

O Truque de Rosser

Rogério Augusto dos Santos Fajardo

Resumo

Este texto – voltado para a disciplina de Lógica para graduação e pós-graduação – complementa a prova do 1º Teorema da Incompletude de Gödel-Rosser apresentado no capítulo 7 de [1], ao apresentar o argumento de Rosser – conhecido como *truque de Rosser* – para provar a versão do teorema que troca a hipótese de ω -consistência por apenas consistência.

1 Preliminares

Vamos retomar a notação feita em [1] para a prova do Teorema de Gödel para Lógica de Primeira Ordem.

Partimos de uma teoria $T = (L, \Gamma)$ – onde L é uma linguagem de primeira ordem e Γ é um conjunto de sentenças – recursivo e capaz de expressar a aritmética. A definição formal de sistema recursivo pode ser encontrada em [1] e [2], mas podemos compreender de forma intuitiva da seguinte forma: existe um algoritmo para decidir se uma sequência de símbolos é uma fórmula e se uma sequência de fórmulas é uma demonstração.

- $t(n)$ é o termo constante correspondente a n (isto é, $t(0)$ é a constante 0 e $t(n+1)$ é $t(n) + 1$).
- $A[t]$ é a fórmula obtida pela substituição de todas as ocorrências livres das variáveis em A pelo termo t .
- $g(A)$ é o número de Gödel da fórmula A .
- $g^*(S)$ é o número de Gödel da sequência de fórmulas S .
- Φ_n é a sequência de símbolos do alfabeto de L cujo número de Gödel é n (isto é, $g(\Phi_n) = n$).
- $A(t, s, u)$ é a fórmula obtida substituindo todas as ocorrências livres das variáveis x, y e z em A pelos termos t, s e u , respectivamente.
- $\overline{D}$ é o conjunto das triplas $(p, m, k) \in \mathbb{N}^3$ tais que p é o número de Gödel de uma demonstração de $\Phi_m[t(k)]$ a partir de Γ .
- $\overline{D'}$ é o conjunto das triplas $(p, m, k) \in \mathbb{N}^3$ tais que p é o número de Gödel de uma demonstração de $\neg\Phi_m[t(k)]$ a partir de Γ .

O último item – a definição de $\overline{D'}$ – não consta em [1] e foi adicionado para explicar o argumento de Rosser.

Iremos assumir que o universo da nossa teoria é o conjunto dos números naturais. Assim, quando escrevemos $\exists xA$ queremos dizer que “Existe um número natural x tal que A ”. Da mesma forma, $\forall xA$ significa “Para todo número natural x vale A ”. Essa hipótese, no entanto, é só para simplificar a escrita. Se T é uma teoria mais abrangente que a aritmética – como, por exemplo, a teoria dos conjuntos de ZFC – podemos adaptar tudo que será feito neste texto com as seguintes conversões, sendo $N(x)$ uma fórmula que significa “ x é um número natural”:

- Onde estiver escrito $\exists xA$ lê-se $\exists x(N(x) \wedge A)$.
- Onde estiver escrito $\forall xA$ lê-se $\forall x(N(x) \rightarrow A)$.

Lema 1 *Existem fórmulas D e D' da linguagem L que possuem x, y e z como as únicas variáveis livres e tais que, para todos $(p, m, k) \in \mathbb{N}^3$, temos*

- Se $(p, m, k) \in \overline{D}$ então $\Gamma \vdash D(t(p), t(m), t(k))$;
- Se $(p, m, k) \notin \overline{D}$ então $\Gamma \vdash \neg D(t(p), t(m), t(k))$;
- Se $(p, m, k) \in \overline{D'}$ então $\Gamma \vdash D'(t(p), t(m), t(k))$;
- Se $(p, m, k) \notin \overline{D'}$ então $\Gamma \vdash \neg D'(t(p), t(m), t(k))$.

⊢

Demonstração Essa é a parte mais trabalhosa da prova dos teoremas de incompletude e não faremos aqui. Está enunciada na Proposição V de [2], só com as duas primeiras partes, e em [3], na forma aqui apresentada.

A ideia da demonstração é codificar na aritmética, usando que o sistema é recursivo, a relação “ x é o número de Gödel da demonstração da fórmula cujo número de Gödel é y ”, mostrando que essa relação é recursiva (equivalentemente, Turing-computável). Para isso é necessário codificar cada axioma e regra de inferência como uma propriedade aritmética. ■

Introduzimos mais uma notação: se A é a fórmula $\Phi_m(t(k))$, denotamos por

- $Dem(x, \ulcorner A \urcorner)$ a fórmula $D(x, t(m), t(k))$;
- $Dem(x, \ulcorner \neg A \urcorner)$ a fórmula $D'(x, t(m), t(k))$;

Ou seja, $Dem(x, \ulcorner B \urcorner)$ significa “ x é o número de Gödel da prova de B ”, tanto no caso em que B é A quanto no caso em que B é $\neg A$.

