O TEOREMA DE HERBRAND

FERNANDO FERREIRA

Definição 1. *Uma fórmula ϕ duma linguagem do cálculo de predicados diz-se* universal *se é da forma $\forall x_1 \cdots \forall x_n \psi$, onde ψ é uma fórmula sem quantificadores.*

As fórmula universais também são conhecidas por fórmulas Π_1. Seja $\mathcal{L}$ uma linguagem com pelo menos uma constante e Γ um conjunto de fórmulas fechadas universais. Define-se a *expansão de Herbrand* de Γ como o seguinte conjunto $\mathscr{H}(\Gamma)$ de fórmulas fechadas, sem quantificadores, de $\mathcal{L}$:

$$\{\psi(t_1, \ldots, t_n) : t_1, \ldots, t_n \text{ termos fechados, } \psi \text{ sem quantificadores e } \forall x_1 \cdots \forall x_n\, \psi(x_1, \ldots, x_n) \in \Gamma\}$$

Podemos considerar $\mathscr{H}(\Gamma)$ como um conjunto de fórmulas do cálculo proposicional, tendo como letras proposicionais as fórmulas atómicas fechadas de $\mathcal{L}$:

Teorema 1. *Seja Γ um conjunto de fórmulas universais fechadas duma linguagem $\mathcal{L}$ com pelo menos uma constante. O conjunto Γ é satisfazível se, e somente se, a expansão de Herbrand $\mathscr{H}(\Gamma)$ é proposicionalmente satisfazível.*

Demonstração. Suponhamos que Γ tem um modelo $\mathfrak{M}$. Definimos uma valoração proposicional v da seguinte forma: dada ψ uma fórmula atómica fechada de $\mathcal{L}$, dizemos que $v(\psi) = 1$ se, e somente se, $\models_{\mathfrak{M}} \psi$. Facilmente se vê, por indução na complexidade das fórmulas, que se ϕ é uma combinação Booleana de fórmulas atómicas fechadas (i.e., ϕ obtém-se das fórmulas atómicas fechadas por meio de $\neg$, $\rightarrow$, $\wedge$ e $\vee$), então $\tilde{v}(\phi) = 1$ sse $\models_{\mathfrak{M}} \phi$. Conclui-se imediatamente que todas as fórmulas de $\mathscr{H}(\Gamma)$ são verdadeiras para esta valoração v.

Reciprocamente, suponhamos que $\mathscr{H}(\Gamma)$ é proposicionalmente satisfazível, i.e., que existe uma valoração v que torna todas as fórmulas proposicionais de $\mathscr{H}(\Gamma)$ verdadeiras. Vamos definir uma interpretação $\mathfrak{M}$ de $\mathcal{L}$. O domínio de $\mathfrak{M}$ é, por definição, o conjunto de todos os termos fechados de $\mathcal{L}$. Dada c uma constante da linguagem $\mathcal{L}$, define-se $c^{\mathfrak{M}}$ como sendo a própria constante c. Dado f um símbolo funcional n-ário de $\mathcal{L}$, define-se $f^{\mathfrak{M}}$ assim: $f^{\mathfrak{M}}(t_1, \ldots, t_n)$ é o termo $f(t_1, \ldots, t_n)$. Vê-se facilmente, por indução na complexidade dos termos fechados t, que $t^{\mathfrak{M}}$ é o próprio termo t. Finalmente, dado R um símbolo relacional n-ário define-se a relação n-ária $R^{\mathfrak{M}}$ da seguinte forma:

$$(t_1, \ldots, t_n) \in R^{\mathfrak{M}} \quad \text{sse} \quad v(R(t_1, \ldots, t_n)) = 1.$$

Mostra-se, por indução na complexidade fórmulas, que se ψ é uma fórmula sem quantificadores de $\mathcal{L}$, então $\models_{\mathfrak{M}} \psi$ sse $\tilde{v}(\psi) = 1$.

Seja $\forall x_1 \cdots \forall x_n \psi(x_1, \ldots, x_n) \in \Gamma$, onde ψ é uma fórmula sem quantificadores. Vamos argumentar que se tem $\models_{\mathfrak{M}} \forall x_1 \cdots \forall x_n \psi(x_1, \ldots, x_n)$. Sejam $t_1, \ldots, t_n$ termos fechados de $\mathcal{L}$. Queremos ver que $\models_{\mathfrak{M}} \psi(x_1, \ldots, x_n)\,[t_1, \ldots, t_n]$. Ora, pelo corolário do lema da substituição, isto é equivalente a mostrar $\models_{\mathfrak{M}} \psi(t_1, \ldots, t_n)$. Como $\psi(t_1, \ldots, t_n) \in \mathscr{H}(\Gamma)$ vem, por hipótese, $\tilde{v}(\psi(t_1, \ldots, t_n)) = 1$. Daqui sai $\models_{\mathfrak{M}} \psi(t_1, \ldots, t_n)$, como se queria. $\square$

Tem-se o seguinte resultado importante:

Teorema de Herbrand. *Seja $\exists y \phi(y)$ uma fórmula fechada de $\mathcal{L}$, logicamente válida, onde ϕ não tem quantificadores. Admita-se que $\mathcal{L}$ tem pelo menos uma constante. Então existem termos fechados $t_1, \ldots, t_n$ de $\mathcal{L}$ tais que a disjunção*

$$\phi(t_1) \vee \cdots \vee \phi(t_n)$$

é uma tautologia (onde se consideram, como letras proposicionais, as fórmulas atómicas fechadas de $\mathcal{L}$). Consequentemente, é uma verdade lógica.

