a)(x) + j(b)(x) = (j(a) + j(b))(x)$, bem como $j(a \cdot b)(x) = (a \cdot b) \cdot x = a \cdot (b \cdot x) = a \cdot (j(b)(x)) = j(a)(j(b)(x)) = j(a) \circ j(b)(x)$. Finalmente, $j(1)(x) = 1 \cdot x = x$, donde $j(1) = \mathrm{id}_A$, e portanto j é morfismo de anéis. A injetividade segue do fato de que, se $a \neq b$, então $j(a)(1) = a \cdot 1 = a \neq b = b \cdot 1 = j(b)(1)$, ou seja, os endomorfismos $j(a)$ e $j(b)$ diferem em ao menos um ponto. □

Mais tarde, usaremos a idéia de anel de endomorfismos para relacionar polinômios e os anéis sobre os quais são definidos.

1.2.1 Aspectos Formais

Sem mais delongas, damos agora a definição de polinômio que empregamos em toda a Dissertação, mesmo após subseqüentes restrições sobre os domínios e operações considerados.

Definição 10 *Seja $(A, +, \cdot)$ um anel comutativo, e $A^{<\omega}$ o conjunto de todas as seqüências (indexadas por $\mathbb{N}$) de A tais que apenas um número finito de seus elementos é diferente de 0. A* álgebra de polinômios sobre A *é a terna $(A^{<\omega}, +^{<\omega}, \cdot^{<\omega})$, onde $+^{<\omega}, \cdot^{<\omega} : A^{<\omega} \times A^{<\omega} \to A^{<\omega}$ são operações binárias definidas do seguinte modo: para quaisquer $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \in A^{<\omega}$,*

- $(a_n)_{n\in\mathbb{N}} +^{<\omega} (b_n)_{n\in\mathbb{N}} = (a_n + b_n)_{n\in\mathbb{N}}$;
- $(a_n) \cdot^{<\omega} (b_n)_{n\in\mathbb{N}} = (\sum_{i+j=n} a_i \cdot b_j)_{n\in\mathbb{N}}$.

O grau *de um polinômio é um a menos que a posição do maior termo não-nulo na seqüência que o define. (Por exemplo, o grau de $(0, 1, 0, 0, \ldots)$ é um.) Por convenção, o grau da seqüência nula é $-\infty$.*

É simples checar que, se $n_0 \in \mathbb{N}$ é tal que $n > n_0$ implica $a_n = b_n = 0$, então $n > n_0$ implica que $a_n + b_n = 0$, e também que $\sum_{i+j=2n} a_i b_j = 0$. Portanto, $A^{<\omega}$ é mesmo fechado sob estas duas operações.

Por abuso de notação, denotaremos as operações $+^{<\omega}, \cdot^{<\omega}$ pelos mesmos símbolos empregados para as operações sobre A. Também denotaremos os elementos de $A^{<\omega}$ da maneira mais usual $(a_0, a_1, \ldots, a_n, 0, 0, \ldots) = a_0 + a_1 X + \ldots + a_n X^n$, e o conjunto $A^{<\omega}$ por $A[X]$. Esta representação pode ser reconciliada com as operações da álgebra de polinômios colocando $X = (0, 1, 0, 0, \ldots)$; vê-se facilmente, por indução, que X^n é a seqüência com 1 na posição $n+1$, e 0 nas demais. O leitor não terá dificuldade em verificar que a soma e produto em $A^{<\omega}$ não são outros senão a soma e produto usuais de polinômios, aprendidos na juventude.

A seguir, investigamos que tipo de estrutura é $A[X]$, e exatamente como se relaciona com A. Para expressar mais completamente esta relação, que se tornará importante mais tarde, introduziremos outros tipos de estrutura algébrica, explicados no Apêndice.

Proposição 6 *Seja A um anel comutativo e $A[X]$ a álgebra de polinômios sobre A. Então $A[X]$ é uma A-álgebra comutativa que contém uma cópia de A, isto é, há um morfismo injetivo de A-álgebras $A \overset{\iota}{\hookrightarrow} A[X]$.*

Demonstração. A associatividade e comutatividade de $+$ sobre $A[X]$ são claras da Definição 10. O papel de elemento neutro é interpretado pela seqüência identicamente nula $0 = (0)_{n\in\mathbb{N}}$, e o elemento identidade é a seqüência $(1, 0, 0, \ldots)$, com apenas um elemento não-nulo, na primeira posição, igual a 1. Para $p = \sum_{j=0}^{n} a_j X^j$, o elemento $\overline{p} = \sum_{j=0}^{n} \overline{a_j} X^j$ é o oposto de p. A comutatividade de $\cdot$ também segue facilmente da simetria em i, j da condição $i + j = n$.

Sejam agora $p = (p_n)_{n\in\mathbb{N}}, q = (q_n)_{n\in\mathbb{N}}, r = (r_n)_{n\in\mathbb{N}} \in A[X]$; o n-ésimo elemento da seqüência $(p \cdot q) \cdot r$ é:

$$\sum_{i+j=n} (p \cdot q)_i r_j = \sum_{i+j=n} [\sum_{k+l=i} p_k q_l] r_j = \sum_{k+l+j=n} p_k q_l r_j$$

Similarmente, o n-ésimo elemento da seqüência $p \cdot (q \cdot r)$ é:

$$\sum_{i+j=n} p_i (q \cdot r)_j = \sum_{i+j=n} p_i [\sum_{k+l=j} q_k r_l] = \sum_{i+k+l=n} p_i q_k r_l$$

Logo $(p \cdot q) \cdot r = p \cdot (q \cdot r)$. Finalmente, o n-ésimo elemento de $p \cdot (q + r)$ é:

$$\sum_{i+j=n} p_i (q + r)_j = \sum_{i+j=n} (p_i q_j + p_i r_j) = \sum_{i+j=n} p_i q_j + \sum_{i+j=n} p_i r_j$$

Este é justamente o n-ésimo termo da seqüência $p \cdot q + p \cdot r$, o que conclui a verificação de que $A[X]$ é um anel comutativo.

O morfismo injetivo $A \overset{\iota}{\hookrightarrow} A[X]$ é dado pela inclusão de A como "polinômios constantes", isto é, $a \mapsto (a, 0, 0, \ldots) = a + 0X + 0X^2 + \ldots$. A checagem de que este mapa é um morfismo de anéis é rotineira.

A estrutura de $A[X]$ como A-módulo é dada por esta inclusão, da seguinte forma: a imagem de um elemento $a \in A$ em $\mathrm{End}(A[X])$ é a operação de multiplicação por $\iota(a)$. Explicitamente, a ação de $a \in A$ sobre um polinômio $p = \sum_{j=0}^{n} p_j X^j \in A[X]$ é multiplicar todos os seus coeficientes por a: $a \cdot p = \iota(a) \cdot p = \sum_{j=0}^{n} (a p_j) X^j$.

Para ver que $A \overset{\iota}{\hookrightarrow} A[X]$ é um morfismo de A-módulos, basta acompanhar a ação de um elemento $a \in A$ sobre qualquer $b \in A$: temos $\iota(a \cdot b) = (a \cdot b, 0, 0, \ldots) = (a, 0, 0, \ldots) \cdot (b, 0, 0, \ldots) = \iota(a) \cdot \iota(b) = a \cdot \iota(b)$. □

Tendo em vista a pretensa aplicação destes polinômios na expressão de fórmulas lógicas, o leitor pode legitimamente se perguntar se polinômios sobre uma única variável serão suficientes para expressar fórmulas complexas como $p \vee (q \rightarrow r)$. A resposta está contida no próxima Definição, que explica parcialmente a opção pela maior generalidade proporcionada por anéis (ao invés de corpos, mais comumente utilizados).

Definição 11 *Seja A um anel comutativo e n um inteiro positivo. Definimos a* álgebra de polinômios a n variáveis sobre A, *denotada $A[X_1, \ldots, X_n]$, da seguinte forma: $A[X_1] = A[X]$, e $A[X_1, \ldots, X_{n+1}] = A[X_1, \ldots, X_n][X]$.*

A definição é adequada porque, indutivamente, cada $A[X_1, \ldots, X_n]$ é um anel comutativo, e portanto está definida a álgebra de polinômios sobre $A[X_1, \ldots, X_n]$. É importante ter em mente que o X em $A[X]$ é apenas uma notação, significando a construção dada na Definição 10. Se, conforme os comentários sucedendo aquela Definição, desejamos interpretar X como um elemento de $A[X]$, é necessário ter cuidado na Definição 11: obviamente, não interpretaremos o X em cada $A[X_1, \ldots, X_n][X]$ como o mesmo elemento! Quando lidarmos com este tipo de situação, empregaremos os símbolos $X_1, \ldots, X_n$ para denotar os elementos dos diferentes níveis, e um polinômio genérico adquire a forma

$$p(X_1, \ldots, X_n) = \sum_{\overline{\alpha} \in \mathbb{N}^n} p_{\overline{\alpha}} X_1^{\alpha_1} \cdots X_n^{\alpha_n}$$

onde cada $p_{\overline{\alpha}} \in A$ e $p_{\overline{\alpha}} \neq 0$ apenas para um número finito de $\overline{\alpha} \in \mathbb{N}^n$.

Da discussão anterior sobre a imersão $A \hookrightarrow A[X]$ está claro que valem as inclusões

$$A \hookrightarrow A[X_1] \hookrightarrow A[X_1, X_2] \hookrightarrow \ldots \hookrightarrow A[X_1, \ldots, X_n]$$

e que cada $A[X_1, \ldots, X_i]$ age sobre todos os $A[X_1, \ldots, X_j]$ subseqüentes de modo natural, pela multiplicação induzida pela inclusão. Em particular, cada $A[X_1, \ldots, X_n]$ é uma A-álgebra contendo A de modo canônico. Cabe observar, algo pedantemente, que numa expressão do tipo $X_1 X_2^2 + X_3$ deve-se entender os símbolos X_1, X_2 como imersões em $A[X_1, X_2, X_3]$, pela cadeia de imersões acima, dos elementos $X_1 \in A[X_1]$, $X_2 \in A[X_1, X_2]$.

No resto da Dissertação, usaremos o termo "álgebra de polinômios sobre A" para nos referirmos a qualquer das A-álgebras $A[X_1, \ldots, X_n]$, $n \in \mathbb{N} \backslash \{0\}$.

1.2.2 Aspectos Semânticos

Até agora temos lidado com os aspectos "sintáticos" dos polinômios, vendo-os apenas como expressões formais. É tempo de mostrá-los sob a ótica em que são primeiro encontrados na escola, isto é, como *funções* polinomiais.

Começamos com o protótipo de espaço ambiente em que se darão as considerações semânticas.

Lema 7 *Seja X um conjunto qualquer (finito ou não) e A um anel. O conjunto A^X, das funções $X \to A$, é dotado da estrutura de anel pelas operações de soma e produto pontuais, com elemento neutro o mapa $x \mapsto 0$ e identidade o mapa $x \mapsto 1$. Se A é comutativo, assim o é A^X.*

Há uma imersão canônica $A \hookrightarrow A^X$ que leva $a \in A$ no mapa constante $x \mapsto a$, que dá a A^X a estrutura de A-módulo. Se A é comutativo, A^X adquire a estrutura de A-álgebra.

Para cada $X' \subseteq X$ há a operação de restrição $A^X \twoheadrightarrow A^{X'}$ *dada por $f \mapsto f\restriction_{X'}$. Se $X' = \{x\}$ é unitário, confundimos a restrição $A^X \twoheadrightarrow A^{\{x\}} \simeq A$ com o mapa* avaliação em x, *denotado $\varepsilon_x \colon A^X \twoheadrightarrow A$ e definido por $f \mapsto f(x)$. Todas estas operações são morfismos de anéis e A-módulos. Se A é comutativo, são morfismo de A-álgebras.*

Demonstração. Que A^X é um anel com as operações pontuais segue facilmente do fato de que A é um anel, e a imersão $a \mapsto \{x \mapsto a\}$ é um morfismo de anéis. Como visto no Lema 5, o fato de A^X ser um anel resulta na existência de um morfismo natural $A^X \hookrightarrow \mathrm{End}(A^X)$ dado pela multiplicação à esquerda. Compondo a série de morfismos $A \hookrightarrow A^X \hookrightarrow \mathrm{End}(A^X)$ obtemos a ação de A sobre A^X que faz deste um A-módulo. Conforme o comentário após a Definição 4 do Apêndice, quando A é comutativo esta mesma ação faz de A^X uma A-álgebra. □

Vale notar que o conjunto A^X é não-vazio, independente do axioma da escolha, pois as funções constantes $x \mapsto a$ são membros explícitos de A^X.

