

# Lógica informática (2008–09)

## Tema 10: Modelos de Herbrand

José A. Alonso Jiménez  
María J. Hidalgo Doblado

Grupo de Lógica Computacional  
Departamento de Ciencias de la Computación e I.A.  
Universidad de Sevilla

## Tema 10: Modelos de Herbrand

1. Modelos de Herbrand
2. Teorema de Herbrand y decisión de la consistencia

## Tema 10: Modelos de Herbrand

### 1. Modelos de Herbrand

Reducción de la LPO básica a proposicional

Universo de Herbrand

Base de Herbrand

Interpretaciones de Herbrand

Modelos de Herbrand

### 2. Teorema de Herbrand y decisión de la consistencia

## Reducción de la LPO básica a proposicional

- ▶ Observación:
  - ▶ En este tema sólo se consideran lenguajes de primer orden sin igualdad.
- ▶ Reducción de la LPO básica a proposicional
  - ▶ Def.: Una **fórmula básica** es una fórmula sin variables ni cuantificadores.
  - ▶ Prop.: Sea  $S$  un conjunto de fórmulas básicas. Son equivalentes:
    1.  $S$  es consistente en el sentido de la lógica de primer orden.
    2.  $S$  es consistente en el sentido de la lógica proposicional.

## Ejemplos de reducción de la LPO básica a proposicional

- ▶  $\{P(a) \vee P(b), \neg P(b) \vee P(c), P(a) \rightarrow P(c)\}$   
es consistente en el sentido de la lógica de primer orden.
- ▶  $\{P(a) \vee P(b), \neg P(b) \vee P(c), P(a) \rightarrow P(c), \neg P(c)\}$   
es inconsistente en el sentido de la lógica de primer orden.

	$P^I$	$P(a) \vee P(b)$	$\neg P(b) \vee P(c)$	$P(a) \rightarrow P(c)$	$\neg P(c)$
$\mathcal{I}_1$	$\emptyset$	0	1	1	1
$\mathcal{I}_2$	$\{c^I\}$	0	1	1	0
$\mathcal{I}_3$	$\{b^I\}$	1	0	1	1
$\mathcal{I}_4$	$\{b^I, c^I\}$	1	1	1	0
$\mathcal{I}_5$	$\{a^I\}$	1	1	0	1
$\mathcal{I}_6$	$\{a^I, c^I\}$	1	1	1	0
$\mathcal{I}_7$	$\{a^I, b^I\}$	1	0	0	1
$\mathcal{I}_8$	$\{a^I, b^I, c^I\}$	1	1	1	0

## Ejemplos de reducción de la LPO básica a proposicional

- ▶  $\{P(a) \vee P(b), \neg P(b) \vee P(c), P(a) \rightarrow P(c)\}$   
es consistente en el sentido proposicional (con modelos  $\mathcal{I}_4, \mathcal{I}_6, \mathcal{I}_8$ ).
- ▶  $\{P(a) \vee P(b), \neg P(b) \vee P(c), P(a) \rightarrow P(c), \neg P(c)\}$   
es inconsistente en el sentido proposicional.

Se consideran los cambios  $P(a)/p, P(b)/q, P(c)/r$

	$p$	$q$	$r$	$p \vee q$	$\neg q \vee r$	$p \rightarrow r$	$\neg r$
$\mathcal{I}_1$	0	0	0	0	1	1	1
$\mathcal{I}_2$	0	0	1	0	1	1	0
$\mathcal{I}_3$	0	1	0	1	0	1	1
$\mathcal{I}_4$	0	1	1	1	1	1	0
$\mathcal{I}_5$	1	0	0	1	1	0	1
$\mathcal{I}_6$	1	0	1	1	1	1	0
$\mathcal{I}_7$	1	1	0	1	0	0	1
$\mathcal{I}_8$	1	1	1	1	1	1	0

## Notación

- ▶  $L$  representa un lenguaje de primer orden sin igualdad.
- ▶  $\mathcal{C}$  es el conjunto de constantes de  $L$ .
- ▶  $\mathcal{F}$  es el conjunto de símbolos de función de  $L$ .
- ▶  $\mathcal{R}$  es el conjunto de símbolos de relación de  $L$ .
- ▶  $\mathcal{F}_n$  es el conjunto de símbolos de función  $n$ -aria de  $L$ .
- ▶  $\mathcal{R}_n$  es el conjunto de símbolos de relación  $n$ -aria de  $L$ .
- ▶  $f/n$  indica que  $f$  es un símbolo de función  $n$ -aria de  $L$ .
- ▶  $P/n$  indica que  $f$  es un símbolo de relación  $n$ -aria de  $L$ .

