Análise, Exequibilidade e Lógica

Fernando Ferreira

Sumário. Faz-se uma sinopse assaz idiossincrática do estudo da extracção de informação computacional de demonstrações em sistemas axiomáticos formais. Descrevem-se quatro casos paradigmáticos e dá-se ênfase aos sistemas relacionados com classes notáveis da complexidade computacional. Finalmente, apresenta-se um sistema de segunda-ordem, computacionalmente exequível, que permite formalizar certos resultados da análise matemática.

1 Introdução

Quando se demonstra uma asserção numa teoria formal verdadeira não só se fica a saber que essa asserção é verdadeira, como também que essa asserção se demonstra na teoria em causa. Foi Georg Kreisel quem chamou a atenção para este ponto quando formulou, em [1], a seguinte questão: "What more than its truth do we know if we have proved a theorem by restricted means?". Estamos interessados em explorar a estratégia que consiste em extrair informação computacional de demonstrações em sistemas formais.

O desenvolvimento deste estratégia é um corolário natural dos trabalhos fundacionais de Kurt Gödel e Alfred Tarski sobre *verdade* e *demonstrabilidade* na década de mil e novecentos e trinta (veja-se [2]). Com efeito, visto que o conceito de demonstrabilidade não exaura o conceito de verdade, é natural tentar explorar matematicamente esta diferença. Porém, por duas ordens de razões, não se pode pôr a questão nestes termos gerais. Por um lado, continua a ser tecnicamente impossível extrair informação computacional genérica de demonstrações em teoria dos conjuntos (ZFC). Por outro lado, mesmo que tal fosse possível, a informação extraida seria irrelevante do ponto de vista computacional (por se obterem classe computacionais que exigem tempo e espaço astronómicos; veja-se, e.g., o primeiro caso abaixo). Devemos contrastar esta atitude com a atitude fundacional - para a qual a extração de informação computacional é um ponto secundário - em que a verdadeira preocupação é efectuar *reduções* de teorias não construtivas a outras menos exigentes ontologicamente. Esta atitude fundacional é a continuação natural do Programa de Hilbert e constitui um dos actuais pólos de investigação em lógica (*vide* [3] e [4]). Neste tipo de investigação têm-se vindo a desenvolver técnicas de estudo das demonstrações formais que, hoje em dia, se aplicam fora do âmbito restrito da fundamentação da matemática.

A questão de Kreisel teve o mérito de desviar alguma investigação em Teoria da Demonstração de preocupações exclusivamente fundacionais para outras mais matemáticas e computacionais. Um dictum que orienta as novas investigações é o seguinte: *mais*

demonstrações $\Rightarrow$ *mais algoritmos*. Por outras palavras, de certas demonstrações em teorias formais podem "ler-se" receitas computacionais. Adicionalmente, quanto menos forte for o sistema axiomático, mais eficientes são as correspondentes receitas. Eis o dictum complementar: *menos axiomas* $\Rightarrow$ *mais eficiência*. Convém esclarecer este segundo dictum: menos axiomas não significa menos quantidade de axiomas mas, sim, menos força dedutiva da axiomática; mais eficiência não significa a obtenção de algoritmos mais concisos ou claros mas, sim, algoritmos cuja implementação necessita de menos recursos de tempo e/ou espaço. Se uma certa asserção, com uma determinada demonstração num sistema axiomático Γ_1, for demonstrável num sistema mais fraco Γ_2 (porventura com uma nova demonstração), então a informação computacional que se extrai da demonstração em Γ_2 é mais rica (no sentido de, genericamente, se extrair um algoritmo mais eficiente do ponto de vista computacional) do que a que provem da teoria Γ_1. A preocupação fundacional em providenciar demonstrações em sistemas tão fracos e construtivos quanto possíveis vem, naturalmente, ao encontro de preocupações mais computacionais. Encontramos, pois, algumas teorias de interesse computacional que surgiram de considerações puramente fundacionais. Hoje em dia têm aparecido novos sistemas, criados especialmente para corresponderem a classes notáveis da complexidade computacional (*vide* o último caso e a última secção deste papel).

Façamos uma derradeira observação introdutória. Os métodos e as ferramentas matemáticas que se utilizam para extrair informação computacional de demonstrações podem ser quaisquer. A restrição metodológica fundacional de apenas utilizar métodos finitistas ou construtivos na metamatemática - a denominada *pureza dos métodos* - deixa de ser relevante nas novas investigações. (Deixamos, porém, um sinal de aviso: a pureza dos métodos poderá ser útil se, no futuro, se considerarem implementações práticas - na forma de linguagens de programação e compiladores - destas investigações.)

Este papel está organizado da seguinte maneira. A seguir à introdução discutem-se quatro casos de sistemas axiomáticos aritméticos, por ordem decrescente de força lógica (e, portanto, por ordem crescente da eficiência algorítmica das receitas computacionais associadas a cada caso). No estudo destes casos procuramos ser precisos na formulação dos resultados. Ficam de fora as demonstrações e, mesmo, as sugestões de demonstrações (ainda que, aqui e ali, atentemos a algumas gemas mais acessíveis). O leitor interessado pode encontrá-las nas indicações bibliográficas que constam do fim de cada secção. Na última parte do papel discutimos brevemente um sistema de análise (i.e., de segunda-ordem) associado à classe computacional dos algoritmos que trabalham em tempo polinomial.

O estudo de cada caso conforma-se a um esquema que passamos a descrever. Seja Γ uma teoria aritmética verdadeira e $A(x, y)$ um predicado recursivo (decidível). Admitamos que a asserção "$\forall x \exists y A(x, y)$" se demonstra em Γ. Em particular esta asserção S é verdadeira, pelo que a função $f(n) = \mu k A(n, k)$ é uma função recursiva total tal que $\forall n A(n, f(n))$ (chamam-se às funções que satisfazem esta condição, funções *testemunhas* da verdade da asserção S). O problema que se põe é o de explicitar - através da limitação dos recursos de tempo e/ou espaço disponíveis para computar funções testemunhas - a informação de que "$\forall x \exists y A(x, y)$" não é somente verdadeira mas, também, demonstrável em Γ. Numa linguagem mais técnica, coloca-se o problema de caracterizar as funções *demonstravelmente totais* da teoria Γ.

Queremos agradecer à organização do *III Encontro de Algebristas Portugueses* o amável convite para apresentar uma comunicação. A versão escrita da comunicação levantou-nos alguns problemas, acabando o resultado final por ser um pouco mais extenso do que aquilo que tinhamos inicialmente previsto. Temos alguma dificuldade em encontrar justificação e motivo para escrever papéis de investigação em português. Sem tirar mérito à presente iniciativa, o público para um papel de Lógica Matemática nestas Actas é, necessariamente, bastante reduzido. A quem se dirige, pois, o papel? Se não é para ser lido, nem para ser criticado pelos meus pares, a sua existência não tem sentido.

Não obstante, o papel aqui está - diante do leitor. Decidimos optar por uma sinopse da nossa área de investigação e da nossa investigação em que, despudoradamente, damos ênfase àquilo que mais nos atrai e apreciamos, a saber: a especial combinação entre o formalismo das linguagens artificiais, o raciocínio matemático rigoroso e as oportunidades de reflexão conceptual. A lógica matemática proporciona amiúde esta combinação e, por isso, é um assunto tão fascinante. Este papel tem a pretensão de transmitir ao leitor um pouco deste fascínio. Tem, também, o objectivo de constituir uma referência para a meia dúzia de pessoas que participa regularmente no Seminário de Lógica Matemática e para quem, com a ajuda deste papel, descubra uma área de interesse e dedicação. Por fim, o Encontro proporcionou-nos a grata oportunidade de tomar contacto com o trabalho de colegas e de conhecer o magnífico *campus* da Universidade de Coimbra.

Indicações bibliográficas

1. Kreisel, G., "Mathematical Significance of Consistency Proofs", *The Journal of Symbolic Logic* 23, pp. 155-182 (1958).

2. Boolos, G. & Jeffrey, R., *Computability and Logic*, Cambridge University Press (1992).

3. Pohlers, W., *Proof Theory: an Introduction*, Lecture Notes in Mathematics 1407, Springer-Verlag (1989).

4. Ferreira, F., "No Paraíso sem Convicção (Uma Explicação do Programa de Hilbert)", a aparecer em Furtado Coelho, J. (org.), *Matemática e Cultura II*.

2 O caso $\boxed{\Gamma = PA}$

A linguagem da aritmética de primeira-ordem compõe-se de uma constante 0, um símbolo funcional unário $'$ (*sucessor*), dois símbolos funcionais binários $+$ e $\cdot$ e um símbolo relacional binário $\leq$. Denotamos por Q a teoria dada pelos seguintes oito axiomas:

$$\begin{array}{rcl} x' & \neq & 0 \\ x' = y' & \rightarrow & x = y \\ x \neq 0 & \rightarrow & \exists y(x = y') \\ x + 0 & = & x \end{array}$$

$$\begin{aligned} x + y' &= (x + y)' \\ x \cdot 0 &= 0 \\ x \cdot y' &= (x \cdot y) + x \\ x \leq y &\leftrightarrow \exists z(z + x = y) \end{aligned}$$

A teoria PA (*Aritmética de Peano*) é formada por Q e pelo seguinte esquema de indução:

$$(IA) \qquad A(0) \wedge \forall x(A(x) \rightarrow A(x+1)) \rightarrow \forall x A(x)$$

onde A é uma fórmula qualquer da linguagem da aritmética, possivelmente com parâmetros (i.e., outras variáveis para além de x). O conjunto dos números naturais ω imbuído da estrutura aritmética usual é um modelo de PA, denominado o *modelo standard*.