Definição 12 *Seja A um anel comutativo e $A[X]$ a álgebra de polinômios sobre A. Definimos a* interpretação canônica de $A[X]$, *um mapa $\phi : A[X] \to A^A$ que associa a cada polinômio uma função $A \to A$, da seguinte forma:*

$$\left[\phi(\sum_{j=0}^{n} p_j X^j)\right](a) = \sum_{j=0}^{n} p_j a^j$$

onde seguimos a convenção $a^0 = 1$.

As funções em A^A que estão na imagem de $A[X]$ sob a interpretação canônica são chamadas *funções polinomiais*. A interpretação canônica $\phi \colon A[X] \to A^A$ é um morfismo de A-álgebras, e portanto as funções polinomiais formam uma sub-A-álgebra de A^A. Para cada $a \in A$, a composição $A[X] \xrightarrow{\phi} A^A \xrightarrow{\varepsilon_a} A$, vista como avaliação de polinômios no ponto $X = a$, é também um morfismo de de A-álgebras.

A extensão desta idéia aos polinômios em várias variáveis é simples, e fornece uma noção de "avaliação parcial", que reduz o número de variáveis de um polinômio e é bastante útil para argumentos indutivos.

Definição 13 *Se A é um anel comutativo e $A[X_1, \ldots, X_n]$ uma álgebra de polinômios, a* interpretação canônica *é o mapa $\phi : A[X_1, \ldots, X_n] \to A^{A^n}$ definido por*

$$\left[\phi(\sum_{\overline{\alpha}\in\mathbb{N}^n} p_{\overline{\alpha}} X_1^{\alpha_1}\cdots X_n^{\alpha_n})\right](a_1,\ldots,a_n) = \sum_{\overline{\alpha}\in\mathbb{N}^n} p_{\overline{\alpha}} a_1^{\alpha_1}\cdots a_n^{\alpha_n}$$

Se $n \geq 2$, ponha $B = A[X_1,\ldots,X_{n-1}]$. Para cada $a_n \in A$, temos a avaliação parcial

$$\varepsilon_{X_n\mapsto a_n} : A[X_1,\ldots,X_n] \to A[X_1,\ldots,X_{n-1}]$$

definida pela composição

$$A[X_1,\ldots,X_n] \xrightarrow{\phi} B^{A[X_1,\ldots,X_{n-1}]} \twoheadrightarrow B^A \xrightarrow{\varepsilon_{a_n}} A[X_1,\ldots,X_{n-1}]$$

em que a seta do meio é o morfismo de restrição do Lema 7.

No Capítulo seguinte estaremos interessados em expressar várias funções como polinômios, e portanto é natural perguntar quando as funções polinomiais compõem *toda* a álgebra de funções.

Definição 14 *Um anel comutativo A é dito* polinomial *se toda função $A \to A$ é polinomial.*

Gostaríamos de classificar completamente os anéis polinomiais, com o objetivo de saber quais são promissores para a aplicação à expressão polinomial de fórmulas lógicas. Um primeiro resultado nesta direção é o seguinte.

Teorema 8 *Seja A um anel comutativo infinito. Então há funções $A \to A$ que não são polinomiais. Isto é, as funções polinomiais formam uma sub-A-álgebra própria de A^A.*

Demonstração. Seja $A_n[X]$ o conjunto de todos os polinômios de grau exatamente n; cada um destes é especificado por $n+1$ coeficientes, sendo que o coeficiente líder é não-nulo. Portanto, o cardinal de $A_n[X]$ é o mesmo de $A\backslash\{0\} \times \prod_{i=0}^{n-1} A$. Se A é infinito, este cardinal é simplesmente $|A^n| = |A|$. Agora, a álgebra de polinômios $A[X]$ é exatamente $\bigcup_{n\in\mathbb{N}} A_n[X]$, logo seu cardinal é $|\aleph_0 \times A| = |A|$. Como as funções polinomiais são a imagem de $A[X]$ sob a interpretação canônica, há no máximo $|A|$ funções polinomiais. (De fato, levando em conta os polinômios constantes, há exatamente $|A|$.)

Por outro lado, há $|A^A| \geq 2^{|A|} > |A|$ funções $A \to A$. Portanto, nem toda função é polinomial. □

O Teorema 8 mostra que, se quisermos ter certeza que todas as funções $A \to A$ são polinômios, devemos restringir nossa atenção ao caso em que A é finito. Porém, esta condição não é suficiente: há anéis finitos em que nem toda função é polinomial. Vemos um caso ilustrativo abaixo.

Exemplo 1 *Seja $\mathbb{Z}_4$ o anel de inteiros módulo 4. Existem funções $\mathbb{Z}_4 \to \mathbb{Z}_4$ que não são polinomiais.*

Demonstração. O argumento é não-construtivo, pois há $4^4 = 256$ funções $\mathbb{Z}_4 \to \mathbb{Z}_4$, número pouco adequado à checagem exaustiva no espaço de uma Dissertação. Porém, a finitude de todos os objetos envolvidos garante que é possível encontrar exemplos concretos sem muita dificuldade, talvez lançando mão de um computador.

Primeiro, uma simples checagem verifica que todo elemento $x \in \mathbb{Z}_4$ satisfaz $x^4 = x^2$, de modo que a imagem de $\mathbb{Z}_4[X]$ sob a interpretação canônica é gerada (enquanto $\mathbb{Z}_4$-módulo) pelas imagens dos elementos $1, X, X^2, X^3 \in \mathbb{Z}_4[X]$: em qualquer expressão em que apareçam potências maiores que 3, podemos substitui-las por potências menores (usando a regra $x^4 = x^2$) até que sobrem apenas potências entre 0 e 3.

Há exatamente 4^4 polinômios de grau no máximo 3 sobre $\mathbb{Z}_4$. Seja $\mathbb{Z}_4^{\leq 3}[X]$ o conjunto destes; trata-se de um grupo comutativo com respeito à soma. Restringindo a interpretação canônica a este, e considerando apenas a soma, temos um morfismo de grupos comutativos $\phi^{\leq 3}\colon \mathbb{Z}_4^{\leq 3}[X] \to \mathbb{Z}_4^{\mathbb{Z}_4}$. Como o domínio e contradomínio têm a mesma cardinalidade, $\phi^{\leq 3}$ será sobrejetiva se, e somente se, for injetiva. Porém, há vários polinômios em $\mathbb{Z}_4^{\leq 3}[X]$ cuja interpretação canônica é a função identicamente nula. Um exemplo é $2X(X+1)$; uma checagem rápida mostra que $2x^2 + 2x = 0$ para qualquer $x \in \mathbb{Z}_4$.

Portanto, $\phi^{\leq 3}$ não é injetiva, e há funções não-polinomiais $\mathbb{Z}_4 \to \mathbb{Z}_4$. □

O Exemplo 1 mostra que é necessário impor mais condições sobre um anel A para que toda função $A \to A$ seja polinomial. Mais ainda: que, se quisermos estabelecer que toda função é polinomial, devemos evitar que polinômios de grau baixo sejam identicamente nulos. O exemplo específico construído, $2X(X+1)$, funciona porque há elementos $a, b \in \mathbb{Z}_4$ tais que $a, b \neq 0$ mas $ab = 0$ (a saber, $a = b = 2$).

O Exemplo 1, aliado ao Lema 1 do Apêndice, sugere que tentemos estabelecer dois resultados, para que possamos considerar resolvida a questão da expressão polinomial de funções arbitrárias. Primeiro, que se um anel finito A não é domínio de integridade, nem toda função $A \to A$ é polinomial; segundo, que se K é um corpo, então toda função $K \to K$ é polinomial. Atacamos estes problemas a seguir; começamos com uma espécie de "compacidade" do A-módulo de funções polinomiais, quando A é finito.

Proposição 9 *Seja A um anel comutativo finito. O sub-A-módulo de funções polinomiais em A^A consiste exatamente das interpretações canônicas dos polinômios de grau no máximo $|A| - 1$.*

Demonstração. Na linha do Exemplo 1, encontramos um polinômio $p_A \in A[X]$, de grau $|A|$ e coeficiente líder 1, cuja interpretação canônica é identicamente nula:

$$p_A(X) = \prod_{a \in A} (X - a)$$

Está claro que, para cada $b \in A$, a expressão de $\phi(p_A)(b)$ contém um fator $b - b$, e portanto é nula.

O polinômio p_0, sendo da forma $X^{|A|} + \sum_{j=0}^{|A|-1} p_j X^j$ e interpretado pela função nula, fornece uma regra de substituição $a^{|A|} = -\sum_{j=0}^{|A|-1} p_j a^j$, válida para todo $a \in A$. Portanto, toda função polinomial em A é igual à interpretação canônica de algum polinômio de grau no máximo $|A| - 1$. □

Teorema 10 *Seja A um anel comutativo finito que não é um domínio de integridade. Então há funções $A \to A$ que não são polinomiais.*

Como este Teorema é importante para o projeto da Dissertação, e matematicamente interessante, damos duas demonstrações. A primeira é mais reveladora, porém mais trabalhosa.

Primeira demonstração. Seja $A_{|A|-1}[X]$ o sub-A-módulo de polinômios de grau no máximo $|A| - 1$; vimos na Proposição 10 que as funções polinomiais são a interpretação canônica de $A_{|A|-1}[X]$. Consideremos a restrição da interpretação canônica,

$$\phi_{|A|-1} \colon A_{|A|-1}[X] \to A^A$$

Trata-se de um morfismo de A-módulos. Tanto o domínio quanto o contradomínio têm exatamente $|A|^{|A|}$ elementos; portanto, $\phi_{|A|-1}$ é sobrejetiva somente se for injetiva. Nosso problema fica reduzido, então, a encontrar um polinômio não-nulo $q \in A_{|A|-1}[X]$ cuja interpretação canônica seja a função nula.

Para cada $a \in A$, ponha $N(a) = \{b \in A \ : \ ab = 0\}$, e seja $a_0 \in A$ tal que $N(a_0) \neq \{0\}$. Tal a_0 existe porque A não é domínio de integridade; temos $|N(a_0)| \geq 2$. Definimos

$$q(X) = a_0 X \cdot \prod_{a \notin N(a_0)} (X - a)$$

Está claro que, se $b \in N(a_0)$, então $q(b) = (a_0 b) \cdot u = 0 \cdot u = 0$; e que, se $b \notin N(a_0)$, então $q(b)$ contém um fator $b - b$, e é portanto nulo. Assim, a interpretação canônica de q é a função nula. O grau de q é $1 + |A| - |N(a_0)| \leq |A| - 1$, como desejávamos. Logo $\phi_{|A|-1}$ não é injetiva, e A não é polinomial. □

Segunda demonstração. Suponha que toda função $A \to A$ seja polinomial, tome $a \in A \backslash \{0\}$ e considere a função $f \colon A \to A$ dada por $f(a) = 1$, $f(x) = 0$ para $x \neq a$. Seja $p = \sum_{j=0}^{n} p_j X^j$ um polinômio cuja interpretação canônica é f. Da hipótese $p(0) = 0$ temos que $p_0 = 0$; portanto $p = X \cdot \left(\sum_{j=1}^{n} p_j X^{j-1} \right)$. Agora, $p(a) = 1$ nos dá $a \cdot \left(\sum_{j=1}^{n} p_j a^{j-1} \right) = 1$; ou seja, a possui inverso multiplicativo. Portanto A é um corpo. □

Nota. Usando a técnica da segunda demonstração, conseguimos facilmente construir funções não-polinomiais explícitas. Por exemplo, em $\mathbb{Z}_4$, a função dada por $\{0,1,3\} \mapsto \{0\}$, $2 \mapsto 1$ não é polinomial.

Tendo resolvido metade do problema de classificação dos anéis polinomiais, voltamo-nos agora para o problema recíproco, também com duas demonstrações.

Lema 11 *Seja D um domínio de integridade arbitrário e $p \in D[X]$ não-nulo de grau d. Então $\phi(p)$ assume o valor 0 em no máximo d pontos de D.*

Demonstração. Por indução no grau d; para $d = 0$, p é um polinômio constante não-nulo, e a conclusão é óbvia. Tome então p de grau $d+1$. Se $\phi(p)$ não assume o valor 0, o Lema vale. Se existe $a \in D$ tal que $\phi(p)(a) = 0$, o bem conhecido algoritmo de divisão de polinômios da Álgebra Elementar mostra que há um polinômio $q \in D[X]$ de grau d tal que $p(X) = (X - a) \cdot q(X)$.

Agora, se $b \in D \backslash \{a\}$ é tal que $p(b) = 0$, então $(b - a) \cdot q(b) = 0$; como $b - a \neq 0$ e D é um domínio de integridade, segue que $q(b) = 0$. Por hipótese de indução, há no máximo d elementos de D com esta propriedade, e portanto p tem no máximo $d + 1$ raízes no total. □

Teorema 12 *Seja K um corpo finito. Então toda função $K \to K$ é polinomial.*

Primeira demonstração. Continuamos na mesma linha de raciocínio, considerando a interpretação canônica restrita

$$\phi_{|K|-1} : K_{|K|-1}[X] \to K^K$$

que é um morfismo de K-módulos, ambos de cardinalidade $|K|^{|K|}$. Porém, agora estabelecemos injetividade de $\phi_{|K|-1}$, mostrando que nenhum polinômio não-nulo é interpretado pela função nula.