## Universo de Herbrand

- ▶ Def.: El **universo de Herbrand** de  $L$  es el conjunto de los términos básicos de  $L$ . Se representa por  $\text{UH}(L)$ .

- ▶ Prop.:  $\text{UH}(L) = \bigcup_{i \geq 0} H_i(L)$ , donde  $H_i(L)$  es el **nivel  $i$**  del  $\text{UH}(L)$  definido por

$$H_0(L) = \begin{cases} \mathcal{C}, & \text{si } \mathcal{C} \neq \emptyset; \\ \{a\}, & \text{en caso contrario. (} a \text{ es una nueva constante).} \end{cases}$$

$$H_{i+1}(L) = H_i(L) \cup \{f(t_1, \dots, t_n) : f \in \mathcal{F}_n \text{ y } t_1, \dots, t_n \in H_i(L)\}$$

- ▶ Prop.:  $\text{UH}(L)$  es finito syss  $L$  no tiene símbolos de función.

## Ejemplos de universo de Herbrand

- ▶ Si  $\mathcal{C} = \{a, b, c\}$  y  $\mathcal{F} = \emptyset$ , entonces
  - $H_0(L) = \{a, b, c\}$
  - $H_1(L) = \{a, b, c\}$
  - $\vdots$
  - $\text{UH}(L) = \{a, b, c\}$
- ▶ Si  $\mathcal{C} = \emptyset$  y  $\mathcal{F} = \{f/1\}$ , entonces
  - $H_0(L) = \{a\}$
  - $H_1(L) = \{a, f(a)\}$
  - $H_2(L) = \{a, f(a), f(f(a))\}$
  - $\vdots$
  - $\text{UH}(L) = \{a, f(a), f(f(a)), \dots\} = \{f^i(a) : i \in \mathbb{N}\}$

## Ejemplos de universo de Herbrand

- ▶ Si  $\mathcal{C} = \{a, b\}$  y  $\mathcal{F} = \{f/1, g/1\}$ , entonces
  - $H_0(L) = \{a, b\}$
  - $H_1(L) = \{a, b, f(a), f(b), g(a), g(b)\}$
  - $H_2(L) = \{a, b, f(a), f(b), g(a), g(b), f(f(a)), f(f(b)), f(g(a)), f(g(b)), g(f(a)), g(f(b)), g(g(a)), g(g(b))\}$
  - $\vdots$
- ▶ Si  $\mathcal{C} = \{a, b\}$  y  $\mathcal{F} = \{f/2\}$ , entonces
  - $H_0(L) = \{a, b\}$
  - $H_1(L) = \{a, b, f(a, a), f(a, b), f(b, a), f(b, b)\}$
  - $H_2(L) = \{a, b, f(a, a), f(a, b), f(b, a), f(b, b), f(a, f(a, a)), f(a, f(a, b)), \dots\}$
  - $\vdots$

## Base de Herbrand

- ▶ Def.: La **base de Herbrand** de  $L$  es el conjunto de los átomos básicos de  $L$ . Se representa por  $BH(L)$ .
- ▶ Prop.:  $BH(L) = \{P(t_1, \dots, t_n) : P \in \mathcal{R}_n \text{ y } t_1, \dots, t_n \in UH(L)\}$ .
- ▶ Prop.:  $BH(L)$  es finita syss  $L$  no tiene símbolos de función.
- ▶ Ejemplos:
  - ▶ Si  $\mathcal{C} = \{a, b, c\}$ ,  $\mathcal{F} = \emptyset$  y  $\mathcal{R} = \{P/1\}$ , entonces
 
$$UH(L) = \{a, b, c\}$$

$$BH(L) = \{P(a), P(b), P(c)\}$$
  - ▶ Si  $\mathcal{C} = \{a\}$ ,  $\mathcal{F} = \{f/1\}$  y  $\mathcal{R} = \{P/1, Q/1, R/1\}$ , entonces
 