Uma classe importante de fórmulas da linguagem da aritmética é a classe das *fórmulas limitadas*, também conhecidas por fórmulas Δ_0. As *quantificações limitadas* são quantificações da forma $\exists x(x \leq t \wedge ...)$ ou da forma $\forall x(x \leq t \rightarrow ...)$, onde t é um termo da linguagem no qual a variável x não ocorre. A classe das fórmulas limitadas é a menor classe de fórmulas que contém as fórmulas atómicas e é fechada para as operações Booleanas e para as quantificações limitadas. Observe-se que estas fórmulas definem conjuntos recursivos no modelo standard.

Para descrever as funções demonstravelmente totais em PA introduzimos a denominada *hierarquia de crescimento rápido* $(F_\alpha)_{\alpha<\epsilon_0}$. (O ordinal ϵ_0 é o supremo da sequência de ordinais $\omega, \omega^\omega, \omega^{\omega^\omega}, ...$) Esta hierarquia define-se por recursão transfinita nos ordinais α inferiores a ϵ_0:

$$\begin{aligned} F_0(n) &= n+1 \\ F_{\lambda+1}(n) &= \underbrace{F_\lambda(F_\lambda(...F_\lambda}_{n+1}(n)...)) \\ F_\alpha(n) &= F_{\{\alpha\}_n}(n) \end{aligned}$$

onde α é um ordinal limite e $\{\alpha\}_n$ é a *sequência fundamental* crescente de ordinais a convergir para α. Definem-se canonicamente estas sequências fundamentais por meio da Forma Normal de Cantor dos ordinais, segundo a qual todo o ordinal α se escreve, de maneira única, do seguinte modo:

$$(*) \qquad \alpha = \omega^{\beta_k} + \omega^{\beta_{k-1}} + ... + \omega^{\beta_1} + \omega^{\beta_0}$$

com $\beta_0 \leq \beta_1 \leq ... \leq \beta_{k-1} \leq \beta_k$. Se $\alpha < \epsilon_0$, o que é o nosso caso, sabe-se que $\beta_k < \alpha$. Podemos, pois, definir (por recursão transfinita) as sequências fundamentais dum ordinal limite $\alpha < \epsilon_0$, escrito na forma $(*)$, assim:

$$\{\alpha\}_n = \omega^{\beta_k} + \omega^{\beta_{k-1}} + ... + \omega^{\beta_1} + \begin{cases} \omega^\gamma \cdot (n+1) & \text{se } \beta_0 = \gamma + 1 \\ \omega^{\{\beta_0\}_n} & \text{se } \beta_0 \text{ é ordinal limite} \end{cases}$$

Devemos avisar o leitor que, apesar de todas as funções F_α serem computáveis (i.e., recursivas), elas têm na prática um crescimento simplesmente astronómico. (Desafiamos o leitor a tentar calcular o número $F_{\omega^{\omega^\omega}+9}(9)$.) Para mais informações sobre este assunto deve consultar-se [1] e a secção 6.3 de [2].

Teorema. *Se $PA \vdash \forall x \exists y A(x,y)$, onde $A \in \Delta_0$, então existe um ordinal $\alpha < \epsilon_0$ e um número natural m tais que $\forall n > m \; \exists k < F_\alpha(n) \; A(n,k)$.*

O teorema afirma que se a teoria PA demonstra a asserção "$\forall x \exists y A(x,y)$" então há funções testemunhas da sua verdade que são dominadas, a partir de certa ordem, por uma função da hierarquia de crescimento rápido. A demonstração do teorema deve-se a contribuições essenciais de Gerhard Gentzen (1935), Georg Kreisel (1952), Helmut Schwichtenberg (1971) e Stan Wainer (1970, 72) e uma exposição encontra-se em [3].

Este resultado entreabre a possibilidade de se obterem resultados de independência. Com efeito, uma asserção matemática verdadeira que se possa exprimir na forma $\forall x \exists y$ $A(x,y)$, com $A \in \Delta_0$, e para a qual a função $f(n) = \mu k A(n,k)$ não seja dominada (a partir de certa ordem) por nenhuma função F_α ($\alpha < \epsilon_0$), é necessariamente independente de PA. Passamos a descrever uma asserção deste tipo, de natureza combinatorial.

Dado k um número natural, denotamos por $[S]^k$ o conjunto de todos os subconjuntos de S com k elementos. Uma *r-coloração* de $[S]^k$ é uma função f de $[S]^k$ no conjunto finito $\{1, 2, ..., r\}$ (as "cores"). Um subconjunto $H \subseteq S$ diz-se *homogéneo* para a coloração f se existir $1 \leq i \leq r$ tal que $f(X) = i$, para todo $X \in [H]^k$ (i.e., todos os subconjuntos de H com k elementos têm a mesma cor). Dadas cardinalidades α e β, escrevemos $\alpha \to (\beta)^k_r$, para significar que toda a r-coloração de $[S]^k$, com S de cardinalidade α, tem um subconjunto homogéneo de cardinalidade β. Em 1930 Frank Ramsey demonstra um resultado lindíssimo: para todos os naturais r e k, $\omega \to (\omega)^k_r$.

O teorema de Ramsey tem uma versão finita, i.e., que não menciona conjuntos infinitos. Reza assim: para quaisquer naturais k, r e m existe um número natural (suficientemente grande) n tal que $n \to (m)^k_r$. Este resultado é uma consequência da versão original do teorema de Ramsey e é interessante saber porquê. Com vista a um absurdo, suponhamos que existem números naturais k, r e m tais que, para nenhum natural n se tem $n \to (m)^k_r$. Seja $T = T_{k,r,m}$ o conjunto de todas as r-colorações de $[\{1, 2, ..., n\}]^k$ para as quais não existe um subconjunto homogéneo de m elementos. Dizemos que uma coloração $\sigma \in T$ precede a coloração $\tau \in T$ se τ é uma extensão (como função) de σ. O conjunto T, com esta relação de precedência, forma uma *árvore*, i.e., uma ordem parcial com mínimo em que o conjunto de predecessores de qualquer elemento é uma cadeia finita. Além disso, T é uma árvore de *ramificação finita*, ou seja, uma árvore em que todo o elemento tem apenas um número finito de sucessores. Ora, por hipótese absurda, T é uma árvore infinita. Assim, pelo lema de König (veja-se [4]), existe um *ramo infinito* em T. No nosso caso este ramo dá origem a uma r-coloração de $[\omega]^k$ que não tem nenhum subconjunto homogéneo de cardinalidade m. Ora, isto contradiz a versão infinita do teorema de Ramsey.

A versão finita do teorema de Ramsey pode exprimir-se na forma $\forall x \exists y A(x,y)$, com $A \in \Delta_0$ e, de facto, demonstra-se em PA (mas, é claro, não pelo argumento acima, pois este usa a versão infinita do teorema de Ramsey cujo enunciado não se pode exprimir na linguagem de PA). No entanto, uma pequena - e, aparentemente, inocente - variação da versão finita do teorema de Ramsey não se demonstra em PA. Um subconjunto finito dos naturais diz-se *grande* se a sua cardinalidade exceder o seu mínimo. A versão modificada afirma que, para quaisquer números naturais k, r e m existe um natural n para o qual qualquer r-coloração de $[\{m, m+1, ..., n\}]^k$ possui um subconjunto homogéneo grande (a notação

abreviada é a seguinte: $n \underset{*}{\rightarrow} (m)^k_r$). Esta versão modificada é uma consequência da versão infinita do teorema de Ramsey (exactamente pelo mesmo argumento acima). Sejam k, r e m números naturais e denotemos por $PH(k, r, m)$ o menor natural n tal que $n \underset{*}{\rightarrow} (m)^k_r$. Em 1980 Jussi Ketonen e Robert Solovay mostraram que a função PH não é dominada (a partir de certa ordem) por nenhuma função da hierarquia de crescimento rápido (há uma exposição deste resultado na secção 6.3 de [2]). Isto implica que a versão modificada do teorema finito de Ramsey é independente da teoria PA. (Nota: este resultado de independência foi demonstrado originalmente por Jeff Paris e Leo Harrington em 1977 com um argumento de teoria dos modelos. Em [5] pode encontrar-se uma exposição desta demonstração.)

Já desde os famosos resultados de incompletude de Gödel em 1931 que se sabe existirem asserções verdadeiras independentes da aritmética de Peano. O interesse dos novos resultados de independência reside no facto deles serem de natureza matemática, ao contrário dos resultados de Gödel que são de natureza metamatemática. Na sequência do resultado pioneiro de Paris e Harrington têm surgido outros resultados de independência, mesmo de teorias mais fortes que PA (*vide* [6]). PA tem, porém, um interesse muito especial. Esta teoria é equivalente (num sentido preciso) à teoria dos conjuntos ZFC com o axioma do infinito substituido pela sua negação constituindo, pois, um candidato natural a uma axiomatização do universo dos conjuntos (hereditariamente) finitos. Ora, a asserção de Paris e Harrington não é, como se viu, demonstrável nesta teoria. Parece que, para se argumentar a verdade desta asserção (que é uma asserção que *não* menciona conjuntos infinitos), é necessário um *detour* pelo infinito actual, nomeadamente através da versão infinita do teorema de Ramsey e fazendo uso do lema de König, que é uma espécie de *viaduto* a ligar o infinito ao finito.