Alguém pode reclamar que a definição dada aqui é ambígua, pois se $\neg A$ é a fórmula $\Phi_m(t(k))$ e A é $\Phi_i(t(j))$, então $Dem(x, \ulcorner \neg A \urcorner)$ pode significar tanto $D(x, t(m), t(k))$ quanto $D'(x, t(i), t(j))$. No entanto, nos contextos em que utilizaremos essa notação fica claro qual fórmula estamos considerando, e o ganho em clareza compensa essa pequena imprecisão.

Lembramos que a fórmula G construída por Gödel é definida como

$$\neg \exists x D(x, t(n), t(n)),$$

onde $n = g(\neg \exists x D(x, y, y))$. Assim, $\Phi_n(t(n))$ é a própria fórmula G e, portanto, G coincide com a fórmula $\neg \exists x Dem(x, \ulcorner G \urcorner)$. Isto é, G diz “Eu não posso ser provada”.

Assumindo que $T \vdash G$, podemos concluir que $T \vdash \neg G$, de onde segue que T é inconsistente. Essa parte não usa a hipótese de T ser ω -consistente. Porém, a outra direção – se $T \vdash \neg G$, então $T \vdash G$ – requer a hipótese de que T é ω -consistente.

2 A sentença de Rosser

Como assumimos que a linguagem é capaz de expressar a aritmética, podemos também que essa possui o símbolo relacional $\leq$. Ou entendemos $x \leq y$ como abreviatura de $\exists z (x + z = y)$.

Também iremos considerar uma fórmula da forma $\exists x \leq y A$ como abreviatura de $\exists x (x \leq y \wedge A)$. Analogamente, $\forall x \leq y A$ significa $\forall x (x \leq y \rightarrow A)$.

Definição 2 Definimos r o número de Gödel da fórmula

$$\forall x (D(x, y, y) \rightarrow \exists z \leq x D'(z, y, y))$$

e R a fórmula

$$\forall x (D(x, t(r), t(r)) \rightarrow \exists z \leq x D'(z, t(r), t(r))).$$

⊥

Usando a notação Dem , podemos notar que

$$R \equiv \forall x (Dem(x, \ulcorner R \urcorner) \rightarrow \exists z \leq x Dem(z, \ulcorner \neg R \urcorner)).$$

Ou seja, a fórmula de Rosser (que é uma sentença, pois não possui variáveis livres), que estamos denotando por R , diz o seguinte:

Se eu posso ser provada com uma prova de número n , então minha negação pode ser provada com uma prova de número menor do que n .

Se a teoria for consistente, a fórmula de Rosser será indecidível. Esboçemos aqui a ideia intuitiva desse argumento. Se o sistema prova R , então existe um número natural n que codifica essa prova e, pela validade de R , existe $m \leq n$ que codifica uma prova da negação de R . Reciprocamente, se o sistema prova $\neg R$, então existe um número natural n tal que a implicação

$$Dem(n, \ulcorner R \urcorner) \rightarrow \exists z \leq n Dem(z, \ulcorner \neg R \urcorner)$$

é falsa. Em particular, o antecedente da implicação é verdadeiro. Isto é, R pode ser provada.

A formalização desse argumento, no entanto, requer um cuidado maior. Note que no argumento final, estamos presumindo que, se provamos a existência de um n que “prova” R , então conseguimos de fato esse n para codificar a prova de R . Mas, aparentemente, aqui estamos precisando novamente da hipótese de ω -consistência, como na prova feita por Gödel, pois n poderia ser não *standard*. Não é tão trivial perceber por que a fórmula de Rosser funciona melhor que a de Gödel

3 O Teorema da Incompletude de Gödel-Rosser

Antes de provarmos o teorema principal, vamos provar alguns lemas com especial cuidado em relação ao quantificador existencial. Lembremos que todo o problema na hora de formalizar o argumento original de Gödel é considerar uma fórmula existencial e escolher um número natural que “testemunha” a validade da fórmula. Se o sistema for consistente mas ω -inconsistente, podemos não encontrar um número *standard* que valide a fórmula. Isto é, podemos provar $\exists x P(x)$ mas não provarmos $P(0)$, $P(1)$ etc. (e até o contrário: provarmos $\neg P(0)$, $\neg P(1)$ etc).