De fato, sabemos pelo Lema 11 que uma função polinômial, interpretação canônica de um polinômio não-nulo de grau d sobre um corpo K, pode assumir o valor 0 no máximo d vezes. Portanto, a interpretação canônica de cada elemento de $K_{|K|-1}[X]$ tem no máximo $|K| - 1$ zeros, e cada uma destas funções deve assumir ao menos um valor não-nulo. Em particular, nenhuma delas é identicamente nula, exceto a interpretação canônica de $0 \in K_{|K|-1}[X]$. □

Segunda demonstração. É possível dar um polinômio explícito cuja interpretação canônica seja uma função $f : K \to K$ dada, com a chamada "fórmula de interpolação de Lagrange". É a seguinte:

$$p_f(X) = \sum_{a \in K} f(a) \cdot \left(\prod_{b \neq a} \frac{X - b}{a - b} \right)$$

Para ver que $\phi(p_f) = f$, ponha $\delta_a(X) = \prod_{b \neq a} \frac{X-b}{a-b}$; está claro que $\phi(\delta_a)(a) = 1$, ao passo que $\phi(\delta_a)(b) = 0$ se $b \neq a$. A fórmula para p_f acima nada mais é que a soma $\sum_{a \in K} f(a)\delta_a$, e evidentemente $\phi(p_f)(a) = f(a)$ para cada $a \in K$. □

A fórmula de Lagrange pode ser aplicada mais geralmente sobre um corpo K arbitrário, onde permite construir um polinômio de grau d que assume valores pré-determinados em d pontos distintos de K. É interessante notar uma simplificação que a fórmula admite sobre um corpo finito.

Lema 13 *Seja K um corpo finito. O produto de todos os elementos não-nulos de K é -1.*

Demonstração. Seja $P = \prod_{a \in K \setminus \{0\}} a$; nossa estratégia é explorar a existência de um inverso para cada elemento, levando ao cancelamento da maior parte dos termos do produto. De fato, se a e a^{-1} aparecem no produto, por comutatividade e associatividade podemos removê-los.

O único empecilho ao cancelamento completo é a existência de elementos a que são seus próprios inversos, de modo que apenas $a = a^{-1}$ aparece no produto. Felizmente, estes são poucos: são as soluções de $X^2 = 1$, que, da Álgebra Elementar, sabemos serem apenas -1 e 1 (que possivelmente são iguais). Temos então $\prod_{a \in K \setminus \{-1,0,1\}} a = 1$ por cancelamento; e, quer $1 = -1$ ou não, $\prod_{a \in \{-1,1\}} a = -1$. Deste modo, $P = -1$. □

De posse deste Lema, vemos que o produto $\prod_{b \neq a}(b - a)$, ocorrendo no denominador de cada termo do somatório no Teorema 12, é justamente o produto dos elementos não-nulos, isto é, -1. Temos então a expressão simplificada

$$p_f(X) = \sum_{a \in K} -f(a) \cdot \left(\prod_{b \neq a} (X - b) \right)$$

A seguir, investigamos a possibilidade de expressar polinomialmente funções $K^n \to K$, visto que pretendemos polinomizar fórmulas lógicas a várias variáveis. Daremos duas demonstrações, cada uma ao longo das mesmas linhas de uma daquelas do Teorema 12. Para a primeira, precisamos de um resultado preliminar sobre os zeros de polinômios de grau baixo.

Lema 14 *Seja D um domínio de integridade arbitrário e $p \in D[X_1, \ldots, X_n]$ de grau no máximo $|D| - 1$ em cada variável. (Grau arbitrário se D for infinito.) Se $\phi(p)$ é a função nula, então p é o polinômio nulo.*

Demonstração. Procedemos por indução em n. Para $n = 1$, o resultado é basicamente o Lema 11, se observarmos que D tem mais elementos que o grau de p. Agora escreva $p = q_{|K|-1} X_{n+1}^{|K|-1} + \ldots + q_1 X_{n+1} + q_0$, com cada $q_i \in D[X_1, \ldots, X_n]$, e fixe valores arbitrários $X_1 = a_1, \ldots, X_n = a_n$; obtemos assim um polinômio $p_{a_1,\ldots,a_n}(X_{n+1}) \in D[X_{n+1}]$.

Por hipótese, este polinômio de grau $|D| - 1$ assume o valor zero nos $|D|$ pontos de D, e portanto é identicamente nulo, isto é, tem coeficientes nulos. Isto vale para cada escolha de $a_1, \ldots, a_n$; portanto, a interpretação canônica de cada um dos q_i é a função nula. Por hipótese de indução, todos os q_i são polinômios nulos, e assim p também é o polinômio nulo. □

Teorema 15 *Seja K um corpo finito e n um inteiro positivo. Então toda função $K^n \to K$ é a interpretação canônica de um polinômio em $K[X_1, \ldots, X_n]$.*

Primeira demonstração. Seja $K_{|K|-1}[X_1, \ldots, X_n]$ o sub-K-módulo de polinômios de $K[X_1, \ldots, X_n]$ cujo grau em cada variável é no máximo $|K| - 1$. A restrição da interpretação canônica

$$\phi_{|K|-1} : K_{|K|-1}[X_1, \ldots, X_n] \to K^{K^n}$$

é um morfismo de K-módulos. Há $|K|^n$ diferentes monômios $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ com expoentes $\alpha_i \in \{0, \ldots, |K| - 1\}$, e portanto há $|K|^{|K|^n}$ diferentes polinômios em $K_{|K|-1}[X_1, \ldots, X_n]$. Como este é o número de funções $K^n \to K$, basta mostrar que $\phi_{|K|-1}$ é injetiva, isto é, que nenhum polinômio não-nulo de $K_{|K|-1}[X_1, \ldots, X_n]$ é interpretado pela função nula. Este é o Lema 14. □

Segunda demonstração. Dada uma função arbitrária $f : K^n \to K$, construimos explicitamente um polinômio $p_f \in K_{|K|-1}[X_1, \ldots, X_n]$ cuja interpretação canônica é f. Como antes, começamos por construir polinômios que se anulam em todos os pontos de K^n exceto um: dado $\overline{a} = (a_1, \ldots, a_n) \in K^n$, defina

$$\delta_{\overline{a}}(X_1, \ldots, X_n) = (-1)^n \cdot \prod_{i=1}^{n} \prod_{b \neq a_i} (X_i - b)$$

Está claro que, se $\overline{a}' = (a'_1, \ldots, a'_n) \neq \overline{a}$, há um índice i tal que $a_i \neq a'_i$, e portanto $\prod_{b \neq a_i}(a'_i - b) = 0$, por conter um fator $a'_i - a'_i$. Deste modo, um dos fatores do produtório externo é nulo, e $\delta_{\overline{a}}(\overline{a}') = 0$. Por outro lado, para cada i, $\prod_{b \neq a_i}(a_i - b) = -1$ pelo Lema 13, de modo que $\delta_{\overline{a}}(\overline{a}) = (-1)^n \cdot (-1)^n = 1$. Finalmente, dada f, escrevemos

$$p_f(X) = \sum_{\overline{a} \in K^n} f(\overline{a}) \cdot \delta_{\overline{a}}(X)$$

que manifestamente satisfaz $\phi(p_f)(\overline{a}) = f(\overline{a})$ para cada $\overline{a} \in K^n$. □

Com isto, encerramos a classificação de anéis comutativos de acordo com a polinomialidade de suas funções. Apenas em corpos finitos K tem-se todas as funções $K \to K$ polinomiais, e nestes tem-se o bônus de todas as funções $K^n \to K$ serem polinomiais.

Um último refinamento sobre anéis não-polinomiais pode ser relevante. Argumentamos que, se A é um anel comutativo finito que não é domínio de integridade, então há funções $A \to A$ que não são polinomiais, e usamos isto como justificativa para, no próximo Capítulo, polinomizar lógicas principalmente sobre corpos.

Contudo, pode-se levantar a questão de que as lógicas usuais não possuem grande número de conectivos, e que provavelmente não seria necessário expressar *todas* as funções em termos de polinômios. Por isso, seria interessante obter uma medida de qual proporção das funções $A \to A$, ou mais geralmente $A^n \to$

A, podem ser expressas por polinômios. O resto deste Capítulo trata deste tema. Por não ser este um objetivo central, e de interesse mais matemático do que lógico, adotamos um tom um pouco mais avançado e assumimos alguma bagagem algébrica da parte do leitor.

Proposição 16 *Seja A um anel comutativo finito que não é domínio de integridade. A fração de funções $A \to A$ que é polinomial é no máximo $2^{-|A|/2}$. Se k é o menor fator primo de $|A|$, esta fração é no máximo $2^{-\frac{k-1}{k}|A|}$.*

Demonstração. Seja $a_0 \in A$ um elemento não invertível, que existe pelo Lema 1 do Apêndice. O mapa $a_0\cdot : A \to A$ dado por $x \mapsto a_0 x$ é um morfismo de grupos comutativos (em relação à soma), e não é sobrejetivo porque $a_0 x \neq 1$ para todo $x \in A$. Logo, o núcleo $\mathrm{Nuc}(a_0\cdot)$ é um subgrupo não-trivial, e contém pelo menos k elementos, onde k é o menor fator primo de $|A|$.

Apesar de $a_0\cdot$ ser morfismo de grupos, e em geral não de anéis, verifica-se facilmente que $\mathrm{Nuc}(a_0\cdot)$ é um ideal de A; podemos então formar o anel quociente $A/\mathrm{Nuc}(a_0\cdot)$, que tem no máximo $|A|/k$ elementos. Sejam $b_1, \ldots, b_m \in A$ um sistema de representantes distintos das classes de equivalência módulo $\mathrm{Nuc}(a_0\cdot)$ (de modo que $m \leq |A|/k$), e defina

$$p_{a_0}(X) = a_0 \prod_{j=1}^{m} (X - b_j)$$

Vejamos que $\phi(p_{a_0})$ é a função nula. Por construção, para cada $a \in A$ há um b_j tal que $a - b_j \in \mathrm{Nuc}(a_0\cdot)$; deste modo, $\prod_{j=1}^{m}(a - b_j)$ está em $\mathrm{Nuc}(a_0\cdot)$, e multiplicando por a_0 obtém-se 0. Ou seja, $\phi(p_{a_0})(a) = 0$ para qualquer $a \in A$.

Deste modo, obtemos um polinômio p, de grau no máximo $|A|/k$, cuja interpretação canônica é nula. Na notação do Teorema 10, consideremos agora o morfismo de A-módulos

$$\phi_{|A|-1} : A_{|A|-1}[X] \to A^A$$

Mostraremos que $\mathrm{Nuc}(\phi_{|A|-1})$ tem pelo menos $2^{\frac{k-1}{k}|A|}$ elementos, e, do Primeiro Teorema do Isomorfismo para grupos comutativos, resultará que a imagem $\phi_{|A|-1}(A_{|A|-1}[X])$ (isto é, as funções polinomiais) terá no máximo $2^{-\frac{k-1}{k}|A|}$ vezes a cardinalidade de $A_{|A|-1}[X]$. Uma vez que este último conjunto tem a mesma cardinalidade de A^A, teremos o resultado.

Está claro que, para qualquer $q \in A_{|A|-1}[X]$, então $p \cdot q \in \mathrm{Nuc}(\phi_{|A|-1})$. Se q tiver grau no máximo $\frac{k-1}{k}|A| - 1$, então $p \cdot q \in A_{|A|-1}[X]$. Nossa estratégia é construir uma família $Q = \{q_1, \ldots, q_N\}$ de polinômios de grau no máximo $\frac{k-1}{k}|A| - 1$, de modo que todos os produtos $p \cdot q_i$ sejam distintos. Deste modo, estabeleceremos que $\mathrm{Nuc}(\phi_{|A|-1})$ tem pelo menos N elementos. Conseguiremos levar a cabo esta tarefa com $N = 2^{\frac{k-1}{k}|A|}$.