Indicações bibliográficas

1. Smorynski, C., "Some Rapidly Growing Functions", in Harrington L. e al. (orgs.), *Harvey Friedman's Research on the Foundations of Mathematics*, North-Holland, pp. 367-380 (1985).

2. Graham, R., Rothschild B. & Spencer J., *Ramsey Theory*, Wiley-Interscience (1980).

3. Buchholz, W. & Wainer, S., "Provably Computable Functions and the Fast Growing Hierarchy", in Simpson S. (org.), *Logic and Combinatorics*, Contemporary Mathematics, AMS, pp. 179-198 (1987).

4. Hrbacek K. & Jech T., *Introduction to Set Theory*, Marcel Dekker (1984).

5. Ferreira, F., *Models of Arithmetic: An Exposition*, trabalho de aptidão científica, Departamento de Matemática, Universidade de Lisboa (1986)

6. Simpson, S., "Unprovable Theorems and Fast Growing Functions", in Simpson S. (org.), *Logic and Combinatorics*, Contemporary Mathematics, AMS, pp. 359-394 (1987).

3 O caso $\boxed{\Gamma = I\Sigma_1}$

A análise da teoria PA pode ser refinada. Definem-se as *fórmulas* Σ_n, com $n \geq 1$, como sendo as fórmulas da linguagem da aritmética com a seguinte forma:

$$\exists x_1 \forall x_2 \exists x_3 ... Q x_n \; A$$

onde A é uma fórmula limitada e onde o quantificador Q é um $\forall$ ou um $\exists$ conforme n é par ou ímpar (respectivamente). As *fórmulas* Π_n, com $n \geq 1$, são as negações destas, i.e., são da forma:

$$\forall x_1 \exists x_2 \forall x_3 ... Q x_n \; A$$

Toda a fórmula pode ser escrita em forma normal prenexa e, portanto, aparece algures classificada como Σ_n ou como Π_n. No modelo standard estas fórmulas definem os conjuntos da *hierarquia da aritmética* (a notação em vigor é ambígua: conforme o contexto, Σ_n refere-se a fórmulas da linguagem da aritmética ou aos subconjuntos de ω que estas fórmulas definem). Seja Δ_n a classe dos subconjuntos de ω que estão simultaneamente em Σ_n e em Π_n (o leitor mais preparado sabe que a classe Σ_1 é a classe dos conjuntos *recursivamente enumeráveis* e que, portanto, a classe Δ_1 é a classe dos conjuntos recursivos). Temos a seguinte figura:

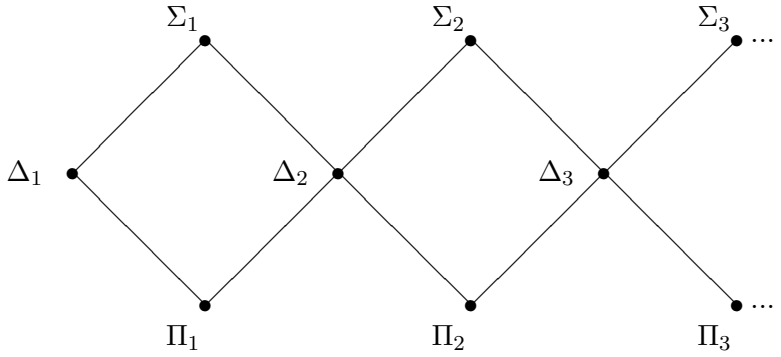

Fig. 1 A hierarquia da aritmética

Nesta figura, o facto de duas classes estarem ligadas por uma linha significa que a classe da esquerda está contida na classe da direita; mormente, pelo teorema de Post (*vide* [1]), estas inclusões são próprias. É natural considerar os subsistemas da teoria PA em que o esquema de indução (IA) se restringe a fórmulas A, com $A \in \Sigma_n$: são as teorias $I\Sigma_n$. Estas teorias formam uma cadeia própria $I\Sigma_1 \subset I\Sigma_2 \subset ...$ cuja união é a teoria PA (no décimo capítulo de [2] há uma exposição deste resultado). A esta hierarquia de teorias associa-se uma correspondente hierarquia de funções:

$I\Sigma_1$	$I\Sigma_2$	...	$I\Sigma_n$	...	$PA = \bigcup_{n\in\omega} I\Sigma_n$
F_α, $\alpha < \omega$	F_α, $\alpha < \omega^\omega$	...	F_α, $\alpha < \omega^{[n]}$	...	F_α, $\alpha < \epsilon_0$

onde $\omega^{[1]} = \omega$ e $\omega^{[n+1]} = \omega^{\omega^{[n]}}$.

Teorema. *Se* $I\Sigma_n \vdash \forall x \exists y A(x,y)$*, onde* $A \in \Delta_0$*, então existe um ordinal* $\alpha < \omega^{[n]}$ *e um número natural* m *tais que* $\forall n > m \; \exists k < F_\alpha(n) \; A(n,k)$.

O caso da teoria $I\Sigma_1$ merece um pouco mais de atenção. A tese central do Programa de Hilbert era a de que se uma asserção finitista (uma asserção *real*, na terminologia de Hilbert) se demonstra utilizando argumentos e asserções infinitistas (asserções *ideais*, na terminologia de Hilbert) então também se demonstra dentro das severas restrições finitistas. Esta tese não é correcta, como demonstrou Gödel em 1931. Tal não impede que uma parte (significativa?) da matemática não se possa efectuar sob uma perspectiva Hilbertiana. Posto de outro modo: seja F a teoria finitista das asserções e argumentos reais de Hilbert; uma extensão T de F é *admissível* do ponto de vista de Hilbert se, sempre que uma asserção real se deduz de T, então essa asserção também se deduz de F (diz-se que a teoria T é uma extensão *conservativa* de F com respeito às asserções reais). Surge aqui uma dificuldade, pois Hilbert nunca definiu formalmente a teoria F (isso não seria necessário *caso* o Programa de Hilbert fosse bem sucedido!). William Tait argumenta persuasivamente em [3] que a teoria formal conhecida por PRA ("Primitive Recursive Arithmetic") captura o finitismo de Hilbert. Ora, se assim for, a teoria $I\Sigma_1$ é admissível do ponto de vista de Hilbert. Uma parte deste resultado está enunciada no seguinte teorema:

Teorema. *Se* $I\Sigma_1 \vdash \forall x \exists y A(x,y)$*, onde* $A \in \Delta_0$*, então existe uma função* f, recursiva primitiva*, tal que* $\forall n A(n, f(n))$.

(A classe das funções recursivas primitivas é a menor classe de funções que inclui a função identicamente nula, a função sucessor e as funções projecções (i.e., as funções $P^i_k(x_1, ..., x_k) = x_i$) e que é fechada para as operações de composição e de definição por *recursão primitiva*. Uma função $(k+1)$-ária f diz-se definida por recursão primitiva a partir da função inicial k-ária g e da função $(k+2)$-ária h se $f(0, x_1, ..., x_k) = g(x_1, ..., x_k)$ e $f(n+1, x_1, ..., x_k) = h(n, f(n, x_1, ..., x_k), x_1, ..., x_k)$. O teorema deve-se, independentemente (à volta do ano de 1970), a Gregory Mints, Charles Parsons e Gaisi Takeuti e [4] expõe uma demonstração.)

Stephen Simpson apresenta em [5] evidência de que uma parte razoável da matemática pode ser efectuada sob uma perspectiva Hilbertiana. O ensaio de Simpson coloca toda esta problemática num âmbito mais geral, nomeadamente no projecto de investigar a *força extensional* exacta dos principais teoremas da matemática. Este projecto, conhecido por *Matemática Recíproca* ("Reverse Mathematics"), procura classificar os teoremas da matemática em vários sistemas axiomáticos de modo a que um determinado teorema não só se demonstre no sistema a ele associado (a parte *directa* da matemática) mas, também, que a própria axiomática seja - num certo sentido - consequência do teorema (a parte *recíproca*). As investigações do programa da matemática recíproca desenrolam-se em sistemas aritméticos de *segunda-ordem*, na senda das próprias investigações de Hilbert e seus

continuadores. O sistema base é a teoria RCA_0 ("Recursive Comprehension Axiom"). Sem pretender descrever completamente esta teoria, que o leitor pode encontrar definida em [5], podemos acrescentar que ela se formula na linguagem da aritmética de segunda-ordem (i.e., com um símbolo relacional binário extra "$\in$" e dois tipos de variáveis: as variáveis numéricas $x, y, z, ...$, que no modelo standard variam nos elementos de ω, e as variáveis de conjunto $X, Y, Z, ...$, que variam em subconjuntos de ω) e que tem como axiomas Q, o esquema (IA) de indução restrito a fórmulas Σ_1 (admitindo-se, agora, também parâmetros de segunda-ordem) e o seguinte postulado:

$$\forall x(A(x) \leftrightarrow B(x)) \rightarrow \exists X \forall x(x \in X \leftrightarrow A(x))$$

onde $A \in \Sigma_1$ e $B \in \Pi_1$ (permitem-se parâmetros de primeira e de segunda ordem). Este axioma esquema pode parecer pouco natural ao não especialista. O que há a ter em conta é que se trata dum axioma de *abstracção* ("comprehension"), i.e., de existência de conjuntos que se obtêm pela extensão duma propriedade. (A força extensional duma teoria mede-se, precisamente, pelos seus postulados de existência de conjuntos.) No modelo standard este axioma afirma simplesmente que se podem formar conjuntos recursivos. O sistema RCA_0 é uma extensão conservativa da teoria $I\Sigma_1$ com respeito a asserções de primeira-ordem e, em particular, com respeito a asserções Π_2. Assim, as funções demonstravelmente totais de RCA_0 são as funções recursivas primitivas e - de facto - esta teoria é admissível do ponto de vista de Hilbert.