$$UH(L) = \{a, f(a), f(f(a)), \dots\} = \{f^i(a) : i \in \mathbb{N}\}$$

$$BH(L) = \{P(a), Q(a), R(a), P(f(a)), Q(f(a)), R(f(a)), \dots\}$$

## Interpretaciones de Herbrand

- ▶ Def.: Una **interpretación de Herbrand** es una interpretación  $\mathcal{I} = (U, I)$  tal que
  - ▶  $U$  es el universo de Herbrand de  $L$ ;
  - ▶  $I(c) = c$ , para cada constante  $c$  de  $L$ ;
  - ▶  $I(f) = f$ , para cada símbolo de función  $f$  de  $L$ .
- ▶ Prop.: Sea  $\mathcal{I}$  una interpretación de Herbrand de  $L$ . Si  $t$  es un término básico de  $L$ , entonces  $\mathcal{I}(t) = t$ .
- ▶ Prop.: Una interpretación de Herbrand queda determinada por un subconjunto de la base de Herbrand, el conjunto de átomos básicos verdaderos en esa interpretación.

## Modelos de Herbrand

- ▶ Nota: Las definiciones de universo de Herbrand, base de Herbrand e interpretación de Herbrand definidas para un lenguaje se extienden a fórmulas y conjuntos de fórmulas considerando el lenguaje formado por los símbolos no lógicos que aparecen.
- ▶ Def.: Un **modelo de Herbrand de una fórmula**  $F$  es una interpretación de Herbrand de  $F$  que es modelo de  $F$ .
- ▶ Def.: Un **modelo de Herbrand de un conjunto de fórmulas**  $S$  es una interpretación de Herbrand de  $S$  que es modelo de  $S$ .
- ▶ Ejemplo: Los modelos de Herbrand de  $\{P(a) \vee P(b), \neg P(b) \vee P(c), P(a) \rightarrow P(c)\}$  son  $\{P(b), P(c)\}$ ,  $\{P(a), P(c)\}$  y  $\{P(a), P(b), P(c)\}$  (ver página 5).
- ▶ Ejemplo: Sea  $S = \{\forall x \forall y [Q(b, x) \rightarrow P(a) \vee R(y)], P(b) \rightarrow \neg \exists z \exists u Q(z, u)\}$ .  
Entonces,  

$$UH(S) = \{a, b\}$$

$$BH(S) = \{P(a), P(b), Q(a, a), Q(a, b), Q(b, a), Q(b, b), R(a), R(b)\}$$
 Un modelo de Herbrand de  $S$  es  $\{P(a)\}$ .

## Tema 10: Modelos de Herbrand

### 1. Modelos de Herbrand

### 2. Teorema de Herbrand y decisión de la consistencia

Interpretación de Herbrand de una interpretación

Consistencia mediante modelos de Herbrand

Extensiones de Herbrand

Teorema de Herbrand

Semidecisión mediante el teorema de Herbrand

## Interpretación de Herbrand de una interpretación

Sea  $S = \{\{\neg Q(b, x), P(a), R(y)\}, \{\neg P(b), \neg Q(z, u)\}\}$  e  $\mathcal{I} = (\{1, 2\}, I)$  con  $a^I = 1, b^I = 2, P^I = \{1\}, Q^I = \{(1, 1), (2, 2)\}, R^I = \{2\}$ . Entonces,  $\mathcal{I} \models S$ .

Cálculo de la interpretación de Herbrand  $\mathcal{I}^*$  correspondiente a  $\mathcal{I}$ :

$$\mathcal{I}^* = (\text{UH}(S), I^*)$$

$$\text{UH}(S) = \{a, b\}$$

$$\text{BH}(S) = \{P(a), P(b), Q(a, a), Q(a, b), Q(b, a), Q(b, b), R(a), R(b)\}$$