Seja Q o conjunto de todos os polinômios em $A_{|A|-1}[X]$ de grau no máximo $\frac{k-1}{k}|A| - 1$, cujos coeficientes são todos 0 ou 1; está claro que $|Q| == 2^{\frac{k-1}{k}|A|}$. Se $q, q' \in Q$ são distintos, cada coeficiente de $q - q'$ é um dentre $\{-1, 0, 1\}$, e

não são todos 0 porque $q \neq q'$. Em particular, o coeficiente líder de $q - q'$ é -1 ou 1, e portanto $p \cdot (q - q') \neq 0$. Deste modo, se $q \neq q'$, temos $p \cdot q \neq p \cdot q'$; como vimos, isto implica que $\mathrm{Nuc}(\phi_{|A|-1})$ tem pelo menos $2^{\frac{k-1}{k}|A|}$ elementos, e que as funções polinomiais são no máximo $2^{-\frac{k-1}{k}|A|}$ de todas as funções $A \to A$. □

Nota. A Proposição acima mostra que, conforme consideramos anéis com mais elementos, e estes não são corpos, a proporção de funções polinomiais decresce muito rápido. Por exemplo, menos de um milésimo das funções $\mathbb{Z}_{15} \to \mathbb{Z}_{15}$ são polinomiais!

Capítulo 2

Aplicações e Variações

De posse de teoremas de estrutura muito gerais sobre polinômios, passamos a alguns exemplos particulares. A aplicação mais imediata do método de polinomização é sobre lógicas semânticas (Definição 1.8) finitamente valoradas.

2.1 Polinomização de Lógicas Semânticas

A idéia básica é encarar os valores-verdade como elementos de um anel (com operações ainda a definir), e escrever fórmulas lógicas, como $(a \wedge b) \vee c$, na forma de um polinômio nas variáveis a, b, c, como $ab + c + abc$. O *desideratum* é que, ao interpretarmos os valores-verdade de a, b, c como elementos de um anel, calcularmos o valor do polinômio no ponto (a, b, c), e reinterpretarmos o elemento resultante como um valor-verdade, este será exatamente aquele fornecido pela fórmula lógica original.

Talvez o exemplo mais simples seja o da lógica proposicional clássica, com apenas dois valores-verdade: "verdadeiro" e "falso". Aqui, temos um conectivo unário de negação, $\neg$, e conectivos binários de disjunção $\vee$, conjunção $\wedge$, implicação $\rightarrow$ e equivalência $\leftrightarrow$. É bem sabido que o par $\{\neg, \rightarrow\}$ é suficiente para expressar os demais, ou mesmo apenas o conectivo de Sheffer, definido por $a|b = \neg(a \wedge b)$. Porém, é ilustrativo expressar todos estes por polinômios, para obter uma idéia da complexidade dos polinômios envolvidos.

Iniciamos nossa investigação pelo caminho mais simples: havendo apenas dois valores-verdade, procuramos um anel comutativo com apenas dois elementos. A Álgebra Elementar nos revela que existe apenas um anel desta cardinalidade, e este resulta comutativo (de fato, um corpo): $\mathbb{Z}_2 = \{0, 1\}$, os inteiros módulo 2. As operações são induzidas pela soma e multiplicação usual de inteiros: $0 + x = x$, $1 + 1 = 0$, $1 \cdot x = x$ e $0 \cdot 0 = 0$.

Escolhemos arbitrariamente interpretar 0 como "falso" e 1 como "verdadeiro", e, para expressar a negação, procuramos um polinômio $p_\neg \in \mathbb{Z}_2[X]$ tal que $p_\neg(0) = 1$ e $p_\neg(1) = 0$. Sendo $\mathbb{Z}_2$ corpo, o Teorema 2.13 afirma que existe tal polinômio.

A tarefa prática de encontrar um exemplo é bastante simplificada pela Proposição 2.10, que restringe nossa busca aos polinômios de grau $|\mathbb{Z}_2| - 1$, ou seja, da forma $aX + b$ com $a, b \in \mathbb{Z}_2$. Tendo em vista os requerimentos do parágrafo anterior, é simples deduzir que $a = b = 1$ é a única solução:

$$p_\neg(X) = 1 + X$$

É interessante notar que $p_\neg$ não depende da escolha de qual elemento de $\mathbb{Z}_2$ representa qual valor-verdade lógico; isto não ocorre para os conectivos $\vee, \wedge, \rightarrow$.

Procuramos agora um polinômio $p_\wedge(X, Y)$ que desempenhe para a conjunção o mesmo papel que $p_\neg$ para a negação; devemos ter $p_\wedge(x, y) = 1$ somente se $x = y = 1$, e 0 nos outros casos. A demonstração do Teorema 2.16 indica que há um tal $p_\wedge$ de grau no máximo 1 em cada variável, isto é, da forma $aXY + bX + cY + d$. Este pequeno sistema de 4 equações e 4 variáveis pode ser resolvido sem maiores problemas, resultando em

$$p_\wedge(X, Y) = XY$$

De modo análogo, encontramos

$$p_\vee(X, Y) = X + Y + XY$$

$$p_\rightarrow(X, Y) = 1 + X + XY$$

$$p_\leftrightarrow(X, Y) = 1 + X + Y$$

Conforme apontado em [?], este artifício simples facilita bastante algumas demonstrações no Cálculo Proposicional clássico, desde que o número de variáveis envolvidas não seja muito grande. Por exemplo, se quisermos verificar que a fórmula $a \rightarrow (b \rightarrow a)$ é uma tautologia, escrevemos o polinômio correspondente:

$$\begin{aligned} p_\rightarrow(a, p_\rightarrow(b, a)) &= 1 + a + a \cdot (1 + b + ba) \\ &= 1 + a + a + ab + a^2 b \\ &= 1 + ab + ab \\ &= 1. \end{aligned}$$

Como o resultado é 1 para quaisquer valores de a, b, e determinamos $p_\rightarrow$ com a convenção de que 1 representa o valor-verdade "verdadeiro", concluímos que $a \rightarrow (b \rightarrow a)$ é verdadeira para quaisquer valores de a, b, ou seja, é uma tautologia.

Vale notar a importância do uso de todas as propriedades algébricas que definem $\mathbb{Z}_2$: a saber, $a + a = 0$ e $a^2 = a$ para qualquer $a \in \mathbb{Z}_2$. De fato, este é um componente importante do método de polinomização enquanto recurso

computacional, e é possível demonstrar que ele fornece um mecanismo de decisão para lógicas semânticas, em circunstâncias razoavelmente gerais. Expressamos este fato numa seqüência de proposições, para maior clareza.

Comecemos por uma definição razoavelmente geral de *polinomização*, que será comum a todas as proposições subseqüentes.

Definição 1 *Seja $\mathcal{L}$ uma lógica semântica com conectivos $(\mathcal{C}_n)_{n\in\mathbb{N}}$. Uma* polinomização *de $\mathcal{L}$ é uma semântica verofuncional sobre um anel comutativo A tal que todos os conectivos de $\mathcal{L}$ são dados por polinômios sobre A.*

Mais formalmente, é uma interpretação $\mathcal{P} = (\mathcal{P}_n)_{n\in\mathbb{N}}$ onde cada $\mathcal{P}_n(\mathcal{C}_n)$ está contido na sub-A-álgebra das funções polinomiais.

Teorema 1 *Seja $\mathcal{L}$ uma lógica semântica finitamente valorada. Então $\mathcal{L}$ possui uma polinomização.*

Demonstração. Se $\mathcal{L}$ possui uma semântica verofuncional sobre um conjunto S de k elementos, pela discussão que segue a Definição 1.8 e pela Proposição 2 do Apêndice, podemos supor que S está propriamente contido em um corpo finito K. Para cada conectivo n-ário c de $\mathcal{L}$, a interpretação $\phi_n(c)$ é uma função $S \to S$, que precisamos completar a uma função polinomial $\overline{\phi_n}(c) : K \to K$ de modo a respeitar relação de conseqüência $\vdash_{\mathcal{L}}$. Tomamos $u_0 \in K\backslash S$ arbitrário e definimos $\overline{\phi_n}(c)(a_1, \ldots, a_n) = u_0$ se algum $a_i \notin S$. Pelo Teorema 1.12, toda função $K^n \to K$ é polinomial.

Resta checar se a nova interpretação $\overline{\phi} = (\overline{\phi_n})_{n\in\mathbb{N}}$ caracteriza a relação de conseqüência de $\mathcal{L}$. Mostraremos que a extensão $\overline{\phi}$ é "conservativa". Primeiro, por uma simples indução na complexidade, notamos que, se α é qualquer fórmula e a interpretação $\overline{\phi}(\alpha)$ é uma função n-ária, temos que $\overline{a} \notin S^n$ implica $\alpha(\overline{a}) \notin S$.

Agora sejam $\Gamma \subseteq \mathrm{For}(\mathcal{L})$ e $\alpha \in \mathrm{For}(\mathcal{L})$; as considerações acima mostram que, para qualquer valoração ν, só pode ocorrer $\nu_{\overline{\phi}}(\Gamma) \subseteq D$ se todas as variáveis que ocorrem em fórmulas de Γ obtêm, por ν, valores em S. Deste modo, a implicação $\nu_{\overline{\phi}}(\Gamma) \subseteq D \Longrightarrow \nu_{\overline{\phi}}(\alpha) \in D$ é equivalente a $\nu_{\phi}(\Gamma) \subseteq D \Longrightarrow \nu_{\phi}(\alpha) \in D$; e esta última, por hipótese, é equivalente a $\Gamma \vdash_{\mathcal{L}} \alpha$. □

Proposição 2 *Seja $\mathcal{L}$ uma lógica semântica q^n-valorada, onde q é um número primo, e suponha que, em alguma interpretação sobre um conjunto de q^n elementos, há apenas um valor-verdade distinguido. Seja $\mathcal{P}$ uma polinomização de $\mathcal{L}$ sobre um corpo K de q^n elementos, em que $D = \{0\}$. Então, para qualquer $\alpha \in \mathrm{For}(\mathcal{L})$, tem-se $\vdash_{\mathcal{L}} \alpha$ se, e somente se, $\mathcal{P}(\alpha) = 0$.*

Podemos olhar a Proposição acima como uma primeira expressão do *calculemus* de Leibniz; sob certas condições (bastante restritas), a validade de uma fórmula fica reduzida à checagem de que um certo polinômio é identicamente nulo. Considerando a Proposição 1.9 e a demonstração do Teorema 1.12, fica claro que os expedientes computacionais elementares de substituir ocorrências de X^q por X, e eliminar somas do tipo $a + \ldots + a$, com q termos, são suficientes para determinar se $\mathcal{P}(\alpha) = 0$.

Conforme apontado por Carnielli [?], se $\mathcal{L}$ goza de um metateorema da dedução, a Proposição acima se estende para o caso $\alpha \vdash_{\mathcal{L}} \beta$.

Como exemplo, oferecemos uma polinomização da lógica trivalente de Lukasiewicz, com conectivo binário $\rightarrow$ e conectivo unário $\neg$. Uma semântica verofuncional para esta lógica, com valores-verdade $\{0, 1, 2\}$ e único valor distinguido 0, é dada pelas matrizes

$\rightarrow$	0	1	2
0	0	0	2
1	0	0	2
2	0	0	0

	$\neg$
0	2
1	0
2	0

Polinomizando esta lógica sobre o corpo $\mathbb{Z}_3$ e aplicando o Teorema 1.12, obtemos os polinômos $p_{\rightarrow}(X, Y) = 2X(Y + 1)(XY + Y + 1)$ e $p_{\neg}(X) = 2X + 2$. Podemos então checar que $x \rightarrow x$ é um teorema desta lógica, da seguinte forma:

$$\begin{aligned} p_{\rightarrow}(X, X) &= 2X(X + 1)(X^2 + X + 1) \\ &= 2X^4 + 4X^3 + 4X^2 + 2X \\ &= 2X^2 + 4X + 4X^2 + 2X \\ &= 6X^2 + 6X \\ &= 0. \end{aligned}$$

Onde usamos as propriedades $X^3 = X$ na segunda passagem, e $3X = 0$ na quarta.

Podemos generalizar este procedimento para lógicas cujas semânticas verofuncionais não têm apenas um valor distinguido, como segue.

Proposição 3 *Seja $\mathcal{L}$ uma lógica semântica finitamente valorada e $\mathcal{P}$ uma polinomização sobre um corpo finito K, com valores distinguidos D. Então, para qualquer $\alpha \in \mathrm{For}(\mathcal{L})$, tem-se $\vdash_{\mathcal{L}} \alpha$ se, e somente se, $\prod_{d \in D}(\mathcal{P}(\alpha) - d) = 0$.*

A Proposição acima, aliada ao Teorema 1, mostra que lógicas verofuncionais finitamente valoradas admitem uma versão particularmente poderosa do *calculemus*.

2.2 Polinomização de Algumas Lógicas Modais

Apesar de muito bem-sucedida no caso de lógicas verofuncionais finitamente valoradas, a polinomização, conforme definida acima, não tem muito a dizer sobre lógicas não-verofuncionais, nem sobre lógicas não finitamente valoradas. O problema de estender estes métodos para outros contextos, proposto por Carnielli, está em aberto há vários anos.

Um caso particular deste problema já é bastante interessante: o de polinomizar, em algum sentido, alguma das lógicas proposicionais modais bem conhecidas, como K, S4 ou S5. Sabe-se, em particular, por um argumento de Dugundji, que S5 não pode ser caracterizada por matrizes finitas.