Em 1976, Harvey Friedman definiu um sistema axiomático muito interessante: a teoria WKL_0. Esta teoria adiciona a RCA_0 o postulado que afirma que toda a sub-árvore infinita da árvore binária tem um ramo infinito (é o denominado *lema fraco de König*).

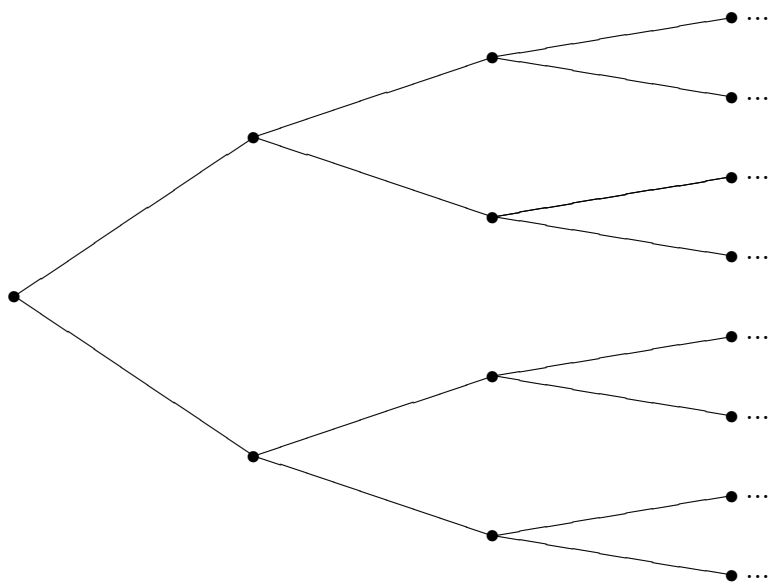

Fig. 2 A árvore binária $\{0,1\}^*$

O lema fraco de König é uma asserção não construtiva, como se pode pressentir pela demonstração que, por etapas, define indutivamente o ramo infinito: a ideia essencial é passar dum nódulo do ramo para o nódulo seguinte de modo a que a parte da sub-árvore a seguir ao novo nódulo se mantenha infinita. Este procedimento não é construtivo pois exige a resposta a um número infinito de perguntas não recursivas, nomeadamente se a parte da sub-árvore a seguir a um determinado nódulo é infinita. De facto, demonstra-se que há sub-árvores infinitas recursivas da árvore binária sem nenhum ramo infinito recursivo. Apesar de tudo, e um pouco surpreendentemente, Friedman demonstra em 1976 que WKL_0 é um sistema admissível do ponto de vista de Hilbert. No seu ensaio Simpson menciona uma lista de teoremas da matemática que são equivalentes ao lema fraco de König (na presença da teoria RCA_0) e que, portanto, se podem formalizar sob uma perspectiva Hilbertiana. Eis a lista: o teorema da cobertura finita de Heine-Borel para conjuntos fechados e limitados de R^n; o teorema da continuidade uniforme e da existência de máximo para funções contínuas definidas num intervalo compacto de R; o teorema da existência local de soluções de equações diferenciais ordinárias; os teoremas de Hahn-Banach e de Alaoglu para espaços de Banach separáveis; o teorema da existência de ideais primos para anéis comutativos contáveis e o teorema de existência e unicidade do fecho algébrico dum corpo contável.

Vamos encerrar esta secção com algumas observações a respeito das duas versões do lema de König. A versão original, enunciada na secção anterior, difere da versão fraca na medida em que admite árvores cuja ramificação, sendo finita, não é limitada uniformemente por uma constante. Pelo contrário, na versão fraca do lema apenas se permitem árvores de ramificação limitada por dois (o caso mais geral, de ramificação limitada por uma constante n, não acrescenta força extensional). Como observámos no parágrafo anterior, a versão fraca do lema de König é admissível do ponto de vista de Hilbert. Porém, demonstra-se que isso já não é o caso com a versão original a qual, portanto, tem uma força extensional estritamente superior à força extensional do lema fraco de König.

Indicações bibliográficas

1. Soare, R., *Recursively Enumerable Sets and Degrees*, Springer-Verlag (1987).
2. Kaye, R., *Models of Peano Arithmetic*, Oxford Logic Guides 15, Claredon Press (1991).
3. Tait, W., "Finitism", *The Journal of Philosophy* 78, pp. 524-546 (1981).
4. Sieg, W., "Herbrand Analysis", *Archive for Mathematical Logic* 30, pp. 409-441 (1991).
5. Simpson, S., "Partial Realizations of Hilbert's Program", *The Journal of Symbolic Logic* 53, pp. 349-363 (1988).

4 O caso $\boxed{\Gamma = I\Delta_0 + exp}$

A teoria $I\Delta_0$ tem como axiomas Q e o esquema de indução restrito a fórmulas Δ_0.

Nesta secção estamos interessados na teoria que se obtem desta adicionando o axioma $\forall x \forall y \exists z$"$z = x^y$": é a denominada teoria $I\Delta_0 + exp$. Não queremos esconder a existência duma dificuldade técnica na formulação deste axioma adicional, nomeadamente na definição da relação "$z = x^y$" na linguagem da aritmética. Gödel mostrou em 1931 como se podem introduzir todas as funções recursivas primitivas na teoria PA (mesmo na teoria $I\Sigma_1$) e, em particular, a função $exp(x, y) = x^y$. A demonstração utiliza a chamada função β de Gödel e não se pode formalizar em $I\Delta_0$. O problema não é grave e a maneira mais simples de o contornar é admitir um novo símbolo funcional binário $exp(x, y)$, na linguagem da aritmética, juntamente com os axiomas $exp(x, 0) = 0'$ e $exp(x, y') = exp(x, y) \cdot x$ (permitindo-se que este novo símbolo funcional apareça na definição das fórmulas Δ_0). Porém, é interessante (e por vezes conveniente) saber que existe uma definição Δ_0 do gráfico da função exponencial para a qual é possível deduzir em $I\Delta_0$ as relações de recorrência da exponenciação. A demonstração deste facto é muito delicada e deve-se a James Bennett (1962, no que concerne à definição Δ_0 do gráfico da exponenciação) e a Jeff Paris (1980, no que concerne à derivação das relações de recorrência da exponenciação na teoria $I\Delta_0$). Eis, precisamente, o que se demonstra: há uma fórmula $\theta(x, y, z) \in \Delta_0$ que satisfaz as três seguintes propriedades,

$$I\Delta_0 \vdash \forall x \forall y \forall z \forall w(\theta(x, y, z) \wedge \theta(x, y, w) \rightarrow z = w)$$

$$I\Delta_0 \vdash \forall x \theta(x, 0, 0')$$

$$I\Delta_0 \vdash \forall x \forall y \forall z(\theta(x, y, z) \rightarrow \theta(x, y', z \cdot x))$$

Com estes factos (recomendamos [1] para uma sua exposição) podemos enunciar rigorosamente o axioma exp: é a asserção "$\forall x \forall y \exists z \theta(x, y, z)$". O leitor deve compreender que exp não é consequência das três propriedades acima listadas. É fácil e instrutivo exibir um modelo de $I\Delta_0$ que não satisfaz exp, i.e., um modelo com as operações usuais de adição e multiplicação para o qual a função exponencial não é total (claro que este modelo apenas satisfaz formas restritas de indução: no caso presente, indução para fórmulas Δ_0; em contrapartida, com indução para fórmulas Σ_1 deduz-se imediatamente exp). Seja, então, M um modelo não standard de PA e a um seu elemento não standard. A sub-estrutura I de M definida por $I = \{c \in M : \exists n \in \omega(c \leq a^n)\}$ é um segmento inicial de M, donde se conclui que as fórmulas limitadas não mudam de valor verdade entre M e I (no dizer técnico, são *absolutas* entre M e I). Por consequência, a estrutura menor I herda todas as verdades Π_1 da estrutura maior M. Como $I\Delta_0$ tem axiomatizações constituidas somente por asserções Π_1, infere-se que I é modelo de $I\Delta_0$. (Um exemplo é a axiomatização que consiste em Q e na seguinte reformulação do esquema de indução: $\forall x(A(0) \wedge \neg A(x) \rightarrow \exists y < x\ (A(y) \wedge \neg A(y + 1)))$, onde A é uma fórmula limitada qualquer.) Não obstante, exp falha em I. Admitamos, com vista a um absurdo, o contrário. Então existe um elemento $c \in I$, e - consequentemente - um natural n com $c \leq a^n$, tal que $I \models \theta(a, a, c)$. Ora, a teoria $I\Delta_0$ demonstra que $\forall x > 0'\ \forall y_1 \forall y_2 \forall z_1 \forall z_2(\theta(x, y_1, z_1) \wedge \theta(x, y_2, z_2) \wedge y_1 < y_2 \rightarrow z_1 < z_2)$. Em particular, podemos concluir que $a^n < c$, o que dá origem a uma contradição.