Demonstração. Como $\exists y\phi(y)$ é uma verdade lógica, conclui-se que a fórmula fechada $\forall y\neg\phi(y)$ de $\mathcal{L}$ não é satisfazível. Seja Γ o conjunto singular $\{\forall y\neg\phi(y)\}$. Pelo teorema anterior, o conjunto de fórmulas $\mathscr{H}(\Gamma) = \{\neg\phi(t) : t \text{ é termo fechado}\}$ não é proposicionalmente satisfazível. Pelo teorema da compacidade do cálculo proposicional, $\mathscr{H}(\Gamma)$ não é finitamente satisfazível. Logo, existe um número finito de termos fechados $t_1, \ldots, t_n$ tais que $\{\neg\phi(t_1), \ldots, \neg\phi(t_n)\}$ não é proposicionalmente satisfazível. Conclui-se que $\phi(t_1) \vee \cdots \vee \phi(t_n)$ é uma tautologia. □

Eis um exemplo. A fórmula $\exists y(P(c, y) \vee \neg P(c, g(y)))$ é logicamente válida. Neste caso o par de termos c e $g(c)$ dá origem à tautologia

$$P(c, c) \vee \neg P(c, g(c)) \vee P(c, g(c)) \vee \neg P(c, g(g(c))).$$

Note-se que os dois termos fechados que obtivemos foram construídos a partir das constantes e símbolos funcionais que ocorrem na fórmula da partida. Uma reflexão rápida mostra que isto é sempre assim (com a possível adição duma constante nova caso na fórmula de partida não ocorram constantes).

O recíproco do teorema de Herbrand é uma trivialidade. Com efeito, se uma disjunção do tipo $\phi(t_1) \vee \cdots \vee \phi(t_n)$, com $t_1, \ldots, t_n$ termos fechados de $\mathcal{L}$, é uma tautologia então a fórmula $\exists y\phi(y)$ é verdadeira em todas as estruturas.

O teorema de Herbrand pode generalizar-se de várias maneiras. Uma extensão imediata consiste em considerar mais do que uma variável existencial. Suponhamos que $\exists y_1 \cdots \exists y_k \phi(y_1, \ldots, y_k)$ é uma verdade lógica expressa numa linguagem com pelo menos uma constante. Um raciocínio semelhante ao acima permite concluir a existência de termos fechados $t_{i,1}, \ldots, t_{i,k}$, com $1 \leq i \leq n$, tais que a disjunção

$$\bigvee_{i=1}^{n} \phi(t_{i,1}, \ldots, t_{i,k})$$

é uma tautologia.

Outra forma de generalizar o teorema de Herbrand consiste em permitir que $\exists y\phi(y)$ seja consequência dum conjunto de fórmulas universais fechadas. A definição de consequência é análoga à do cálculo proposicional:

Definição 2. *Seja Γ um conjunto de fórmulas fechadas duma determinada linguagem. Uma fórmula fechada ϕ dessa linguagem diz-se* consequência semântica *de Γ, e escreve-se $\Gamma \models \phi$, se para toda a interpretação $\mathfrak{M}$ da linguagem, ϕ é verdadeira sob $\mathfrak{M}$ sempre que todas as fórmulas de Γ são verdadeiras sob $\mathfrak{M}$. Se $\Gamma = \emptyset$, escreve-se $\models \phi$ (neste caso, ϕ é uma verdade lógica).*

Seja Γ um conjunto de fórmulas universais fechadas e suponhamos que $\Gamma \models \exists y\phi(y)$, onde ϕ não tem quantificadores. Neste caso, o conjunto de fórmulas universais $\Gamma \cup \{\forall y\neg\phi(y)\}$ não é satisfazível. Pelo raciocínio da demonstração do teorema de Herbrand não é difícil concluir que existem:

i. fórmulas $\psi_1(\underline{z_1}), \ldots, \psi_r(\underline{z_r})$ (admitem-se repetições) sem quantificadores tais que $\forall\underline{z_1}\psi_1(\underline{z_1}), \ldots, \forall\underline{z_r}\psi_r(\underline{z_r}) \in \Gamma$;
ii. termos fechados $\underline{q_1}, \ldots, \underline{q_r}$; e
iii. termos fechados $t_1, \ldots, t_n$,

tais que o condicional

$$\psi_1(\underline{q_1}) \wedge \cdots \wedge \psi_{r-1}(\underline{q_r}) \to \phi(t_1) \vee \cdots \vee \phi(t_n),$$

é uma tautologia (em que se consideram, como letras proposicionais, as fórmulas atómicas fechadas de $\mathcal{L}$). Observe-se que $\psi_1(\underline{q_1}), \ldots, \psi_r(\underline{q_r})$ são instanciações de fórmulas (universais) de Γ. Consequentemente,

$$\Gamma \models \phi(t_1) \vee \cdots \vee \phi(t_n).$$