Mesmo este caso particular esteve aberto por alguns anos até que Agudelo e Carnielli o resolveram em [?]. A seguir, detalhamos os aspectos mais importantes desta solução.

As chamadas lógicas modais normais são lógicas proposicionais que incluem conectivos da Lógica Proposicional Clássica, aos que se acrescenta os conectivos modais unários $\Diamond, \Box$, denominados "possibilidade" e "necessidade", respectivamente. Na terminologia do Capítulo 1, são lógicas hilbertianas, cujos axiomas incluem todas as tautologias proposicionais e todas as instâncias da regra de Kripke

$$\Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$$

As regras de inferência são modus ponens e a regra de necessitação, de acordo com a qual $\vdash \alpha$ permite concluir $\vdash \Box\alpha$. A lógica normal mais restrita, K, é definida apenas por estas exigências. A lógica S4 é obtida de K adicionando-se os axiomas

$$\Box\alpha \to \alpha$$

$$\Box\alpha \to \Box\Box\alpha$$

A lógica S5 inclui todos os axiomas de S4 e ainda $\alpha \to \Box\Diamond\alpha$.

A discussão filosófica em torno das lógicas K, S4 e S5 é acalorada e sofisticada, envolvendo concepções epistemológicas e metafísicas das idéias de possibilidade e necessidade. Não é nosso objetivo entrar nessa discussão, mas apenas ilustrar o poder da técnica de Agudelo e Carnielli, que, decerto, ainda admitirá muitas explorações e extensões.

Um fator importante sobre as lógicas K, S4 e S5, que facilita seu tratamento por meio de polinômios, é estarem baseadas na Lógica Proposicional Clássica. Por isso, a idéia é polinomizá-las sobre o corpo $\mathbb{Z}_2$, e lançar mão de um artifício reminiscente da Geometria Algébrica para tratar os conectivos modais. Explicitaremos o procedimento para o caso de S5, que é ilustrativo dos demais.

Como na Lógica Proposicional Clássica, deixamos $1 \in \mathbb{Z}_2$ ser o único valor distinguido, e definimos polinômios para os conectivos como no início deste Capítulo, a saber:

$$p_\neg(X) = 1 + X$$

$$p_\to(X, Y) = 1 + X + XY$$

Das Proposições 2 e 3, sabemos que esta polinomização caracteriza completamente a relação de conseqüência da Lógica Proposicional Clássica. Aos operadores modais, contudo, não associamos polinômios, isto é, não há $p_\Box, p_\Diamond$ desempenhando papel semelhante ao dos polinômios acima. A solução é associar a cada fórmula $\Box\alpha$ uma nova variável, denotada $X_{\Box\alpha}$, e restringir a atenção a uma variedade algébrica sobre $\mathbb{Z}_2$.

Mais formalmente, a polinomização de uma destas lógicas modais, $\mathcal{L}$, é uma transformação $\mathcal{P}_M$ que leva elementos de $\mathrm{For}(\mathcal{L})$ em polinômios sobre $\mathbb{Z}_2$, da seguinte forma. Fixamos conjuntos disjuntos de variáveis $\overline{X} = \{X_v : v \in \mathcal{V}_\mathcal{L}\}$ e $\overline{X}' = \{X_{\Box\alpha} : \alpha \in \mathrm{For}(\mathcal{L})\} \cup \{x_{\neg\Box\alpha} : \alpha \in \mathrm{For}(\mathcal{L})\}$, e definimos:

$$\begin{aligned}
\mathcal{P}_M(v) &= x_v, \text{ para } v \in \mathcal{V}_\mathcal{L} \\
\mathcal{P}_M(\neg\alpha) &= \mathcal{P}_M(\alpha) + 1 \\
\mathcal{P}_M(\alpha \to \beta) &= \mathcal{P}_M(\alpha) \cdot (\mathcal{P}_M(\beta) + 1) + 1 \\
\mathcal{P}_M(\Box\alpha) &= X_{\Box\alpha}
\end{aligned}$$

Uma variedade algébrica nada mais é que o conjunto de soluções de um sistema de equações polinomiais. Por exemplo, o conjunto de todos os pares de números reais (x, y) que são solução do sistema de uma única equação $y - x^2 = 0$ é uma variedade algébrica sobre $\mathbb{R}$; trata-se da parábola, familiar aos alunos do Ensino Médio. Também o conjunto de pontos (x, y) que são solução simultânea do par de equações $y - x^2 = 0, x^2 + y^2 - 1 = 0$ é uma variedade algébrica; neste caso, ela consiste de apenas dois pontos.

A variedade algébrica que empregaremos é definida de modo a expressar os axiomas da lógica. Trata-se do subconjunto de $\mathbb{Z}_2^{\overline{X} \cup \overline{X}'}$ definido pelas seguintes equações polinomiais:

$$\begin{aligned}
X_{\Box(\alpha\to\beta)}(X_{\Box\alpha}(X_{\Box\beta} + 1)) &= 0 \\
X_{\Box\alpha}(\mathcal{P}_M(\alpha) + 1) &= 0 \\
\mathcal{P}_M(\alpha)(X_{\Box\Diamond\alpha} + 1) &= 0 \\
X_{\Box\alpha}(X_{\Box\Box\alpha} + 1) &= 0 \\
X_{\neg\Box\alpha} &= X_{\Box\alpha} + 1 \\
\text{Para cada } \alpha \text{ tal que } \mathcal{P}_M(\alpha) = 1, \quad &\text{tem} - \text{se} \quad X_{\Box\alpha} = 1
\end{aligned}$$

Convém notar que, em sua completa generalidade, a construção acima não fornece uma variedade algébrica no sentido usual do termo, visto que o número infinito de variáveis garante que o espaço ambiente, bem como a "variedade", possuem dimensão infinita. Não obstante, em aplicações e cálculos concretos, apenas um número finito de fórmulas está em jogo de cada vez, de modo que podemos restringir a atenção a um número finito de variáveis.

Em termos computacionais, a restrição a uma variedade algébrica significa que há mais regras de manipulação de polinômios disponíveis, para além daquelas fornecidas simplesmente pela estrutura de corpo, viz. $X^{q^n} = X$ e $q \cdot X = 0$. Surpreendentemente, esta linguagem geométrica é suficiente para a expressividade de S5.

Agudelo e Carnielli estabelecem correção e completude para esta semântica polinomial por um argumento bastante técnico, porém inspirado em demonstrações clássicas de completude, como a da Lógica Proposicional Clássica.

Damos a seguir um exemplo de uso deste cálculo polinomial que estabelece que $(\Diamond p \to p) \vee (\Diamond p \to \Box\Diamond p)$ é um teorema em S5. Por questões de espaçoes, abreviamos a polinomização $\mathcal{P}_M(\alpha)$ de uma fórmula α por α^*. Temos:

$$\begin{aligned}((\Diamond p \to p) \vee (\Diamond p \to \Box\Diamond p))^* &= \\ (\Diamond p \to p)^*(\Diamond p \to \Box\Diamond p)^* + (\Diamond p \to p)^* + (\Diamond p \to \Box\Diamond p)^* &= \\ (\Diamond p \to \Box\Diamond p)^*((\Diamond p \to p)^* + 1) + (\Diamond p \to p)^* &= \\ ((\Diamond p)^*((\Box\Diamond p)^* + 1) + 1)((\Diamond p)^*((p)^* + 1)) + (\Diamond p)^*((p)^* + 1) + 1.&\end{aligned}$$

Daí obtemos o polinômio

$$((X_{\Box\neg p} + 1)(X_{\Box\Diamond p} + 1) + 1)((X_{\Box\neg p} + 1)(X_p + 1)) + (X_{\Box\neg p} + 1)(X_p + 1) + 1$$

Lançando mão das regras algébricas de $\mathbb{Z}_2$ e dos polinômios que definem a variedade algébrica, logramos reduzir o polinômio acima a

$$((X_{\Box\neg p} + 1)(X_{\Box\neg p}) + 1)((X_{\Box\neg p} + 1)(X_p + 1)) + (X_{\Box\neg p} + 1)(X_p + 1) + 1,$$

em seguida a

$$(X_{\Box\neg p} + 1)(X_p + 1)(X_{\Box\neg p} + 1)(X_p + 1) + (X_{\Box\neg p} + 1)(X_p + 1) + 1$$

e por fim ao polinômio constante 1, mostrando que se trata, de fato, de uma tautologia.

2.3 Um Teorema de Compacidade

A seguir expomos outra aplicação do método polinomização. Trata-se de um teorema de compacidade, de cunho semântico, válido para lógicas verofuncionais finitamente valoradas bastante gerais.

Teorema 4 *Seja $\mathcal{L}$ uma lógica verofuncional m-valorada, com semântica verofuncional sobre um conjunto A de m elementos e valores distinguidos $D \subseteq A$. Seja $\Gamma \subseteq \mathrm{For}(\mathcal{L})$ e $\alpha \in \mathrm{For}(\mathcal{L})$. Se $\Gamma \vdash_{\mathcal{L}} \alpha$, então existe um subconjunto finito Γ^{fin} de Γ tal que $\Gamma^{\mathrm{fin}} \vdash_{\mathcal{L}} \alpha$.*

Demonstração. Seja $\mathcal{P}$ uma polinomização de $\mathcal{L}$ sobre um corpo finito $K \supseteq A$. Para cada $\alpha \in \mathrm{For}(\mathcal{L})$ e $\Gamma \subseteq \mathrm{For}(\mathcal{L})$, defina

$$\alpha_D = \prod_{d \in D} (\mathcal{P}(\alpha) - d)$$

$$\Gamma_D = \{\gamma_D : \gamma \in \Gamma\}$$

Suponha que $\Gamma \vdash_{\mathcal{L}} \alpha$, onde α depende das variáveis $X_1, \ldots, X_n$, e sejam Γ_D, α_D como acima. Seja $\phi : K[X] \to K^K$ a interpretação canônica de polinômios como funções polinomiais.
Por hipótese, temos que $\phi(\Gamma_D) = \{0\}$ implica $\phi(\alpha_D) = 0$. Se há $\Gamma_D^{\text{fin}} \subseteq \Gamma_D$ finito tal que $\phi(\Gamma_D^{\text{fin}}) = \{0\}$ implica $\phi(\alpha_D) = 0$, qualquer subconjunto Δ de Γ tal que $\Delta_D = \Gamma_D^{\text{fin}}$ satisfaz $\Delta \vdash_{\mathcal{L}} \alpha$; em particular, há um subconjunto finito com esta propriedade. A idéia é demonstrar a existência de um tal Γ_D^{fin}.

Suponha, por absurdo, que não existe Γ_D^{fin}, e defina $\mathcal{F} = \{F \subseteq \Gamma_D : F \text{ é finito }\}$. Considere o conjunto finito K munido da topologia discreta, e $K^{\mathcal{V}_{\mathcal{L}}}$ com a topologia produto; pelo teorema de Tychonoff [?], K^P é compacto.

Agora, por hipótese, para cada $F \in \mathcal{F}$ há $\overline{a}_F \in K^{\mathcal{V}_{\mathcal{L}}}$ tal que $F(\overline{a}_F) = 0$ mas $\alpha_D(\overline{a}_F) \neq 0$. Encarando $\overline{a}_F$ como elemento de $K^{\mathcal{V}_{\mathcal{L}}}$, temos uma rede $(\overline{a}_F)_{F \in \mathcal{F}}$ neste espaço, que por compacidade possui uma subrede convergente; seja $\overline{a}$ um seu ponto de acumulação. Veremos que valem $\Gamma_D(\overline{a}) = 0$ e $\alpha(\overline{a}) \neq 0$.

Seja $\beta \in \text{For}(\mathcal{L})$ uma fórmula nas variáveis $Y_1, \ldots, Y_m$ e $Z_\beta = \{\overline{b} \in K^m : \beta(\overline{b}) = 0\}$; Z_β é um conjunto finito, e portanto $Z_\beta^* = \bigcup_{\overline{b} \in Z_\beta} \overline{b} \times K^{\mathcal{V}_{\mathcal{L}} \setminus \{Y_1, \ldots, Y_m\}}$ é aberto em $K^{\mathcal{V}_{\mathcal{L}}}$. Analogamente, o conjunto $N_\beta = \{\overline{b} \in V^m : \beta(\overline{b}) \neq 0\}$, é finito, logo $N_\beta^* = \bigcup_{\overline{b} \in N_\beta} \overline{b} \times K^{\mathcal{V}_{\mathcal{L}} \setminus \{Y_1, \ldots, Y_m\}}$ é aberto. Uma vez que $K^{\mathcal{V}_{\mathcal{L}}} = Z_\beta^* \cup N_\beta^*$, a união sendo disjunta, temos que Z_β^* e N_β^* são também fechados em $K^{\mathcal{V}_{\mathcal{L}}}$.