A classe das funções *elementares à Kalmar* é uma classe de funções mais restrita do que a classe das funções recursivas primitivas. A definição desta nova classe difere da definição

da classe das funções recursivas primitivas em dois pontos: primeiro, incluem-se, *ab initio*, as funções de adição, de multiplicação e exponencial; segundo, substitui-se a operação de recursão primitiva pela operação de *recursão primitiva limitada*. Uma função $(k+1)$-ária f diz-se definida por recursão primitiva limitada a partir da função inicial k-ária g, da função $(k+2)$-ária h e da função de limitação $(k+1)$-ária l se $f(0, x_1, ..., x_k) = g(x_1, ..., x_k)$ e $f(n+1, x_1, ..., x_k) = min\{h(n, f(n, x_1, ..., x_k), x_1, ..., x_k), l(n, x_1, ..., x_k)\}$. O seguinte resultado faz parte do "folclore" da lógica:

Teorema. *Se $I\Delta_0 + exp \vdash \forall x \exists y A(x, y)$, onde $A \in \Delta_0$, então existe uma função f, elementar à Kalmar, tal que $\forall n A(n, f(n))$.*

O trabalho [2] é uma boa referência para este assunto.

Indicações bibliográficas

1. Wilkie, A., "Modèles nonstandard de l'arithmétique et complexité algorithmique", in Ressayre, J. & Wilkie, A. (orgs.), *Modèles non standard en arithmétique et théorie des ensembles*, Publications Mathématiques de l'Université Paris VII, vol. 22, pp. 5-45 (1987).

2. Isabel Matos, A., *Aritmetização e Hierarquia de Funções em Subsistemas da Aritmética de Peano*, trabalho de aptidão científica, Departamento de Matemática, Universidade de Lisboa (1991).

5 O caso $\boxed{\Gamma = I\Delta_0 + \Omega_1}$

Os sistemas estudados nos três casos anteriores satisfazem o axioma *exp* e, portanto, incluem a função exponencial como função demonstravelmente total. Ora, a função exponencial tem um crescimento de tal ordem que não se considera ser *computacionalmente exequível* ("computationally feasible"). Assim, se pretendermos estudar sistemas aritméticos relacionados com classes notáveis da complexidade computacional (*vide* adiante) temos que nos restringir a teorias que não demonstrem *exp*: são os denominados sistemas *fracos* da aritmética.

A teoria $I\Delta_0$ é um candidato natural a considerar. Há, porém, motivos técnicos que tornam esta teoria demasiado fraca e pouco natural. Por exemplo, a teoria $I\Delta_0$ não é fechada para o produto de *comprimentos* de números (o comprimento de um número n, que denotamos por $c(n)$, é o número de símbolos da sua representação em notação binária): ou seja, há modelos de $I\Delta_0$ que incluem elementos a e b para os quais não existe um elemento c no modelo cujo comprimento seja o produto dos comprimentos de a e b. O princípio que é necessário introduzir para impedir esta possibilidade é o axioma Ω_1, que reza: $\forall x \exists z$ "$z = x^{\lfloor log_2 x \rfloor}$" ($\lfloor w \rfloor$ denota o menor inteiro que não excede w). Este axioma pode ser rigorosamente formulado com o auxílio da função θ discutida na secção anterior:

$$\forall x \forall y \forall w (w \leq x \wedge \theta(0'', y, w) \rightarrow \exists z\, \theta(x, y, z))$$

A formulação da teoria $I\Delta_0 + \Omega_1$ não se adapta facilmente a um estudo mais aprofundado dos seus aspectos computacionais. Na sua dissertação de doutoramento [1], Samuel Buss introduz a teoria S_2, que é uma reformulação de $I\Delta_0 + \Omega_1$. Não vamos descrever esta teoria em detalhe pois, mais adiante, vamos preferir trabalhar com outra que (a nosso ver) é-lhe notacionalmente superior. Interessa-nos, de momento, apontar que a introdução dum novo símbolo funcional binário # (leia-se *sharp*) na linguagem de S_2 substitui o papel do axioma Ω_1 (com vista a deixar de haver axiomas Π_2 na teoria). A interpretação deste novo símbolo no modelo standard é a seguinte: $x\#y = 2^{c(x)\cdot c(y)}$ - estamos, pois, em presença de uma função com a ordem de grandeza exacta para obter produtos de comprimentos. Nestas circunstâncias, um argumento clássico (e simples) permite-nos demonstrar que, se $S_2 \vdash \forall x \exists y A(x,y)$, com A uma fórmula limitada, então existe um termo $t(x)$ da linguagem de S_2 tal que $S_2 \vdash \forall x \exists y \leq t(x)\ A(x,y)$. Apesar de apenas termos descrito parcialmente a linguagem da teoria S_2 e de termos deixado em branco as especifidades desta teoria, é possível - e cremos ser interessante - esboçar a demonstração deste resultado (conhecido, na literatura, por Teorema de Parikh). Suponhamos que a conclusão é falsa, i.e., que de S_2 não se deduz $\forall x \exists y \leq t(x) A(x,y)$ para nenhum termo $t(x)$ da linguagem. Aumente-se a linguagem com uma nova constante c e considere-se o conjunto de fórmulas $\Lambda = \{\forall y \leq t(c)\ \neg A(c,y) : t \text{ termo da linguagem de } S_2\}$. Devido à nossa suposição inicial, é fácil de argumentar que Λ é finitamente consistente. Assim, por compacidade, Λ tem um modelo M. A sub-estrutura I de M definida por $I = \{x \in M : \text{existe um termo } t \text{ tal que } M \models x \leq t(c)\}$ é um segmento inicial de M e, por conseguinte, herda todas as verdades Π_1 de M. Como S_2 tem uma axiomatização Π_1 (contrariamente ao que se passa com a teoria $I\Delta_0 + \Omega_1$), I é modelo de S_2. É evidente que neste modelo a asserção "$\exists y A(c,y)$" é falsa. Assim, "$\forall x \exists y A(x,y)$" também é falsa em I e, portanto, não é consequência de S_2. Como se queria demonstrar.

Uma função f diz-se computável em espaço polinomial (ou em $Pspace$) se existir um polinómio $p(x) \in N[x]$ e um algoritmo (uma máquina de Turing, se quisermos ser rigorosos) de tal modo que, para qualquer "input" n, o algoritmo computa o "output" $f(n)$ utilizando, no máximo, $p(c(n))$ "bits" de memória (i.e., $p(c(n))$ casas da fita da máquina de Turing). Qualquer predicado ou relação limitada é decidível em espaço polinomial (i.e., tem a função característica computável em tempo polinomial). Desta observação conclui-se, sem muita dificuldade, que todas as funções testemunhas $f(x) = \mu y \leq t(x)\ A(x,y)$, onde A é uma fórmula limitada, são computáveis em espaço polinomial.

Da discussão dos dois últimos parágrafos conclui-se o seguinte:

Teorema. *Se $I\Delta_0 + \Omega_1 \vdash \forall x \exists y A(x,y)$, onde $A \in \Delta_0$, então existe uma função f, computável em espaço polinomial, tal que $\forall n A(n, f(n))$.*

A conclusão deste teorema não é a melhor possível e pode ser aperfeiçoada em duas direcções. O melhoramento numa das direcções é relativamente trivial e torna-se patente com definições apropriadas. Estamos mais interessados num melhoramento diferente, devido a Buss na sua dissertação. Para poder explicar o resultado de Buss necessitamos de introduzir certas sub-teorias de S_2 e algumas noções básicas de complexidade computacional. *En passant*, mencionaremos o referido melhoramento mais trivial.

A classe computacional mais central da Ciência da Computação é a classe *Ptime* das funções computáveis em tempo polinomial (tomamo-la como a caracterização teórica do conceito "ser computacionalmente exequível"). Esta classe é constituida pelas funções f para as quais existe um polinómio $p(x) \in N[x]$ e um algoritmo (máquina de Turing) de tal modo que, para qualquer "input" n, o algoritmo computa $f(n)$ em, no máximo, $p(c(n))$ passos. (Observe-se que Ptime$\subseteq$Pspace, pois um passo de computação - numa máquina de Turing - pede, quando muito, um novo "bit" de memória. A propósito, não se sabe se esta inclusão é estrita.) Outras classes de complexidade computacional vão ser necessárias, mas introduzi-las-emos em simultâneo com as definições das sub-teorias de S_2 que são relevantes para compreender o resultado de Buss.

Contrariamente ao tratamento de Buss, vamos utilizar a notação binária, ou seja, o alfabeto $\{0,1\}^*$ de todas as sequências finitas (palavras) de zeros e uns. Esta é a notação dos cientistas da computação e pensamos que também é a notação mais adequada no contexto dos sistemas fracos de aritmética. Assim, enquanto para Buss o modelo standard das suas teorias é ω, para nós vai ser a árvore binária $\{0,1\}^*$. A linguagem de primeira-ordem com igualdade das teorias de $\{0,1\}^*$ compõe-se de três constantes ϵ, 0 e 1, dois símbolos funcionais binários $\frown$ (para a *concatenação*) e $\times$, e um símbolo relacional binário $\subseteq$ (para *sub-palavra inicial*). Há catorze axiomas básicos:

$$
\begin{array}{rcl}
x \frown \varepsilon = x & & x \times \varepsilon = \varepsilon \\
x \frown (y \frown 0) = (x \frown y) \frown 0 & & x \times (y \frown 0) = (x \times y) \frown x \\
x \frown (y \frown 1) = (x \frown y) \frown 1 & & x \times (y \frown 1) = (x \times y) \frown x \\
x \frown 0 = y \frown 0 \rightarrow x = y & & x \frown 1 = y \frown 1 \rightarrow x = y \\
x \subseteq \varepsilon & \leftrightarrow & x = \varepsilon \\
x \subseteq y \frown 0 & \leftrightarrow & x \subseteq y \vee x = y \frown 0 \\
x \subseteq y \frown 1 & \leftrightarrow & x \subseteq y \vee x = y \frown 1 \\
x \frown 0 & \neq & y \frown 1 \\
x \frown 0 & \neq & \varepsilon \\
x \frown 1 & \neq & \varepsilon
\end{array}
$$