Em alguns passos da demonstração feita por Rosser, parece que precisamos novamente usar assumir a ω -consistência do sistema para encontrar essa “testemunha para a fórmula existencial”, de modo que o truque de Rosser pode parecer malogrado em seu propósito. Mas, essencialmente, o que muda no argumento de Rosser em relação ao de Gödel, é que os existenciais estão limitados a números naturais *standard*, e por isso podem ser realizados por outro natural *standard*.

Assim, mesmo que, neste texto, de modo geral, não tenha muitos detalhes lógicos, onde o fantasma dos naturais não-*standard* assombrar, fazendo o leitor desconfiar – com toda razão – de argumentos utilizando o quantificador existencial, agiremos com excesso de cautela.

Lembramos a seguinte notação: se escrevemos uma fórmula A como $A(x)$, então $A(t)$ significa $[A]_x^t$. Quando A tem uma única variável livre, $A(t)$ significa $[A]_x^t$, onde x é a única variável livre de A (ou seja, nesse caso, não precisamos, antes, destacar a variável x escrevendo $A(x)$ no lugar de A).

Lema 3 *Se $A(x)$ é uma fórmula e n um número natural, então*

$$T \vdash (\exists x \leq t(n) A) \leftrightarrow (A(t(0)) \vee A(t(1)) \vee \dots \vee A(t(n))).$$

Demonstração: Vamos fazer a prova por indução em n , na metalinguagem. Na prática, quando utilizamos o lema, é para um n específico (e *standard*), de modo que o argumento pode ser feito finitariamente, sem usar indução.

Usando que $x \leq 0$ é equivalente a $x = 0$, o passo inicial da indução é provarmos que

$$(1) \quad T \vdash (\exists x (x = 0 \wedge A)) \leftrightarrow A(0).$$

Façamos primeiro a recíproca. Como $0 = 0$ é sempre verdadeiro, temos que $A(0) \rightarrow (0 = 0) \wedge A(0)$ é um teorema lógico. Mas $(0 = 0) \wedge A(0) \rightarrow \exists x (x = 0 \wedge A(0))$ também é teorema (veja [1, Teorema 6.25]). Portanto, concluímos que

$$T \vdash A(0) \rightarrow \exists x (x = 0 \wedge A).$$

Para a outra direção, começamos com o seguinte teorema lógico (veja [1, Teorema 6.15]):

$$T \vdash (x = 0) \rightarrow (A(x) \rightarrow A(0)).$$

Usando uma simples equivalência tautológica (verificável com tabela-verdade) obtemos

$$T \vdash (\neg A(0)) \rightarrow ((x = 0) \rightarrow \neg A(x)).$$

Como $A(0)$ não tem ocorrência livre da variável x , provamos que

$$T \vdash (\neg A(0)) \rightarrow \forall x((x = 0) \rightarrow \neg A(x)).$$

A contrapositiva dessa fórmula é justamente a implicação que faltava provar na expressão (1).

Agora suponhamos que, para um termo t e uma fórmula B (que será $(A(t(0)) \vee \dots \vee A(t(n)))$, quando t for $t(n)$), temos

$$(2) \quad T \vdash (\exists x((x \leq t) \wedge A)) \leftrightarrow B.$$

Por uma propriedade da aritmética, que assumimos ser válida em T , temos

$$T \vdash (x \leq t + 1) \leftrightarrow (x \leq t \vee x = t + 1).$$

Em particular, usando a distributividade da lógica proposicional podemos provar, em T , que $(x \leq t + 1) \wedge A$ é equivalente a $((x \leq t) \wedge A) \vee ((x = t + 1) \wedge A)$. Portanto

$$(3) \quad T \vdash \exists x((x \leq t + 1) \wedge A) \leftrightarrow \exists x(((x \leq t) \wedge A) \vee \exists x((x = t + 1) \wedge A)).$$

Repetindo o argumento usado na prova de (1) concluímos que

$$(4) \quad T \vdash (\exists x(x = t + 1 \wedge A)) \leftrightarrow A(t + 1).$$

Substituindo algumas subfórmulas de (3) por suas equivalentes dadas por (2) e (4) obtemos

$$T \vdash \exists x((x \leq t + 1) \wedge A) \leftrightarrow (B \vee A(t + 1)),$$

que é o que queríamos provar. ■

Convém reforçar que apesar desse lema ser um esquema infinito de teoremas, o fato de a usarmos não significa que estamos fazendo uma demonstração infinita, visto que a usaremos apenas em instâncias específicas.