Para cada $\gamma_D \in \Gamma_D$ e $F \supseteq \{\gamma_D\}$, temos $\overline{a}_F \in Z_{\gamma_D}^*$; como a subrede dos subconjuntos finitos que contêm $\{\gamma_D\}$ é cofinal, temos que $\gamma_D(\overline{a}) = 0$, como desejado. Similarmente, para cada $F \supseteq \{\alpha\}$ temos $\overline{a}_F \in N_\alpha^*$; por cofinalidade, temos $\alpha(\overline{a}) \neq 0$.

Isto é uma contradição; obtivemos um elemento $\overline{a} \in K^{\mathcal{V}_{\mathcal{L}}}$ tal que $\Gamma(\overline{a}) = 0$ mas $\alpha(\overline{a}) \neq 0$. Portanto, deve haver $\Gamma_D^{\text{fin}} \subseteq \Gamma_D$ finito tal que $\Gamma_D^{\text{fin}} \vdash_{\mathcal{L}} \alpha$. □

2.4 Categorificação

Neste Capítulo, foi apresentada uma definição precisa de polinomização, e, não obstante, a extensão de Carnielli e Agudelo, que não se resume à definição formal. Por isso, torna-se desejável dar uma definição mais abstrata de polinomização.

Uma vez que sua característica básica é a tradução de certos objetos formais em outros, o âmbito mais natural para a formulação abstrata é a Teoria de Categorias; adotamos a maior parte das escolhas subjacentes ao trabalho de Reis [?], porém atemo-nos a conectivos tradicionais, e não tratamos dos conectivos flexíveis por ele investigados. Seria interessante estudar a interação da polinomização com conectivos flexíveis, obtendo alguma espécie de "polinômio flexível"; porém, isto vai além do escopo do presente trabalho.

Reproduzimos aqui, ligeiramente adaptadas, duas das definições que nos interessarão: a noção de tradução entre assinaturas, e a categoria $\mathbf{Log_1}$. Não citaremos explicitamente os lemas técnicos que justificam as diversas construções, mas referimos o leitor à Dissertação de Mestrado de Reis [?].

Definição 2 *Seja $\mathcal{C} = (\mathcal{C}_n)$ uma família de conectivos de diversas aridades, e*

$\mathcal{L}, \mathcal{L}'$ duas lógicas proposicionais com conectivos $\mathcal{C}$. Uma tradução *de $\mathcal{L}$ em $\mathcal{L}'$ é uma família de aplicações $\tau_n : \mathcal{C}_n \to \mathrm{For}(\mathcal{L}')$ que leva cada elemento de $\mathcal{C}_n$ em uma fórmula de $\mathcal{L}'$ com n variáveis.*

Por considerações gerais sobre objetos livres em categorias algébricas, sabemos que qualquer tradução pode ser estendida a uma aplicação $\tau : \mathrm{For}(\mathcal{L}) \to \mathrm{For}(\mathcal{L}')$.

Definição 3 *Seja $\mathcal{C}$ uma família de conectivos. A categoria de lógicas proposicionais $\mathbf{Log_1}^{\mathcal{C}}$ tem como objetos lógicas proposicionais sobre os conectivos de $\mathcal{C}$, e como morfismos traduções $\tau : \mathcal{L} \to \mathcal{L}'$ que preservam a relação de conseqüência, no seguinte sentido:*

$$\Gamma \vdash_{\mathcal{L}} \alpha \implies \tau(\Gamma) \vdash_{\mathcal{L}'} \tau(\alpha)$$

Reis mostra que a composição de morfismos é associativa; o morfismo identidade de uma lógica $\mathcal{L} \in Obj(\mathbf{Log_1}^{\mathcal{C}})$ é a tradução trivial $\mathrm{For}(\mathcal{L}) \to \mathrm{For}(\mathcal{L})$ dada pela aplicação identidade.

Seguindo sugestão de Costa-Leite (comunicação pessoal), tomamos $\mathbf{Log_1}^{\mathcal{C}}$ como categoria-base para sistemas lógicos, e buscamos identificar uma boa noção de polinomização como um funtor $\mathbf{Log_1}^{\mathcal{C}} \stackrel{\mathcal{F}}{\to} \mathbf{C}$; o problema é identificar uma categoria adequada $\mathbf{C}$.

No contexto da Definição 2.1, a escolha natural seria a categoria $\mathbf{A}$ de todas as álgebras polinomiais $A[X_1, \ldots, X_n]$, em que A é um anel e $n \in \mathbb{N}$, com morfismos ditados pela polinomização; não parece haver uma definição intrínseca e natural de morfismos em $\mathbf{A}$ de modo a permitir a definição de um funtor $\mathbf{Log_1}^{\mathcal{C}} \stackrel{\mathcal{F}}{\to} \mathbf{A}$.

O problema de formular a polinomização categorialmente, de modo mais satisfatório, é extremamente interessante, e continua aberto.

Capítulo 3

Limitações

Neste capítulo, investigaremos algumas limitações do método de polinomização. Como bem se sabe, é quase sempre impossível estabelecer, além de qualquer dúvida, que um objetivo não pode ser atingido por um certo método. Sempre resta a possibilidade de que alguém, mais competente que o autor, encontrará uma maneira de driblar os obstáculos.

Não obstante, é prática corriqueira em Matemática, em particular Lógica Matemática, apresentar argumentos sobre a *improbabilidade* de se atingir um objetivo por um certo método, mostrando que o emprego *mais natural* do método necessariamente falha. Este é o tom adotado no presente capítulo.

3.1 Lema de Craig e Nullstellensatz

Começamos por enunciar dois resultados famosos, um da Álgebra e outro da Lógica, sem nos preocuparmos com sua demonstração. O primeiro é o famoso *Nullstellensatz* de Hilbert.

Teorema 1 *Seja K um corpo algebricamente fechado e $K[X_1, \ldots, X_n]$ uma álgebra de polinômios sobre K. Sejam $g, f_1, \ldots, f_m \in K[X_1, \ldots, X_n]$ cujas interpretações canônicas $\phi(g), \phi(f_1), \ldots, \phi(f_m)$ satisfazem a seguinte condição de "dependência": para qualquer $\overline{a} \in K^n$, se $\phi(f_1)(\overline{a}) = \ldots = \phi(f_m)(\overline{a}) = 0$, também $\phi(g)(\overline{a}) = 0$. Então existe um natural N e polinômios $p_1, \ldots, p_m \in K[X_1, \ldots, X_n]$ tais que $g^N = p_1 f_1 + \ldots + p_m f_m \in K[X_1, \ldots, X_n]$.*

Nossa denominação do segundo teorema é um pouco menos precisa. Nos referiremos a ele como Lema de Craig, apesar deste termo ser reservado, em Lógica, a uma vasta coleção de resultados, válidos em diversas lógicas diferentes. De fato, a Lógica de Primeira Ordem goza de um certo Lema de Craig bastante poderoso.

Teorema 2 *Seja $\mathcal{L}$ uma lógica verofuncional funcionalmente completa. Sejam $\alpha, \beta \in \mathrm{For}(\mathcal{L})$ tais que $\alpha \vdash_{\mathcal{L}} \beta$. Então existe uma fórmula θ, cujas variáveis*

ocorrem tanto em α quanto em β, tal que $\alpha \vdash_{\mathcal{L}} \theta$ e $\theta \vdash_{\mathcal{L}} \gamma$. Uma tal fórmula é chamada "interpolante".

Um dos objetivos do presente trabalho é investigar se a polinomização abre ligações entre a Lógica e a Álgebra. Em comunicação pessoal, Carnielli observou que, à primeira vista, o Nullstellensatz de Hilbert parece promissor, enquanto ferramenta algébrica, para demonstrar versões bastante gerais do Lema de Craig. Se tomamos apenas dois polinômios f_1, g tais que $\phi(f_1)(\overline{a}) = 0$ implica $g(\overline{a}) = 0$, a conclusão do Nullstellensatz os relaciona de modo estreito: para certos $N \in \mathbb{N}$, $p_1 \in K[X_1, \ldots, X_n]$ tem-se $g^N = p_1 f_1$. Não é implausível que seja possível extrair desta equação informações sobre as variáveis que f_1 e g têm em comum. Esta avenida de ataque seria levada a cabo com o ferramental das Proposições 2.2 e 2.3.

Dadas fórmulas α, β como no enunciado deste Lema, a idéia seria polinomizá-las, obtendo $\mathcal{P}(\alpha), \mathcal{P}(\beta)$. Somos deliberadamente vagos sobre a natureza exata desta polinomização, exceto que é sobre algum corpo K, e transforma as fórmulas em algum tipo de polinômio. Como visamos argumentar que uma geral linha de ataque não rende frutos, não convém prender-nos a detalhes.

Lançando mão de algum artifício como o da Proposição 2.3, reduziríamos a relação $\alpha \vdash_{\mathcal{L}} \beta$ a uma relação algébrica da forma $\mathcal{P}(\alpha) = 0 \Longrightarrow \mathcal{P}(\beta) = 0$. Se pudéssemos supor que o corpo-base K é algebricamente fechado, resultaria do Nullstellensatz a existência de um natural N e um polinômio p tal que

$$\mathcal{P}(\beta)^N = p \cdot \mathcal{P}(\alpha)$$

Sejam agora $\overline{X} = (X_1, \ldots, X_l)$ as variáveis que ocorrem apenas em α, $\overline{Y} = (Y_1, \ldots, Y_m)$ aquelas comuns a α, β, e $\overline{Z} = (Z_1, \ldots, Z_n)$ as que só ocorrem em β. Do lado esquerdo da equação acima, só ocorrem as variáveis $\overline{Y}, \overline{Z}$. Sendo K algebricamente fechado, será necessariamente infinito, e o Lema 1.14 mostra então que os dois lados da equação acima devem coincidir como polinômios formais, coeficiente a coeficiente. Segue daí que as variáveis $\overline{X}$ não podem ocorrer do lado direito; em particular, não podem ocorrer em α.

Deste modo, se o argumento natural para o Lema de Craig via Nullstellensatz tiver alguma chance de funcionar, deverá demonstrar mais até do que o pedido: se duas fórmulas $\alpha(\overline{X}, \overline{Y}), \beta(\overline{Y}, \overline{Z})$ são tais que $\alpha \vdash_{\mathcal{L}} \beta$, então o próprio $\mathcal{P}(\alpha)$ deverá ser o interpolante no âmbito polinomial. Ao polinomizarmos α, sua dependência em $\overline{X}$ deve naturalmente cancelar-se através apenas das manipulações algébricas elementares envolvidas no *calculemus*.

Porém, isto não pode acontecer em geral, pois há exemplos de fórmulas α cujo valor-verdade depende das variáveis $\overline{X}$ de maneira não-trivial, e portanto cuja polinomização necessariamente envolve estas variáveis.

O exemplo mais simples se dá na Lógica Proposicional Clássica, polinomizada sobre $\mathbb{Z}_2$, desta vez com valor distinguido 0. Tomemos fórmulas $\alpha = p \wedge q \wedge r$ e $\beta = q \wedge r$, em que $\overline{X} = p$, $\overline{Y} = (q, r)$ e não há variáveis $\overline{Z}$. As polinomizações destas fórmulas são, respectivamente:

$$\begin{aligned} p_\alpha(X_p, X_q, X_r) &= (X_p + 1)(X_q + 1)(X_r + 1) + 1 \\ p_\beta(X_q, X_r) &= (X_q + 1)(X_r + 1) + 1 \end{aligned}$$

Agora, está claro que $\alpha \vdash_{LPC} \beta$, e o próprio β é um interpolante. Porém, o que se deve notar aqui é que

$$p_\alpha(X_p, X_q, X_r) = X_p X_q X_r + X_p X_q + X_p X_r + X_q X_r + X_p + X_q + X_r$$

que não admite maiores simplificações. Em particular, nenhuma manipulação algébrica válida sobre $\mathbb{Z}_2[X_p, X_q, X_r]$ pode eliminar os termos em X_p. Isto coloca sérias dúvidas na estratégia de demonstrar o Lema de Craig através do Nullstellensatz.

Na verdade, há uma explicação simples para este fenômeno. Uma condição necessária para a validade do Nullstellensatz é que o corpo-base seja algebricamente fechado, e necessariamente infinito. Porém, em lógicas verofuncionais finitamente valoradas trabalhamos apenas sobre corpos finitos e assim perdemos informação sobre os polinômios.