As interpretações dos símbolos da linguagem no modelo standard (i.e., na árvore binária) são claras: observamos, apenas, que $x \times y$ é a palavra x concatenada consigo própria um número de vezes igual ao comprimento da palavra y. Diz-se que x é uma *sub-palavra* de y, e escreve-se $x \subseteq^* y$, se $\exists z \subseteq y\,(z \frown x \subseteq y)$. A classe das *sw.q.-fórmulas* é a menor classe de fórmulas que inclui as fórmulas atómicas da linguagem e que é fechada para as operações Booleanas e para as *quantificações de sub-palavra* (as quantificações de sub-palavra são da forma $\forall x \subseteq^* t\,(...)$ ou $\exists x \subseteq^* t\,(...)$, onde t é um termo da linguagem no qual a variável x não ocorre). Não é difícil convencermo-nos de que as sw.q.-fórmulas definem, no modelo standard, relações decidíveis em tempo polinomial. Por exemplo, a operação de quantificação existencial de sub-palavra não nos atira para fora da classe Ptime: com efeito, dada uma relação $A(x,y)$ em Ptime, para decidir se, dado um certo σ, se tem $\exists x \subseteq^* t(\sigma)\, A(x,\sigma)$, basta calcular o valor $t(\sigma)$ e, seguidamente, percorrer todas as

sub-palavras ρ desse valor inquirindo, de cada vez, se $A(\rho, \sigma)$ vale. Assim que obtivermos uma resposta afirmativa podemos concluir que $\exists x \subseteq^* t(\sigma)\ A(x, \sigma)$; se todas as respostas forem negativas, podemos concluir o contrário. Como o número de sub-palavras de $t(\sigma)$ é limitado por $1 + \frac{n(n+1)}{2}$, onde n é o comprimento da palavra $t(\sigma)$, e como cada pergunta se responde (por hipótese) em tempo polinomial, conclui-se facilmente que este processo exaustivo está em Ptime. (Podemos acrescentar que a classe das sw.q.-relações é parte *própria* de Ptime: este facto está longe de ser trivial, sendo consequência de um importante resultado em Teoria dos Circuitos devido a M. Furst, J. Saxe & M. Sipser em 1984 e exposto e melhorado por J. Håstad em [2].)

A relação que vale entre x e y se, e somente se, o comprimento de x não exceder o comprimento de y, pode exprimir-se na linguagem formal do seguinte modo: $1 \times x \subseteq 1 \times y$, abreviado por $x \leq y$. (Esta notação é, por vezes, criticada por se confundir com a notação usual da ordem natural de ω. Acreditamos, no entanto, que esta confusão tem algumas virtualidades sob o ponto de vista computacional. Adiantamos uma observação cuja pertinência se tornará mais clara com o prosseguimento da discussão: o número de números naturais que não excedem um dado número n é da ordem de grandeza de $2^{c(n)}$, exactamente a mesma ordem de grandeza do número de todos os elementos de $\{0,1\}^*$ de comprimento menor ou igual que o comprimento duma palavra σ de comprimento n.) As quantificações limitadas são, na terminologia binária, quantificações da forma $\forall x \leq t\ (...)$ ou $\exists x \leq t\ (...)$, onde t é um termo no qual a variável x não ocorre; a classe das fórmulas limitadas, aqui também denominadas fórmulas Σ^b_∞, é a menor classe de fórmulas que inclui as sw.q.-fórmulas e que é fechada para as operações Booleanas e de quantificação limitada. O célebre problema "$P = NP$?" é a questão de saber se a classe Ptime é fechada para quantificações limitadas. O método de exaustão que utilizámos para ver que as quantificações de sub-palavra não nos levam para fora de Ptime falha, desta vez, miseravelmente. Com efeito, uma busca exaustiva com vista a decidir se $\exists x \leq t(\sigma)\, A(x, \sigma)$ percorre $2^{n+1} - 1$ elementos, tantos quantas as palavras que não excedem o comprimento n de $t(\sigma)$. Uma busca exponencial, portanto! Claro que esta análise não afasta uma resposta afirmativa à questão "$P = NP$?", pois pode ser que exista um algoritmo (necessariamente diferente do método exaustivo) que decida se $\exists x \leq t(\sigma)\, A(x, \sigma)$ em tempo polinomial. Tal algoritmo até nem tem que ser "uniforme", podendo depender crucialmente do predicado $A(x, y)$. Não obstante, a maioria dos especialistas inclina-se para uma resposta negativa ao problema. Segundo muitos matemáticos, a questão "$P = NP$?" é o mais importante novo problema da matemática (veja-se, por exemplo, o testemunho de Steve Smale em [3]). O problema foi formulado há vinte e três anos e acredita-se que seja uma questão muito, muitíssimo, difícil.

Definem-se as *fórmulas* Σ^b_n, com $n \geq 1$, como sendo as fórmulas limitadas da linguagem da árvore binária com a seguinte forma:

$$\exists x_1 \leq t_1 \forall x_2 \leq t_2 \exists x_3 \leq t_3 \ ... \ Q x_n \leq t_n\ A$$

onde A é uma sw.q.-fórmula, t_1,...,t_n são termos da linguagem e o quantificador Q é um $\forall$ ou um $\exists$ conforme n é par ou ímpar (respectivamente). As *fórmulas* Π_n, com $n \geq 1$, são

as negações destas, i.e., são da forma:

$$\forall x_1 \leq t_1 \exists x_2 \leq t_2 \forall x_3 \leq t_3 \ldots Q x_n \leq t_n \, A$$

No modelo standard $\{0,1\}^*$ estas fórmulas definem os conjuntos da *hierarquia (do tempo) polinomial*. Aos conjuntos definidos por fórmulas Σ_n^b chamam-se conjuntos Σ_n^p; aos conjuntos definidos por fórmulas Π_n^b chamam-se conjuntos Π_n^p. A definição corrente da classe NP usa o conceito de máquina de Turing *não determinista* (o que explica o "N" de NP) que trabalha em tempo polinomial: o leitor pode encontrar esta definição e uma introdução acessível à complexidade computacional em [4]. Sabe-se que $NP = \Sigma_1^p$; por conseguinte, a classe dos conjuntos complementares de conjuntos em NP, denotada por $co{-}NP$, é a classe Π_1^p. ∇_n^p é a classe dos subconjuntos de $\{0,1\}^*$ que estão simultaneamente em Σ_n^p e em Π_n^p. No caso $n = 1$ ficamos com a classe $NP \cap co - NP$ dos conjuntos *bem-caracterizados*, segundo a terminologia de alguns matemáticos (usada, por exemplo, pelo meu amigo Orestes Cerdeira).

Nesta fase da exposição necessitamos de um conceito da Teoria das Funções Recursivas: o conceito de computação com o auxílio dum *oráculo* (veja-se a referência [1] da seção 3). Um oráculo é um subconjunto Ω de $\{0,1\}^*$; uma computação com o auxílio de Ω consiste numa computação ordinária em que se permite um novo procedimento: *consultar* o oráculo. Quando este procedimento é chamado durante o processo computacional pergunta-se se $\sigma \in \Omega$, onde σ é um valor entretanto já calculado: a computação prossegue conforme a resposta (esta resposta é suposta ser dada imediatamente, contando como *um* passo da computação). Definimos Δ_{n+1}^p como a classe dos conjuntos que são decidíveis em tempo polinomial por meio de uma computação auxiliada por um oráculo em Σ_n^p; Δ_1^p (também denotada por classe P) é a classe dos conjuntos decidíveis em tempo polinomial. Temos a seguinte figura:

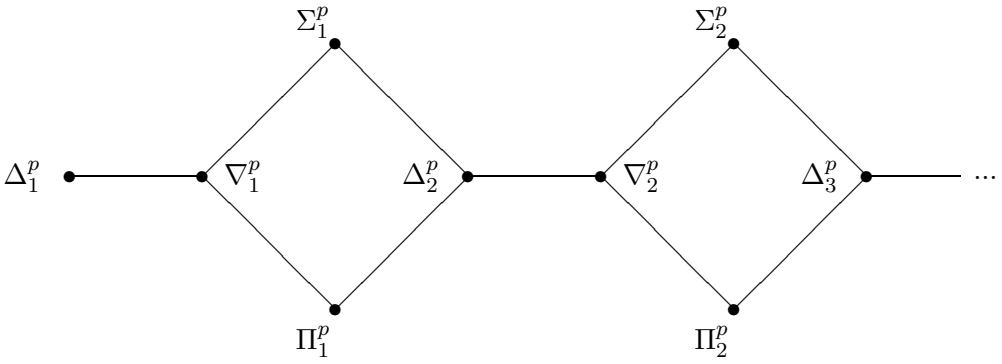

Fig. 3 A hierarquia (do tempo) polinomial

A hierarquia polinomial tem semelhanças superficiais com a hierarquia aritmética: aos conjuntos decidíveis correspondem os conjuntos decidíveis em tempo polinomial; aos conjuntos recursivamente enumeráveis, os conjuntos NP; às quantificações, as quantificações

limitadas; e às fórmulas limitadas da linguagem da aritmética correspondem as sw.q.-fórmulas da linguagem da árvore binária. O que salta à vista quando se comparam as figuras 1 e 3 são as linhas horizontais que ligam os *deltas* com os *nablas*. Não se poderia, também, ter desenhado estas linhas na figura 1 definindo, para isso, as classes dos subconjuntos de ω decidíveis com o auxílio dum oráculo em Σ_n? Claro que sim! O que acontece é que estas classes não acrescentam nada de novo, pois coincidem com as classes Δ_{n+1} ([1] da secção 3 é, novamente, uma boa referência). No caso da hierarquia polinomial as questões "$\nabla_n^p = \Delta_n^p$?" estão em aberto: em particular, para $n = 1$, temos a importante questão "$P = NP \cap co - NP$?". Mas a diferença principal é a seguinte: as inclusões da hierarquia aritmética são todas estritas, enquanto que no caso da hierarquia polinomial não se sabe se elas o são, ou não. Por exemplo, saber se $\Delta_1^p = \Sigma_1^p$ é o problema "$P = NP$?".