O próximo lema seria muito fácil provar se nos permitíssemos utilizar uma forma coloquial muito comum no discurso matemático: se eu provo que existe um número x satisfazendo uma determinada fórmula, diz-se comumente “pegue (ou tome/seja/escolha) x tal que...” e a partir daí trabalhamos com essa variável como se fosse uma constante com a propriedade assumida. Já justificamos nosso receio com as conclusões apresadas envolvendo o quantificador existencial, e aqui novamente seremos cautelosos.

Lema 4 *Sejam A e B fórmulas e n um número natural. Suponha que $T \vdash \exists x(A \wedge \forall y \leq x B(y))$ e $T \vdash \neg B(t(n))$. Então*

$$T \vdash \exists x < t(n)(A \wedge \forall y \leq x B(y)).$$

Demonstração: Pelo axioma A3 da lógica de primeira ordem (veja [1]), temos

$$T \vdash \forall y(y \leq x \rightarrow B(y)) \rightarrow (t(n) \leq x \rightarrow B(t(n))).$$

Da hipótese de que $T \vdash \neg B(t(n))$, usando alguma instância de tautologia adequada e *modus ponens* obtemos

$$T \vdash \forall(y \leq x \rightarrow B(y)) \rightarrow \neg(t(n) \leq x).$$

Usando generalização e a definição de $\forall y \leq x B(y)$ obtemos

$$T \vdash \forall x((\forall y \leq x B(y)) \rightarrow \neg(t(n) \leq x)),$$

o que equivale a

$$T \vdash \neg \exists x(t(n) \leq x \wedge \forall y \leq x B(y)).$$

Em particular

$$T \vdash \neg \exists x(t(n) \leq x \wedge A \wedge \forall y \leq x B(y)).$$

Agora considere C a fórmula $A \wedge \forall y \leq x B(y)$. Mostramos que $T \vdash \neg \exists x(t(n) \leq x \wedge C)$ e, pela hipótese do lema, temos $T \vdash \exists x C$. Note que

$$\begin{aligned} \exists x C &\equiv \exists x(((t(n) \leq x) \wedge C) \vee (\neg(t(n) \leq x) \wedge C)) \equiv \\ &\equiv \exists x((t(n) \leq x) \wedge C) \vee \exists x(\neg(t(n) \leq x) \wedge C). \end{aligned}$$

Como provamos que a primeira parte dessa última disjunção é falsa, concluímos que a segunda é verdadeira. Isto é:

$$T \vdash \exists x((\neg(t(n) \leq x)) \wedge A \wedge \forall y \leq x B(y))$$

Usando a propriedade de linearidade dos números naturais, $\neg(t(n) \leq x)$ é equivalente a $x \leq t(n)$, o que conclui a prova do lema. ■

Teorema 5 *Se uma teoria T é consistente, recursiva e capaz de expressar a aritmética, então T é incompleta.*

Demonstração Considere R a sentença de Rosser, já descrita na seção anterior. Suponha que $T \vdash R$ e seja n o número de Gödel de uma demonstração de R . Isto é, temos $(n, r, r) \in \overline{D}$, onde r é como na Definição 2. Pelo Lema 1, temos $T \vdash D(t(n), t(r), t(r))$. Retomando a definição das fórmulas Dem e R , temos

$$(1) \quad T \vdash Dem(t(n), \ulcorner R \urcorner).$$

Por outro lado, da hipótese $T \vdash R$ e usando o axioma A3 e Modus Ponens, temos

$$(2) \quad T \vdash Dem(t(n), \ulcorner R \urcorner) \rightarrow \exists x \leq t(n) Dem(x, \ulcorner \neg R \urcorner).$$

Logo, de (1) e Modus Ponens concluímos que

$$(3) \quad T \vdash \exists x \leq t(n) Dem(x, \ulcorner \neg R \urcorner).$$

Portanto, do Lema 3, obtemos

$$(4) \quad T \vdash Dem(t(0), \ulcorner \neg R \urcorner) \vee \dots \vee Dem(t(n), \ulcorner \neg R \urcorner).$$

Suponha que $T \not\vdash \neg R$. Isso significa que, para cada número natural $k \leq n$, $(k, r, r) \notin \overline{D'}$. Portanto, pelo Lema 1,

$$(5) \quad T \vdash \neg Dem(t(0), \ulcorner \neg R \urcorner) \wedge \dots \wedge \neg Dem(t(n), \ulcorner \neg R \urcorner),$$

que é justamente a negação da fórmula (4), provando que T é inconsistente.