Uma opção é considerar os fechos algébricos dos corpos finitos sobre os quais trabalhamos; porém, em geral esta passagem não preserva a relação que nos interessa, a saber, $\mathcal{P}(\alpha) = 0 \Longrightarrow \mathcal{P}(\beta) = 0$. Por exemplo, para as fórmulas α, β da Lógica Proposicional Clássica acima, e trabalhando em $\mathbb{Z}_2$, temos que

$$p_\alpha(X_p, X_q, X_r) = 0 \Longrightarrow X_p = X_q = X_r = 0 \Longrightarrow p_\beta(X_q, X_r) = 0;$$

porém, ao passar para o fecho algébrico de $\mathbb{Z}_2$, a primeira implicação deixa de valer. De fato, tomando $X_p = X_q = X_r$ iguais a uma raiz ω de $Y^2 + Y + 1 = 0$, obtemos $p_\alpha(\omega, \omega, \omega) = 0$ sem que $p_\beta(\omega, \omega) = 0$.

Em conclusão, a exigência do Nullstellensatz por um corpo infinito permite que o Lema 1.14 valha em sua versão mais forte, e faz com que os únicos interpolantes que se pode encontrar pelo Nullstellensatz sejam interpolantes triviais. (Isto é, um dentre α, β.)

3.2 Grandes Cardinais e Tabelas de Laver

Do final dos anos 80 ao início dos 90, Laver [27] estudou certas álgebras finitas, em que se define uma única operação binária $\star$ com propriedades pouco usuais. Estas álgebras surgiram, de maneira natural, nas investigações conduzidas por Laver e Dougherty [25, 17] em Teoria de Conjuntos, em particular sobre certos tipos de grandes cardinais: os chamados *rank-into-rank*.

Conforme se vê a Teoria de Modelos, se $\mathfrak{A}$ é uma estrutura sobre a linguagem L, e A é o conjunto subjacente, uma função $j : A \to A$ é chamada *imersão elementar* se:

1. $j(A)$ é o domínio de uma subestrutura de $\mathfrak{A}$, denotada por $j(\mathfrak{A})$;

2. Para cada tupla finita $a \in j(A)^n$, e cada fórmula $\varphi(x_1, \ldots, x_n)$ em L, tem-se $j(\mathfrak{A}) \models \varphi(a)$ se e somente se $\mathfrak{A} \models \varphi(a)$.

Simbolicamente, denota-se este fato por $j : \mathfrak{A} \prec \mathfrak{A}$. Na Teoria de Conjuntos, geralmente são estudadas estruturas sem símbolos de função nem constante, e com o único símbolo de relação $\in$. Por isso, costuma-se abreviar a notação para $j : A \prec A$, já que os demais componentes da estrutura estão subentendidos [24].

As imersões elementares diferentes da identidade são chamadas não-triviais.

Como é de praxe, seja V_α o nível α da hierarquia de conjuntos de von Neumann [23]. Os axiomas de *rank-into-ranks* postulam a existência de cardinais δ tais que existem imersões elementares não-triviais $j : V_\delta \prec V_\delta$. Estas imersões podem ser vistas como uma extensão natural do conceito de infinito como objeto auto-similar. Assim como o menor conjunto infinito, $\aleph_0$, é caracterizado pela existência de uma injeção não-sobrejetiva $\aleph_0 \to \aleph_0$ — isto é, $\aleph_0$ é equipotente a uma sua parte própria — os mais poderosos axiomas de grandes cardinais postulam a existência de cardinais que são elementarmente equivalentes a uma sua parte própria.

A imposição de várias condições sobre tais imersões dá origem a axiomas ainda mais fortes. Os três mais freqüentemente estudados na literatura, em ordem decrescente de força, são:

I1. Existe δ tal que há uma imersão elementar não-trivial $j : V_{\delta+1} \prec V_{\delta+1}$.

I2. Existe uma imersão elementar não-trivial $j : V \prec M$ tal que $V_\delta \subseteq M$ para algum $\delta > crit(j)$ satisfazendo $j(\delta) = \delta$.

I3. Existe δ tal que há uma imersão elementar não-trivial $j : V_\delta \prec V_\delta$.

O significado técnico preciso de I2 não nos preocupará aqui.

O poder dedutivo destes axiomas foi intensamente investigado durante a década de 70, e um estudo bastante completo deste tema encontra-se em [24], particularmente nas seções 5 e 24. O trabalho subseqüente de Laver e Dougherty pode ser visto como extensão natural desta linha de pesquisa: se existem *rank-into-ranks*, qual é a estrutura algébrica que acompanha as imersões elementares associadas a eles? É neste contexto que surge a operação $\star$ de Laver, em particular no estudo de I3, em que se assume que δ é ordinal limite.

A operação é definida da seguinte forma. Dado um ordinal limite λ e imersões elementares não-triviais $j, k : V_\lambda \prec V_\lambda$, define-se $j \star k : V_\lambda \prec V_\lambda$ por:

$$j \star k = \bigcup_{\alpha<\lambda} j(k \cap V_\alpha)$$

Ou seja, $j \star k$ é uma espécie de limite das imagens de segmentos iniciais da imersão k; a idéia é que gostaríamos de encarar k como uma função ordinária (portanto um conjunto) e considerar a imagem de k pela imersão elementar j.

A operação $\star$ é em geral não-associativa, não-comutativa, mas o fato de estar definida entre imersões elementares torna simples demonstrar que satisfaz a propriedade distributiva à esquerda: $j\star(k\star l) = (j\star k)\star(j\star l)$. Esta propriedade também é conhecida como *autodistributividade*.

Para demonstrar certas propriedades técnicas das imersões, Laver foi levado a definir e analisar álgebras finitas autodistributivas (cf. [27]). Estas álgebras, apresentadas efetivamente nas *tabelas de Laver*, possuem propriedades inesperadamente difíceis de demonstrar, e são objeto de diversas conjecturas que parecem se resolver apenas através de hipóteses muito fortes sobre o infinito: a saber, os axiomas I1–I3, acima. A seguir, descrevemos brevemente estas tabelas, reformulando os conceitos envolvidos em termos algébricos mais apropriados ao espírito do presente projeto.

Para cada número natural n, define-se a operação binária $\star_n$ sobre $\mathbb{Z}_n$ da seguinte forma:

$$\forall a \in \mathbb{Z}_n : a \star_n 1 = a + 1,$$

$$\forall a, b, c \in \mathbb{Z}_n : a \star_n (b \star_n c) = (a \star_n b) \star_n (a \star_n c)$$

A operação assim definida, tendo por domínio um conjunto finito, pode ser representada numa tabela $n \times n$: estas são as tabelas de Laver. Curiosamente, as sentenças acima possuem modelos se e somente se n é potência de 2. Isto é, existem tais operações se e só se $n = 2^m$. Abaixo, apresentamos alguns exemplos.

$\star_2$	0	1
0	1	1
1	0	1

$\star_4$	0	1	2	3
0	1	3	1	3
1	2	3	2	3
2	3	3	3	3
3	0	1	2	3

Figura 3.1: Tabelas de Laver para os casos $n = 2$ e $n = 4$

O aspecto mais interessante das tabelas de Laver, do ponto de vista tanto matemático quanto filosófico, é o fato de serem objetos eminentemente finitários, computáveis, e ao mesmo tempo guardarem relação com as noções mais extremas de infinito. De fato, dentre os axiomas de grandes cardinais amplamente estudados, I1–I3 são os mais fortes cuja inconsistência com a teoria de conjuntos ZFC ainda não se demonstrou (cf. [24], seção 24).

Em comunicação pessoal, Carnielli sugeriu que as heurísticas subjacentes ao método de polinomização poderiam ser úteis na investigação de propriedades gerais destas tabelas. Com efeito, a cardinalidade do conjunto subjacente é forçosamente a potência de um número primo, ou seja, é a cardinalidade de algum corpo finito. Isto torna particularmente conveniente a expressão polinomial da operação auto-distributiva: identificamos o conjunto $\mathbb{Z}_{2^m}$ com o corpo finito de 2^m elementos, obtido por extensão algébrica do corpo $\mathbb{Z}_2$. No caso em que $m = 1$, este é o próprio corpo $\mathbb{Z}_2$, e no caso em que $m = 2$ é o corpo $\mathbb{Z}_2[X]/(X^2 + X + 1)$, dos polinômios sobre $\mathbb{Z}_2$ módulo o polinômio irredutível $X^2 + X + 1$. Escrevemos as tabelas de Laver nesta notação, abaixo:

$\star_2$	0	1
0	1	1
1	0	1

$\star_4$	0	1	X	$1+X$
0	1	$1+X$	1	$1+X$
1	X	$1+X$	X	$1+X$
X	$1+X$	$1+X$	$1+X$	$1+X$
$1+X$	0	1	X	$1+X$

Figura 3.2: Tabelas de Laver sobre os corpos $\mathbb{F}_2$ e $\mathbb{F}_4$

Obtemos expressões polinomiais para estas operações como segue:

$$\star_2(P,Q) = P \cdot Q + P + 1$$

$$\begin{aligned}
\star_4(P,Q) \quad = \quad & (1+X) + P(P+1)(P+X)Q(Q+1)(Q+1+X) \\
& +X(P+1)(P+X)(P+1+X)(Q+1)(Q+X)(Q+1+X) \\
& +(1+X)P(P+1)(P+X)(Q+1)(Q+X)(Q+1+X) \\
& +X(P+1)(P+X)(P+1+X)Q(Q+1)(Q+1+X) \\
& +P(P+X)(P+1+X)(Q+1)(Q+X)(Q+1+X) \\
& +P(P+X)(P+1+X)Q(Q+1)(Q+1+X) \\
& +XP(P+1)(P+X)Q(Q+X)(Q+1+X)
\end{aligned}$$

Ambos os polinômios acima foram obtidos pelo mecanismo da segunda demonstração do Teorema 1.15. A característica mais proeminente dos polinômios obtidos por este método é sua complexidade: são o somatório de 2^m termos, cada um dos quais é o produto de $m \times (2^m - 1)$ fatores. Quando se trata de polinômios simples (e.g. lineares), a maior parte destes termos se cancela mutuamente; porém, não há como determinar se este é o caso, exceto pelo expediente de multiplicar todos os termos e obter a expressão do polinômio em soma de monômios. Por exemplo, uma rápida checagem da expressão acima mostra que o monômio P^3Q^3 aparece na expressao de $\star_4(P,Q)$ com coeficiente nulo. Ou seja, o polinômio que expressa $\star_4(P,Q)$ não tem o grau mais alto possível que se poderia esperar da Proposição 1.9, o que parece apontar para algum tipo de simplicidade nos polinômios que expressam as diversas operações $\star_{2^m}$.

Um problema famoso sobre as tabelas de Laver, contra o qual se poderia testar a eficácia da polinomização, é o seguinte. Para cada $m \in \mathbb{N}$, seja L_m o número de elementos distintos que aparecem na primeira linha da tabela de Laver sobre $\mathbb{Z}_{2^m}$, conforme definida acima — isto é, a cardinalidade da imagem da função $x \mapsto 0 \star_{2^m} x$. É verdade que $L_m \to \infty$ quando $m \to \infty$? Este resultado foi demonstrado com técnicas conjuntistas bastante sofisticadas; poderia o método polinomial fornecer uma demonstração mais elementar?

Uma possível estratégia seria assumir o contrário, de modo que L_m permanecesse limitado enquanto $m \to \infty$; daí seguiria que o polinômio a uma variável $\star_{2^m}(0,Q)$ teria muitos valores repetidos. Mais precisamente, se $L_m \leq M$ para todo m, então há algum valor Q_0 tal que $\star_{2^m}(0,Q) = Q_0$ para pelo

menos $2^m/M$ diferentes valores de Q. Isto equivale a dizer que o polinômio $\star_{2^m}(0,Q) - Q_0$ possui pelo menos $2^m/M$ raízes, o que significa, pelo Lema 1.11, que tem grau no mínimo $2^m/M$. (É simples estabelecer que $\star_{2^m}(0,Q)$ não é constante.)

Em linhas gerais, então, vemos que a limitação de L_m implica em grau alto para os polinômios que expressam $\star_{2^m}$; e vimos também, em um caso específico, que $\star_{2^m}$ não tem o grau mais alto possível.

Um obstáculo heurístico, porém tecnicamente formidável, a esta abordagem, é o seguinte. Grosso modo, demonstrações e manipulações que lidam exclusivamente com polinômios sofrem de limitações nos resultados que podem estabelecer: em construções diretas, em geral demonstra-se relações algébricas ou racionais entre algumas quantidades (por exemplo, $L_m \geq \sqrt{m}$); em construções recursivas, às vezes é possível obter relações exponenciais (por exemplo, $L_m \geq \log m$). No entanto, sabe-se que a dependência de L_m cresce com m tão ou mais lentamente que a inversa da função de Ackermann. O fato desta dependência extrapolar o âmbito das funções primitivas recursivas deixa claro que técnicas bastante sutis serão necessárias para permitir que o método polinomial se mostre frutífero nesta questão.