$$I^*(P(a)) = P^I(a^I) = P^I(1) = \text{V}$$

$$I^*(P(b)) = P^I(b^I) = P^I(2) = \text{F}$$

$$I^*(Q(a, a)) = Q^I(a^I, a^I) = Q^I(1, 1) = \text{V}$$

$$I^*(Q(a, b)) = Q^I(a^I, b^I) = Q^I(1, 2) = \text{F}$$

$$I^*(Q(b, a)) = Q^I(b^I, a^I) = Q^I(2, 1) = \text{F}$$

$$I^*(Q(b, b)) = Q^I(b^I, b^I) = Q^I(2, 2) = \text{V}$$

$$I^*(R(a)) = R^I(a^I) = R^I(1) = \text{F}$$

$$I^*(R(b)) = R^I(b^I) = R^I(2) = \text{V}$$

$$I^* = \{P(a), Q(a, a), Q(b, b), R(b)\} \text{ y } \mathcal{I}^* \models S.$$

## Interpretación de Herbrand de una interpretación

Sea  $S$  el conjunto de cláusulas  $\{\{P(a)\}, \{Q(y, f(a))\}\}$  e  $\mathcal{I} = (\{1, 2\}, I)$  con  $a^I = 1, f^I = \{(1, 2), (2, 1)\}, P^I = \{1\}, Q^I = \{(1, 2), (2, 2)\}$ . Entonces,  $\mathcal{I} \models S$ .

Cálculo de la interpretación de Herbrand  $\mathcal{I}^*$  correspondiente a  $\mathcal{I}$ :

$$\mathcal{I}^* = (\text{UH}(S), I^*)$$

$$\text{UH}(S) = \{a, f(a), f(f(a)), \dots\} = \{f^i(a) : i \in \mathbb{N}\}$$

$$\text{BH}(S) = \{P(f^n(a)) : n \in \mathbb{N}\} \cup \{Q(f^n(a), f^m(a)) : n, m \in \mathbb{N}\}$$

$$I^*(P(a)) = P^I(a^I) = P^I(1) = \text{V}$$

$$I^*(P(f(a))) = P^I(f^I(a^I)) = P^I(f^I(1)) = P^I(2) = \text{F}$$

$$I^*(P(f(f(a)))) = P^I(f^I(f^I(a^I))) = P^I(1) = \text{V}$$

$$I^*(P(f^n(a))) = \begin{cases} \text{V}, & \text{si } n \text{ es par;} \\ \text{F}, & \text{en caso contrario.} \end{cases}$$

$$I^*(Q(f^n(a), f^m(a))) = \begin{cases} \text{V}, & \text{si } m \text{ es impar;} \\ \text{F}, & \text{en caso contrario.} \end{cases}$$

$$I^* = \{P(f^{2n}(a)) : n \in \mathbb{N}\} \cup \{Q(f^n(a), f^{2m+1}(a)) : n, m \in \mathbb{N}\}$$

$$\mathcal{I}^* \models S.$$



## Consistencia mediante modelos de Herbrand

- ▶ Prop.: Sea  $S$  un conjunto de fórmulas básicas. Son equivalentes:
  1.  $S$  es consistente.
  2.  $S$  tiene un modelo de Herbrand.
- ▶ Prop.: Sea  $S$  un conjunto de cláusulas. Si  $\mathcal{I}^*$  es una interpretación de Herbrand correspondiente a un modelo  $\mathcal{I}$  de  $S$ , entonces  $\mathcal{I}^*$  es un modelo de  $S$ .
- ▶ Prop.: Sea  $S$  un conjunto de cláusulas. Son equivalentes:
  1.  $S$  es consistente.
  2.  $S$  tiene un modelo de Herbrand.
- ▶ Prop.: Sea  $S$  un conjunto de cláusulas. Son equivalentes:
  1.  $S$  es inconsistente.
  2.  $S$  no tiene ningún modelo de Herbrand.
- ▶ Prop.: Existen conjuntos de fórmulas consistentes sin modelos de Herbrand.

17 / 25

## Ejemplo de conjunto consistente sin modelos de Herbrand

- ▶ Sea  $S = \{\exists x P(x), \neg P(a)\}$ . Entonces,
  - ▶  $S$  es consistente.  
 $\mathcal{I} \models S$  con  $\mathcal{I} = (\{1, 2\}, I)$ ,  $a^I = 1$  y  $P^I = \{2\}$ .
  - ▶  $S$  no tiene modelos de Herbrand  

$$\text{UH}(S) = \{a\}$$

$$\text{BH}(S) = \{P(a)\}$$
 Las interpretaciones de Herbrand de  $S$  son  $\emptyset$  y  $\{P(a)\}$ .  

$$\emptyset \not\models S$$

$$\{P(a)\} \not\models S$$
  - ▶ Una forma clausal de  $S$  es  $S' = \{P(b), \neg P(a)\}$ .
  - ▶ Un modelo de Herbrand de  $S'$  es  $\mathcal{I} = (U, I)$  con  $U = \{a, b\}$  y  $P^I = \{b\}$ .