A versão funcional da hierarquia polinomial interessa-nos. Uma função de $\{0,1\}^*$ para $\{0,1\}^*$ diz-se na classe $\square_{n+1}^p$ se for computável em tempo polinomial com o auxílio dum oráculo em Σ_n^p; a classe $\square_1^p$ é a classe Ptime das funções computáveis em tempo polinomial. Tem-se o seguinte resultado:

Teorema. *Se $S_2 \vdash \forall x \exists y A(x,y)$, onde $A \in \Sigma_n^b$, então existe uma função $f \in \square_{n+1}^p$ tal que $\forall n A(n, f(n))$.*

Há um certo abuso da notação no enunciado deste resultado. Definimos fórmulas Σ_n^b para a linguagem da teoria da árvore binária, não para a linguagem de S_2: este abuso é admissível, pois existem definições correspondentes para a linguagem de aritmética - remetemos o leitor interessado para a dissertação de Buss. O teorema precedente melhora o teorema inicial desta secção, pois todas as funções da hierarquia (do tempo) polinomial são computáveis em espaço polinomial. A ideia nuclear da demonstração do teorema utiliza um método importante em algoritmia: *dividir para reinar* ("divide and conquer"). Se $S_2 \vdash \forall x \exists y A(x,y)$, com $A \in \Sigma_n^b$, então o Teorema de Parikh permite-nos concluir que existe um termo $t(x)$ tal que $S_2 \vdash \forall x \exists y \leq t(x) A(x,y)$ e, em particular, a asserção "$\forall x \exists y \leq t(x) A(x,y)$" é verdadeira. Ignorando diferenças de linguagem, admitimos que esta asserção está formulada na terminologia da árvore binária. Por meio dum truque do ofício podemos supor, sem perda de generalidade, que $\forall x \exists y (c(y) = c(t(x)) \wedge A(x,y))$. Então o algoritmo,

```
procedure CalcularTestemunha
var x,y palavra
begin
    read x
    y := ϵ
    while c(y) < c(t(x)) do
        begin
            if ∃w (c(y) + c(w) + 1 = c(t(x)) ∧ A(x, y ⌢ 0 ⌢ w)) then y := y ⌢ 0
            else y := y ⌢ 1
        end
end
```

calcula a testemunha y em tempo polinomial, com o auxílio de um oráculo em Σ_n^p. Assim, este algoritmo define uma $\Box_{n+1}^p$-função testemunha da verdade da asserção "$\forall x \exists y A(x,y)$", como se pretendia.

O esquema da indução na notação (*Notation Induction Axiom*) consiste nas asserções da forma,

$$(NIA) \qquad A(\epsilon) \wedge \forall x (A(x) \rightarrow A(x \frown 0) \wedge A(x \frown 1)) \rightarrow \forall x A(x)$$

onde A é uma fórmula da linguagem da árvore binária, possivelmente com parâmetros. A teoria constituida pelos catorze axiomas básicos, juntamente com o esquema (NIA) aplicado a todas as fórmulas da linguagem, define uma teoria equivalente a PA. Estamos interessados, realmente, no caso em que o esquema de indução na notação apenas se aplica a fórmulas limitadas. Se permitirmos todas as fórmulas de Σ_∞^b, ficamos com a teoria Σ_∞^b-NIA, que é equivalente a S_2. Se somente permitirmos indução na notação para fórmulas Σ_n^b obtemos as sub-teorias Σ_n^b-NIA (estas teorias são equivalentes às teorias S_2^n de Buss). Sem nos preocuparmos com diferenças de linguagem, o teorema de Buss reza assim:

Teorema. *Se Σ_n^b-NIA $\vdash \forall x \exists y A(x,y)$, onde $A \in \Sigma_n^b$, então existe uma função $f \in \Box_n^p$ tal que $\forall \sigma A(\sigma, f(\sigma))$.*

(A demonstração original deste teorema deve-se a Buss na sua dissertação de doutoramento. No capítulo V de [5], Pavel Pudlák expõe uma demonstração alternativa - devida a Alex Wilkie - via teoria dos modelos. Este capítulo de Pudlák é a melhor introdução disponível à aritmética limitada. Em [6] apresentamos ainda outra demonstração do teorema, agora no contexto da notação da árvore binária.)

Corolário. *Se Σ_1^b-NIA $\vdash \forall x \exists y A(x,y)$, onde $A \in \Sigma_1^b$, então existe uma função f, computável em tempo polinomial, tal que $\forall \sigma A(\sigma, f(\sigma))$.*

Corolário. *Se Σ_1^b-NIA $\vdash \forall x (A(x) \leftrightarrow B(x))$, onde $A \in \Sigma_1^b$ e $B \in \Pi_1^b$, então o conjunto $X = \{\sigma : A(\sigma)\}$ é decidível em tempo polinomial.*

Este último corolário é uma boa ilustração do dictum *mais demonstrações $\Rightarrow$ mais algoritmos*. A seguinte leitura torna isto patente: se $X \in NP \cap co-NP$ (i.e., se X for simultaneamente definível por fórmulas $A \in \Sigma_1^b$ e $B \in \Pi_1^b$) e *se a equivalência entre A e B se demonstra na teoria Σ_1^b-NIA*, então $X \in P$. A demonstração é simples. Sejam $A(x) = \exists y \leq t(x) A'(x,y)$ e $B(x) = \forall y \leq q(x) B'(x,y)$, onde A' e B' são sw.q.-fórmulas. Por uma parte da hipótese (a implicação da direita para a esquerda), tem-se que Σ_1^b-NIA $\vdash \forall x \exists y (\neg y \leq q(x) \vee B'(x,y) \rightarrow \exists z \leq t(x) A'(x,z))$. Então, pelo primeiro corolário, existe uma função f, computável em tempo polinomial, tal que $\forall \sigma ((\forall \rho \leq q(\sigma) B'(\sigma,\rho)) \rightarrow f(\sigma) \leq t(\sigma) \wedge A'(\sigma, f(\sigma)))$. Agora não é difícil argumentar que $X = \{\sigma : f(\sigma) \leq t(\sigma) \wedge A'(\sigma, f(\sigma))\}$ e, por conseguinte, $X \in P$.

O teorema de Buss não é, porém, inteiramente satisfatório. Com efeito, admitamos que a asserção "$\forall x \exists y \leq t(x) A(x,y)$" (a que chamamos S), onde $A \in \Sigma^b_1$ e t é um termo da linguagem, é verdadeira. Já argumentámos que a verdade de S pode ser testemunhada por meio duma função em $\Box^p_2$; também sabemos, pelo primeiro corolário acima, que se S for consequência da teoria Σ^b_1-NIA então há uma função em Ptime testemunha da sua verdade. Ora, qual é a informação computacional suplementar que se pode extrair do facto da asserção S se demonstrar numa das teorias Σ^b_n-NIA? Já depois da realização do Encontro obtivémos o seguinte resultado:

Teorema. *Se Σ^b_2-NIA $\vdash \forall x \exists y A(x,y)$, onde $A \in \Sigma^b_1$, então existem $B \in \Delta^p_1$ e f, g e h em Ptime tais que,*

(1) $\forall \sigma \forall \rho \, (B(\sigma,\rho) \wedge \neg A(\sigma, g(\sigma,\rho)) \rightarrow f(\sigma,\rho) < \rho \wedge B(\sigma, f(\sigma,\rho)))$

(2) $\forall \sigma \; B(\sigma, h(\sigma))$

Este teorema providencia informação computacional, ainda que não na forma de funções demonstravelmente totais. Sem embargo, pensamos estar na direcção certa e admitimos que seja necessário trabalhar com uma noção mais geral do conceito de testemunha (o próximo papel [7] discute esta temática). Está em aberto a extensão deste resultado às teorias Σ^b_n-NIA, para $n \geq 3$.

A nosso ver, um dos resultados mais interessantes em Aritmética Limitada é o seguinte:

Teorema. *Se Σ^b_n-NIA $= \Sigma^b_{n+1}$-NIA então $\Sigma^p_{n+2} = \Pi^p_{n+2}$ (e, por conseguinte, $\Box^p_{n+3} = \Box^p_{n+4}$).*

Refraseando: se a teoria Σ^b_n-NIA demonstra o esquema da indução na notação para fórmulas Σ^b_{n+1} (*menos axiomas*) então toda a função em $\Box^p_{n+4}$ já figura em $\Box^p_{n+3}$ (*mais eficiência*). Em suma, desde que a hierarquia de teorias Σ^b_1-NIA $\subseteq \Sigma^b_2$-NIA $\subseteq \Sigma^b_3$-NIA $\subseteq$... não seja estrita, então a hierarquia polinomial também não o é. Este resultado deve-se a Jan Krajíček, Pavel Pudlák e Gaisi Takeuti (o resultado original é mais fino - consulte-se [8] ou a exposição na terceira parte de [9]).