Agora suponha que $T \vdash \neg R$. Isto é:

$$(6) \quad T \vdash \exists x (Dem(x, \ulcorner R \urcorner) \wedge \forall y \leq x \neg Dem(y, \ulcorner \neg R \urcorner)).$$

Seja n o menor número natural que codifica uma prova de $\neg R$. Pela afirmação 1,

$$(7) \quad T \vdash Dem(t(n), \ulcorner \neg R \urcorner).$$

De (6) e (7) e o Lema 4 concluímos que

$$(8) \quad T \vdash \exists x \leq t(n) Dem(x, \ulcorner R \urcorner).$$

Assim, repetindo o argumento das expressões (4) e (5) ao assumir que $T \not\vdash R$, contradizendo que T é consistente.

Portanto, se T é consistente, R é indecidível em T . ■

4 Considerações finais

Tanto a sentença G , de Gödel, quanto a sentença R , de Rosser, faz uma autorreferência afirmando ser ela própria impossível de ser demonstrada. Se o sistema for consistente, ele não pode provar G nem R (essa parte não assume ω -consistência, mesmo na versão de Gödel).

Como esse argumento pode ser formalizado no próprio sistema, pois é uma afirmação aritmética, se T prova sua consistência (que pode ser formalizada pela fórmula $\neg \exists x D(x, \ulcorner 0 = 1 \urcorner, 0)$, que iremos denotar por $Con(T)$), concluímos que T prova que, se ele próprio for consistente, então G (e também R) não pode ser provada. No caso da fórmula de Rosser, temos, então:

$$T \vdash Con(T) \rightarrow \neg \exists x Dem(x, \ulcorner R \urcorner).$$

Logo, se T prova sua própria consistência, por *modus ponens* $T \vdash \neg \exists x Dem(x, \ulcorner R \urcorner)$. Ora, mas isso significa que o antecedente da implicação na sentença de Rosser é sempre falsa, o que implica que R é verdadeira. Ou seja, nesse caso, temos que $T \vdash R$ e, como já provado, disso segue que T é inconsistente.

Portanto, mesmo que no segundo teorema de incompletude de Gödel assumimos apenas a consistência – isto é, no segundo teorema a sentença de Rosser não “melhora” o resultado –, podemos usar R no lugar de G para obter a mesma conclusão.

Temos, portanto, que se T é consistente, então $T \not\vdash G$, $T \not\vdash R$ e $T \not\vdash \neg R$. O que não podemos concluir, sem a hipótese adicional de ω -consistência, que $T \not\vdash \neg G$.

Apesar do segundo teorema de incompletude assegurar que um sistema (sempre estamos considerando recursivo e capaz de expressar a aritmética) consistente prova sua própria consistência, não podemos garantir que a própria fórmula $Con(T)$ é indecidível em T . Isso porque *um sistema consistente pode provar sua própria inconsistência*. De fato, pelo segundo teorema de incompletude e por um corolário do Teorema da Dedução, a teoria T' – definida como T acrescida da sentença $\neg Con(T)$ como axioma – é consistente. É evidente que a inconsistência de T implica a inconsistência de T' , pois uma prova de $0 = 1$ a partir dos axiomas de T continua sendo uma prova da mesma fórmula em T' . Mas, mais do que isso, podemos formalizar essa implicação como uma propriedade aritmética sobre os números de Gödel das provas e das fórmulas. Logo, $T' \vdash \neg Con(T) \rightarrow \neg Con(T')$ e, por *modus ponens*, $T' \vdash \neg Con(T')$.

Nesse caso, T' prova que existe um número que é código de uma prova de $0 = 1$, mas também prova que 0 não é código de uma prova de $0 = 1$, que 1 também não é, que 2 não é, etc. Ou seja, o número natural que testemunha que existe uma prova de $0 = 1$ é *não-standard*, e por isso não conseguimos decodificá-lo para obter de fato uma prova da inconsistência de T' .

Mas isso não funciona ao adicionarmos G ou R (ou sua negação) ao sistema T . Isso porque a construção da fórmula de Gödel ou de Rosser depende do sistema T . Ou seja, se considerarmos o sistema T' como T adicionado a R como axioma, o “ R do sistema T é diferente do R do sistema T' ”. Então o “antigo” R passa a ser teorema do sistema, mas surge um “novo” R codificado nesse outro sistema que é indecidível em T' ,

Referências

- [1] Fajardo, R. *Lógica Matemática*. Edusp, 2017.
- [2] Gödel, K. *On formally Undecidable Propositions of Principia Mathematica and Related Systems* [1931]. Nova York, Dover Publications, 1992.
- [3] Rosser, B. *Extensions of Some Theorems of Gödel and Church*. J. Symb. Logic 1, nº 3, (1936), 87–91.