18 / 25

## Instancias básicas de una cláusula

- ▶ Def.: Una **sustitución**  $\sigma$  (de  $L$ ) es una aplicación  $\sigma : \text{Var} \rightarrow \text{Térm}(L)$ .
- ▶ Def.: Sea  $C = \{L_1, \dots, L_n\}$  una cláusula de  $L$  y  $\sigma$  una sustitución de  $L$ . Entonces,  $C\sigma = \{L_1\sigma, \dots, L_n\sigma\}$  es una **instancia** de  $C$ .
- ▶ Ejemplo: Sea  $C = \{P(x, a), \neg P(x, f(y))\}$ .  
 $C[x/a, y/f(a)] = \{P(f(a), a), \neg P(f(a), f(f(a)))\}$
- ▶ Def.:  $C\sigma$  es una **instancia básica** de  $C$  si todos los literales de  $C\sigma$  son básicos.
- ▶ Ejemplo: Sea  $C = \{P(x, a), \neg P(x, f(y))\}$ .  
 $\{P(f(a), a), \neg P(f(a), f(f(a)))\}$  es una instancia básica de  $C$ .  
 $\{P(f(a), a), \neg P(f(f(a)), f(a))\}$  no es una instancia básica de  $C$ .  
 $\{P(x, a), \neg P(f(f(a)), f(a))\}$  no es una instancia básica de  $C$ .

## Extensiones de Herbrand

- ▶ Def.: La **extensión de Herbrand** de un conjunto de cláusulas  $S$  es el conjunto de fórmulas  

$$\text{EH}(S) = \{C\sigma : C \in S \text{ y,}$$

$$\text{para toda variable } x \text{ en } C, \sigma(x) \in \text{UH}(S)\}$$
- ▶ Prop.:  $\text{EH}(L) = \bigcup_{i \geq 0} \text{EH}_i(L)$ , donde  $\text{EH}_i(L)$  es el nivel  $i$  de la  $\text{EH}(L)$   

$$\text{EH}_i(S) = \{C\sigma : C \in S \text{ y,}$$

$$\text{para toda variable } x \text{ en } C, \sigma(x) \in \text{UH}_i(S)\}$$
- ▶ Ejemplos:
  - ▶ Sea  $S = \{\{P(x)\}, \{\neg P(f(x))\}\}$ . Entonces,  
 $\text{EH}_0(S) = \{\{P(a)\}, \{\neg P(f(a))\}\}$   
 $\text{EH}_1(S) = \text{EH}_0(S) \cup \{\{P(f(a))\}, \{\neg P(f(f(a)))\}\}$   
 $\text{EH}_2(S) = \text{EH}_1(S) \cup \{\{P(f(f(a)))\}, \{\neg P(f(f(f(a))))\}\}$
  - ▶ Sea  $S = \{\{\neg P(x), Q(x)\}, \{P(a)\}, \{\neg Q(z)\}\}$ . Entonces,  
 $\text{EH}(S) = \{\{\neg P(a), Q(a)\}, \{P(a)\}, \{\neg Q(a)\}\}$ .
  - ▶ Sea  $S = \{\{\neg P(x), Q(x)\}, \{\neg Q(y), R(y)\}, \{P(a)\}, \{\neg R(a)\}\}$ .  
 Entonces,  
 $\text{EH}(S) = \{\{\neg P(a), Q(a)\}, \{\neg Q(a), R(a)\}, \{P(a)\}, \{\neg R(a)\}\}$ .

## Teorema de Herbrand

- ▶ **Teorema de Herbrand:** Sea  $S$  un conjunto de cláusulas. Son equivalentes:
  1.  $S$  es consistente.
  2.  $\text{EH}(S)$  es consistente (en el sentido proposicional).
- ▶ **Prop.:** Sea  $S$  un conjunto de cláusulas. Entonces, son equivalentes
  1.  $S$  es inconsistente.
  2.  $\text{EH}(S)$  tiene un subconjunto finito inconsistente (en el sentido proposicional).
  3. Para algún  $i$ ,  $\text{EH}_i(S)$  es inconsistente (en el sentido proposicional).