Os dicta que enunciámos na introdução necessitam de alguma qualificação. Se não, atentemos à seguinte situação. O *esquema da colecção limitada* consiste nas asserções da forma,

$$(B\Sigma^b_\infty) \qquad \forall u \leq x \exists y A(u,y) \rightarrow \exists w \forall u \leq x \exists y \leq w A(u,y)$$

onde A é uma fórmula limitada, possivelmente com parâmetros. É claro que este esquema é verdadeiro no modelo standard (demonstra-se imediatamente, por indução, na teoria $I\Sigma_1$). Jeff Paris e Laurence Kirby mostraram em [10] que $B\Sigma^b_\infty$ não é consequência da teoria Σ^b_∞-NIA. Na parte final de [11] Alex Wilkie e Jeff Paris levantaram a questão "intrigante" (*sic*) de saber se $B\Sigma^b_\infty$ é consequência de Σ^b_∞-NIA$+\neg exp$. Em [12] apresentamos o seguinte resultado:

Teorema. *Se o esquema* $B\Sigma^b_\infty$ *é consequência da teoria* Σ^b_∞*-NIA* $+ \neg exp$*, então a hierarquia (do tempo) polinomial é estrita.*

Assim, sob a hipótese de haverem certas demonstrações (nomeadamente, sob a suposição do esquema da colecção limitada se demonstrar na teoria Σ^b_∞-NIA$+\neg exp$) conclui-se que certos algoritmos não existem (e.g., conclui-se que $P \neq NP$, o que indica que uma resposta positiva à questão de Wilkie e Paris é muito problemática). Não obstante, observe-se que a hipótese do teorema assevera a existência de certas demonstrações *a partir duma teoria que é falsa* no modelo standard.

Indicações bibliográficas

1. Buss, S., *Bounded Arithmetic*, Ph.D. Dissertation, Princeton University (1985). Uma revisão desta tese foi publicada por Bibliopolis/North-Holland (1986).

2. Håstad, J., *Computational Limitations for Small Depth Circuits*, Ph.D. Dissertation, Massachussetts Institute of Technology (1986).

3. Smale, S., "Theory of Computation", in Casacuberta, C. & Castellet, M. (orgs.), *Mathematical Research Today and Tomorrow*, Springer-Verlag (1992).

4. Garey, M. & Johnson, D., *Computers and Intractability: a Guide to the Theory of NP-completeness*, Freeman (1979).

5. Hájek, P. & Púdlak, P., *Metamathematics of First-Order Arithmetic*, Perspectives in Mathematical Logic, Springer-Verlag (1993).

6. Ferreira, F., "Stockmeyer Induction", in Buss, S. & Scott, P. (orgs.), *Feasible Mathematics*, Progress in Computer Science and Applied Logic 9, Birkhäuser, pp. 161-180 (1990).

7. Ferreira, F., "On $\forall\Sigma^b_1$-consequences of S^2_2" (abstract). Estamos a preparar um papel com este título.

8. Krajíček, J., Pudlák, P. & Takeuti, G., "Bounded Arithmetic and the Polynomial Hierarchy", *Annals of Pure and Applied Logic* 52, pp. 143-154 (1991).

9. Oitavem, I., *Três Assuntos de Lógica e Complexidade*, Dissertação de Mestrado (em preparação).

10. Paris, J. & Kirby, L., "Σ_n-collection Schema in Arithmetic", in Macintyre, A. et al. (orgs.), *Logic Coloquium 1977*, Studies in Logic and the Foundations of Mathematics, North-Holland, pp. 199-209 (1978).

11. Paris, J. & Wilkie, A., "On the Existence of End Extensions of Models of Bounded Arithmetic" in Fenstad J. et al. (orgs.), *Logic, Methodology and Philosophy of Science VIII*, Elsevier, pp. 143-161 (1989).

12. Ferreira, F., "Binary Models Generated by their Tally Part", submetido para publicação.

6 Uma Teoria para a Análise

Em finais de 1985, durante um simpósio sobre o Programa de Hilbert (publicado em [1]), Wilfried Sieg levantou o seguinte problema: *to find a mathematically significant subsystem of analysis whose class of provably recursive functions consists only of the computationally "feasible" ones.* A nossa dissertação de doutoramento [2] aborda este problema, com o entendimento de que "computationally feasible" queira dizer "computável em tempo polinomial". O tratamento que seguimos baseia-se no trabalho de Simpson e Friedman em Matemática Recíproca (veja-se a secção 3): procurámos substituir a teoria base RCA_0, ligada à classe das funções recursivas primitivas, por outra mais fraca de modo a que o papel destas funções seja substituido pelas funções computáveis em tempo polinomial.

As novas teorias formulam-se numa linguagem de segunda-ordem, com a terminologia da árvore binária. A teoria base consiste em Σ^b_1-NIA (admitindo-se também parâmetros de segunda-ordem no esquema de indução na notação) e no postulado:

$$(\nabla^b_1 - CA) \qquad \forall x(A(x) \leftrightarrow B(x)) \rightarrow \exists X \forall x(x \in X \leftrightarrow A(x))$$

onde $A \in \Sigma^b_1$ e $B \in \Pi^b_1$ (permitem-se parâmetros de primeira e de segunda ordem). No modelo standard este postulado afirma que se podem formar os conjuntos de $NP \cap co-NP$ e, por maioria de razão, os conjuntos de P. É relativamente simples argumentar que Σ^b_1-NIA+∇^b_1-CA é uma extensão conservativa da teoria de Σ^b_1-NIA com respeito às asserções de primeira-ordem. Em particular, as funções demontravelmente totais (com gráfico Σ^b_1) desta teoria são as funções computáveis em tempo polinomial.

A diferença mais notável entre o caso clássico da Matemática Recíproca e as novas teorias diz respeito ao papel desempenhado pelo esquema da colecção limitada (*vide* a parte final da secção anterior). Há alguns anos demonstrámos que este esquema e o lema fraco de König estão intimamente ligados. Antes de enunciar o teorema que estabelece esta ligação necessitamos de algumas definições.

Dada uma fórmula A da linguagem de segunda-ordem, com uma variável livre distinta x, denotamos por $\mathrm{Tree}_\infty(A_x)$ a seguinte fórmula,

$$\forall x \forall y(A(x) \wedge y \subseteq x \rightarrow A(y)) \wedge \forall u \exists x \equiv u A(x)$$

onde "$x \equiv u$" abrevia "$x \leq u \wedge u \leq x$" (i.e., x e u têm o mesmo comprimento). Observe-se que x é uma variável muda na fórmula $\mathrm{Tree}_\infty(A_x)$. Seja X uma variável de segunda-ordem; $\mathrm{Path}(X)$ é a fórmula,

$$\mathrm{Tree}_\infty((x \in X)_x) \wedge \forall x \forall y(x \in X \wedge y \in X \rightarrow x \subseteq y \vee y \subseteq x)$$

O lema fraco de König para fórmulas limitadas é o esquema

$$(\Sigma^b_\infty - WKL) \qquad \mathrm{Tree}_\infty(A_x) \rightarrow \exists X(\mathrm{Path}(X) \wedge \forall x(x \in X \rightarrow A(x)))$$

onde A é uma fórmula limitada (permitem-se parâmetros de primeira e segunda ordem) e X é uma nova variável. O seguinte teorema encontra-se em [3]:

Teorema. *Uma asserção de primeira-ordem é consequência da teoria* Σ^b_1*-NIA*$+\nabla^b_1$*-CA* $+\Sigma^b_\infty$*-WKL se, e somente se, for consequência da teoria de primeira-ordem* Σ^b_1*-NIA*$+B\Sigma^b_\infty$.

Ora, sob condições muito gerais, Buss demonstrou que a adjunção do esquema $B\Sigma^b_\infty$ a uma teoria limitada não permite deduzir novas asserções Π_2 (*vide* [4]). Este resultado de Buss, juntamente com o teorema acima, assegura que a teoria Σ^b_1-NIA$+\nabla^b_1$-CA$+\Sigma^b_\infty$-WKL obedece ao requisito de Sieg no que concerne às funções demonstravelmente totais. (De facto, trabalhamos com uma teoria base ligeiramente mais forte do que Σ^b_1-NIA $+\nabla^b_1$-CA, a que démos o nome de *BTFA* - uma sigla para "Base Theory for Feasible Analysis".) Coloca-se com pertinência a questão de saber se BTFA$+\Sigma^b_\infty$-WKL é um subsistema "significativo" da análise. Nesta altura apenas podemos adiantar que estamos a trabalhar nesta questão, nomeadamente no que concerne aos casos do teorema do valor intermédio de Bolzano, do princípio da cobertura de Heine-Borel, do teorema de Cantor da continuidade uniforme e do teorema da existência de máximo para funções contínuas definidas num intervalo compacto (Weierstrass).

Indicações bibliográficas

1. Sieg, W., "Hilbert's Program Sixty Years Later", *The Journal of Symbolic Logic* 53, pp. 338-348 (1988).

2. Ferreira, F., *Polynomial Time Computable Arithmetic and Conservative Extensions*, Ph.D. Dissertation, Pennsylvania State University (1988).

3. Ferreira, F., "A Feasible Theory for Analysis", em publicação no *The Journal of Symbolic Logic*.

4. Ferreira, F., "A Note on a Result of Buss Concerning Bounded Theories and the Collection Scheme", submetido para publicação.

Universidade de Lisboa
Departamento de Matemática
Rua Ernesto de Vasconcelos, Bloco C1, 3
1700 Lisboa
PORTUGAL
(mferferr@ptearn.bitnet)