Conclusão

O método de polinomização de lógicas, iniciado por Carnielli e desenvolvido por Carnielli e Agudelo, ainda guarda perguntas. Desenvolvemos sistematicamente seus detalhes nos casos mais simples, envolvendo lógicas caracterizadas por semânticas finitamente valoradas, e descrevemos a adaptação de Carnielli e Agudelo para englobar certas lógicas não finitamente valoradas, mais notavelmente a lógica modal S5.

Questões ainda em aberto incluem a extensão do método para outras lógicas proposicionais não finitamente valoradas, principalmente a intuicionista; formulação categorial satisfatória da polinomização; e extensão infinitária do método para contemplar a Lógica de Primeira Ordem.

Neste último tópico, é possível que a idéia de função geradora, familiar à Combinatória moderna, seja peça fundamental. Um obstáculo técnico significativo é que as operações e questões usualmente consideradas em Combinatória, como partições de conjuntos e número de estruturas de uma certa cardinalidade, obedecem leis formais semelhantes àquelas que regem a multiplicação de séries formais infinitas. Este é o motivo por trás da eficácia das funções geradoras em Combinatória. Contudo, ainda não se conhece analogia semelhante para as operações e questões consideradas pela Lógica Matemática.

Apêndice

A seguir, revisamos algumas definições e primeiros resultados da Álgebra, para conveniência do leitor.

Definição 15 *Seja G um conjunto e $+ : G \times G \to G$ uma operação binária. Dizemos que $(G, +)$ é um* grupo *se: $+$ é associativa; há um elemento "neutro" $0 \in G$ tal que, para todo $x \in G$, vale $x + 0 = 0 + x = x$; e para cada $x \in G$ existe um "oposto" $\overline{x} \in A$ tal que $x + \overline{x} = \overline{x} + x = 0$. O grupo é dito* comutativo *se $+$ é comutativa.*

Se $(G, +)$ e $(G', +')$ são grupos, uma aplicação $f : G \to G'$ é um morfismo de grupos *se vale $f(x + y) = f(x) +' f(y)$ para quaisquer $x, y \in G$.*

Vale notar que morfismos de grupos automaticamente respeitam os elementos neutros, além das operações. De fato, se $f : G \to G'$ é morfismo de grupos, $0 \in G$ e $0' \in G'$ são os respectivos elementos neutros, tem-se $f(0) = f(0 + 0) = f(0) + f(0)$, donde $f(0) + \overline{f(0)} = f(0) + f(0) + \overline{f(0)}$ e portanto $f(0) = 0'$.

Uma das circunstâncias mais naturais que nos levam a considerar anéis é motivada por grupos.

Definição 16 *Seja G um grupo comutativo. O* anel de endomorfismos de G, *denotado* $\mathrm{End}(G)$, *é o conjunto de todos os morfismos de grupos $f : G \to G$, munido das operações de soma pontual e composição funcional: para cada $x \in G$ e $f, g : M \to M$, definimos $(f + g)(x) = f(x) + g(x)$ e $(f \cdot g)(x) = f \circ g(x) = f(g(x))$.*

É claro da definição que o anel de endomorfismos é de fato um anel com respeito às operações escolhidas; o elemento neutro da soma pontual é o endomorfismo identicamente nulo $x \mapsto 0$, e o elemento identidade da composição funcional é o morfismo identidade $x \mapsto x$. Apesar da comutatividade do grupo subjacente, o anel de endomorfismos é mais freqüentemente não-comutativo.

A seguir, definimos os conceitos de módulo e álgebra. Grosso modo, um módulo é um grupo comutativo sobre o qual age um anel; uma álgebra é a mesma coisa, trocando "grupo comutativo" por "anel".

Definição 17 *Seja $(A, +_A, \cdot_A)$ um anel. Um* A-módulo *é um grupo comutativo $(M, +_M)$ munido de uma operação de "multiplicação por escalares" $\cdot : A \times M \to M$ satisfazendo às seguintes condições, para quaisquer $a, b \in A$, $m, n \in M$:*

- $a \cdot (m +_M n) = (a \cdot m) +_M (a \cdot n)$;
- $(a +_A b) \cdot m = (a \cdot m) +_M (b \cdot m)$;
- $(a \cdot_A b) \cdot m = a \cdot (b \cdot m)$;
- $1 \cdot m = m$.

Se M, M' são A-módulos, uma função $f : M \to M'$ é um morfismo de A-módulos *se é um morfismo de grupos comutativos e, para cada $(a, m) \in A \times M$, tem-se $f(a \cdot m) = a \cdot f(m)$.*

Em outras palavras, um A-módulo é dado por um grupo comutativo M e um morfismo de anéis $A \to \mathrm{End}(M)$, que a cada $a \in A$ associa a operação $a\cdot$, ou "multiplicação por a".

Como é usual, omitiremos os subscritos nos símbolos $+_A, +_M$ e semelhantes, confiando no contexto para deixar claro sobre qual domínio se realizam as operações.

Definição 18 *Seja A um anel. Uma* A-álgebra *é um anel $(M, +_M, \cdot_M)$ em que $(M, +_M)$ é um A-módulo e $\cdot_M : M \times M \to M$ é compatível com a multiplicação por escalares deste módulo, no seguinte sentido: para quaisquer $m, n \in M$ e $a, b \in A$, tem-se $(a \cdot m) \cdot (b \cdot n) = (a \cdot b) \cdot (m \cdot n)$. A A-álgebra é* comutativa *se $\cdot_M$ é comutativa.*

Um morfismo de A-álgebras *é um morfismo de anéis que é também morfismo dos A-módulos correspondentes.*

Observamos que qualquer anel comutativo A é uma A-álgebra (e portanto um A-módulo) de modo canônico: toma-se por soma e produto da álgebra as operações correspondentes de A, e pela multiplicação escalar o produto de A; isto é, o mapa $A \to \mathrm{End}(A)$ necessário à definição de módulo é aquele fornecido pelo Lema 5.

Passamos agora a estruturas mais específicas: os domínios de integridade e corpos.

Definição 19 *Um anel comutativo A é um* domínio de integridade *se, para quaisquer $a, b \in A \backslash \{0\}$, tem-se $a \cdot b \neq 0$. É um* corpo *se, para qualquer $a \in A \backslash \{0\}$, existe a^{-1} tal que $a \cdot a^{-1} = 1$.*

Lema 17 */*

Todo corpo é um domínio de integridade. Todo domínio de integridade finito é um corpo.

Demonstração. Seja K um corpo, $a \in K \backslash \{0\}$ e $b \in K$ tais que $ab = 0$. Então $a^{-1}(ab) = a^{-1} \cdot 0 = 0$; mas $a^{-1}(ab) = (a^{-1}a)b = 1 \cdot b = b$. Logo $b = 0$ e K é um domínio de integridade.

Seja agora D um domínio de integridade finito e $d \in D \backslash \{0\}$. Como já vimos, o mapa $x \mapsto d \cdot x$ é um morfismo $D \to D$ de grupos comutativos. A hipótese $dx = 0 \Rightarrow x = 0$ diz justamente que este morfismo é injetivo; sendo D finito, será também sobrejetivo. Logo, há $d' \in D$ tal que $dd' = 1$. □

Proposição 18 *Se q é um número primo e n é qualquer natural, existe um corpo finito com q^n elementos. Em particular, existem corpos finitos arbitrariamente grandes. Ademais, dois corpos finitos de mesma cardinalidade são isomorfos.*

A Proposição acima é bem conhecida da Teoria de Corpos, e sua demonstração não nos preocupará, pois só fazemos uso (no Capítulo 2) do fato de haverem corpos finitos arbitrariamente grandes. Uma demonstração mais simples deste fato é que, para cada primo q, os inteiros módulo q, com as operações usuais de soma e produto módulo q, formam um corpo com q elementos.

Referências Bibliográficas

[1] T. Ahmed, *Algebraic Logic, Where Does It Stand Today?* Bulletin of Symbolic Logic 11(4):465–516, 2005.

[2] G. Birkhoff, S. MacLane, *Algebra*. American Mathematical Society, 1999.

[3] Aristóteles, W. Ross (ed.), *Aristotle's Prior and Posterior Analytics*. Clarendon Press, Oxford, 1951.

[4] G. Boole, *Calculus of Finite Differences*. Chelsea Publishing, Nova York, 1970.

[5] G. Boole, *The Mathematical Analysis of Logic*. Basil Blackwell, Oxford, 1998.

[6] G. Boole, *An Investigation of the Laws of Thought*. Dover Publications, Nova York, 1958.

[7] X. Caicedo, A. Martín, *Completud de dos Cálculos Lógicos de Leibniz*. Theoria 16(3):539–558, 2001.

[8] C. Caleiro, W. Carnielli, M. Coniglio, J. Marcos, *Two's Company: "The Humbug of Many Logical Values"*. Em Logica Universalis, pp. 169–189, Birkhäuser Verlag, Basel, 2005.

[9] W. Carnielli, *A Systematization of the finite many-valued logics through the method of tableaux*. Journal of Symbolic Logic 52(2):473–493, 1987.

[10] W. Carnielli, M. Coniglio, J. Marcos, *Logics of Formal Inconsistency*. Em Handbook of Philosophical Logic, vol. 14, Kluwer Academic Publishers, Dordrecht, 2006.

[11] W. Carnielli, *Polynomial Ring Calculus for Logical Inference*. CLE e-Prints 5(3):1–17, Campinas, 2005.

[12] W. Carnielli, *Polynomizing: Logic Inference in Polynomial Format and the Legacy of Boole*. CLE e-Prints 6(3):1–16, Campinas, 2006.

[13] M. Clegg, J. Edmonds, R. Impagliazzo, *Using the Gröbner Basis Algorithm to Find Proofs of Unsatisfiability*. Proc. of the 28th Annual ACM Symposium on the Theory of Computing, pp. 174–183, 1996.

[14] J. Corcoran, *Aristotle's Prior Analytics and Boole's Laws of Thought*. History and Philosophy of Logic 24:261–288, 2003.

[15] L. Couturat, *La Logique de Leibniz d'Après des Documents Inédits*. Georg Olms, Hildesheim, 1961.

[16] M. Davis, *The Universal Computer: The Road from Leibniz to Turing*. W. W. Norton & Company, Nova York, 2000.

[17] R. Dougherty. *Critical Points in an Algebra of Elementary Embeddings*. Annals of Pure and Applied Logic 65:211–241, 1993.

[18] R. Dougherty, T. Jech. *Left-Distributive Embedding Algebras*. Electronic Research Announcements of the Amer. Math. Soc. 03:28–37, 1997.

[19] M. Dummett, resenha de *Studies of Logic and Probability by George Boole*. Journal of Symbolic Logic 24:203–209, 1959.

[20] H. Eves, *An Introduction to the History of Mathematics*. Brooks Cole, 1990.

[21] S. Gottwald, *A Treatise on Many-Valued Logics*. Studies in Logic and Computation, Research Studies Press Ltd, Hertfordshire, 2001.

[22] N. Jacobson, *Basic Algebra I*. Dover Publications, 2009.

[23] T. Jech, *Set Theory*. Springer-Verlag, Berlim, 2006.

[24] A. Kanamori, *The Higher Infinite*. Springer-Verlag, Berlim, 2003.

[25] R. Laver, *Elementary Embeddings of a Rank into Itself*. Abstracts of the American Mathematical Society 7:6, 1986.

[26] R. Laver, *The Left-Distributive Law and the Freeness of an Algebra of Elementary Embeddings*. Advances in Mathematics 91:209–231, 1992.

[27] R. Laver, *On the Algebra of Elementary Embeddings of a Rank into Itself*. Advances in Mathematics 110:334–351, 1995.

[28] R. Laver, *A Division Algorithm for the Free Left-Distributive Algebra*. Pré-publicação, disponível em http://arxiv.org/abs/math.LO/9204203

[29] G. Leibniz, G. Parkinson, (ed. e trad.), *Leibniz: Logical Papers*. Oxford University Press, Oxford, 1966.

[30] G. Leibniz, L. Loemker (ed. e trad.), *Leibniz: Philosophical Papers and Letters*. Reidel, Dordrecht, 1967.

[31] B. Mates, *The Philosophy of Leibniz*. Oxford University Press, Oxford, 1986.

[32] H. Matsumura, M. Reid, *Commutative Ring Theory*. Cambridge University Press, 1989.

[33] M. Schroeder, *A brief history of the notation of Boole's algebra.* Nordic Journal of Philosophical Logic 2(1):41–62, 1997.

[34] J. Shoenfield, *Mathematical Logic.* AK Peters, Ltd., 2001.

[35] B. Van Der Waerden, B. Van Der Waerden, F. Blum (trad.), *Algebra: Volume I.* Springer-Verlag, Nova York, 1991.

[36] J.-Z. Wu, H.-Y. Tan, Y. Li, *An Algebraic Method to Decide the Deduction Problem in Many-Valued Logics.* Journal of Applied Non-Classical Logics 8(4):353–360, 1998.