## Semidecisión mediante el teorema de Herbrand

**Entrada:** Un conjunto finito de cláusulas,  $S$ .

**Salida:** *Inconsistente*, si  $S$  es inconsistente.

$i := 0$

**bucle**

**si**  $\text{EH}_i(S)$  es inconsistente (en el sentido proposicional) **entonces**

        Devolver *Inconsistente* y terminar

**en caso contrario**

$i := i + 1$

**fsi**

**fbucle**

## Ejemplos de decisión mediante el teorema de Herbrand

- ▶  $S = \{\{\neg P(x), Q(x)\}, \{P(a)\}, \{\neg Q(z)\}\}$  (p. 20) es inconsistente.  
 $\text{EH}_0(S) = \{\{\neg P(a), Q(a)\}, \{P(a)\}, \{\neg Q(a)\}\}$  es inconsistente.
  - 1  $\{\neg P(a), Q(a)\}$
  - 2  $\{P(a)\}$
  - 3  $\{\neg Q(a)\}$
  - 4  $\{Q(a)\}$                       Res 1, 2
  - 5  $\square$                               Res 3, 4
- ▶  $S = \{\{\neg P(x), Q(x)\}, \{\neg Q(y), R(y)\}, \{P(a)\}, \{\neg R(a)\}\}$  es inconsistente.  
 $\text{EH}_0(S) = \{\{\neg P(a), Q(a)\}, \{\neg Q(a), R(a)\}, \{P(a)\}, \{\neg R(a)\}\}$ .
  - 1  $\{\neg P(a), Q(a)\}$
  - 2  $\{\neg Q(a), R(a)\}$
  - 3  $\{P(a)\}$
  - 4  $\{\neg R(a)\}$
  - 5  $\{Q(a)\}$                       Res 1, 3
  - 6  $\{R(a)\}$                       Res 5, 2
  - 7  $\square$                               Res 6, 4

## Ejemplos de decisión mediante el teorema de Herbrand

- ▶  $S = \{\{P(x)\}, \{\neg P(f(x))\}\}$  es inconsistente (p. 20).
  - $\text{EH}_0(S) = \{\{P(a)\}, \{\neg P(f(a))\}\}$  es consistente  
 $\mathcal{I} = \{P(a)\} \models \text{EH}_0(S)$
  - $\text{EH}_1(S) = \text{EH}_0(S) \cup \{\{P(f(a))\}, \{\neg P(f(f(a)))\}\}$  es inconsistente.
    - 1  $\{P(f(a))\}$
    - 2  $\{\neg P(f(a))\}$
    - 3  $\square$                               Res 1, 2
- ▶  $S = \{\{\neg P(x), Q(f(x), x)\}, \{P(g(b))\}, \{\neg Q(y, z)\}\}$  es inconsistente.  
 Dem.:  
 $S' = \{\{\neg P(g(b)), Q(f(g(b)), g(b))\}, \{P(g(b))\}, \{\neg Q(f(g(b)), g(b))\}\}$   
 $\subset \text{EH}(S)$   
 es inconsistente.
  - 1  $\{\neg P(g(b)), Q(f(g(b)), g(b))\}$
  - 2  $\{P(g(b))\}$
  - 3  $\{\neg Q(f(g(b)), g(b))\}$
  - 4  $\{Q(f(g(b)), g(b))\}$                       Res 1, 2
  - 5  $\square$                               Res 3, 3

## Bibliografía

1. M.L. Bonet *Apuntes de LPO*. (Univ. Politécnica de Cataluña, 2003) pp. 31–34.
2. C.L. Chang y R.C.T. Lee *Symbolic logic and mechanical theorem proving* (Academic Press, 1973) pp. 51–62.
3. M. Ojeda e I. Pérez *Lógica para la computación (Vol. 2: Lógica de Primer Orden)* (Ágora, 1997) pp. 59–74.
4. E. Paniagua, J.L. Sánchez y F. Martín *Lógica computacional* (Thomson, 2003) pp. 160–169.
5. L. Paulson *Logic and proof* (U. Cambridge, 2002) pp. 47–50.
6. U. Schöning *Logic for computer scientists* (Birkäuser, 1989) pp. 70–78.