Logique du premier ordre (2)

F. Alexandre

École Supérieure d'Informatique et Applications de Lorraine

March 3, 2009

Pour définir la sémantique d'une formule :

- définition d'un **domaine** (notion d'interprétation)
- donner si nécessaire un sens aux variables libres des formules (notion de valuation)
- donner un sens aux symboles de fonctions et aux symboles de prédicats (notion d'interprétation)
  définition de la sémantique des symboles de fonctions (comme application effective (du domaine vers le domaine)
  et des symboles de relation (comme application du domaine vers les booléens)

Example

Soit la formule

$$\varphi = p(x,\ a) \wedge \exists y \exists z\ p(y,\ z)$$

donner une sémantique à cette formule, c'est dire si elle vaut 0 (faux) ou 1 (vrai),
pour cela on doit préciser
dans quel ensemble varient les variables libres de la formule (ici x)
ce que sont les symboles **p**, **a** de la formule

Definition (valuation)

Soient X un ensemble de variables et E un ensemble, une valuation δ des variables de X est une application de X vers E : $\delta \ : \ X \longrightarrow E$.

Definition

Soient δ une application de X vers E et e un élément de E, $\delta[x := e]$ est la valuation définie par

$\delta[x := e](y) = \delta(y) \textit{ si } y \neq x$

$\delta[x := e](x) = e$

Autrement dit $\delta[x := e]$ coïncide avec δ sauf en x ou elle vaut e

Example

Soient $X = \{x, \ y, \ z\}$, $E = \mathbb{N}$ et δ définie par $\delta(x) = 0$, $\delta(y) = 0$ et $\delta(z) = 1$.
Soit $\zeta = \delta[x := 2]$, on a $\zeta(x) = 2$, $\zeta(y) = 0$ et $\zeta(z) = 1$.

Remarques

- les valuations permettent de donner des valeurs aux variables libres des formules
- une valuation s'appelle aussi affectation de valeurs aux variables ou environnement
- l'ensemble des applications de X vers E est parfois noté E^X

Definition (Interprétation)

Soit $\mathcal{L}$ un langage du premier ordre, une interprétation I pour $\mathcal{L}$, est déterminée par les données suivantes :

- un ensemble E **non vide** appelé le domaine de l'interprétation I
- à chaque symbole de fonction f d'arité n, on associe une application $I(f) \ : \ E^n \longrightarrow E$
- à chaque constante c on associe $I(c) \in E$
- à chaque symbole de relation r d'arité n on associe une relation $I(r)$ sur E^n, c'est-à-dire une application $I(r) \ : \ E^n \longrightarrow \{0, \ 1\}$
- au symbole d'égalité $=$, on fait correspondre l'égalité $=$ sur E, c'est-à-dire $= \ : \ E \times E \longrightarrow \{0, \ 1\}$

Remarques

- Si I est une interprétation on note $|I|$ son domaine ($|I| = E$)
- les variables et les constantes prennent leur valeur dans le domaine de l'interprétation I
- étant donné un symbole de fonction f d'arité n, $I(f)$ est une fonction comportant n d'arguments. De même si r est un symbole de relation $I(r)$ est une relation comportant le même nombre d'arguments que r.

Definition (Interprétation d'un terme)

Soient t un terme défini sur un langage $\mathcal{L}$ et I une interprétation du langage $\mathcal{L}$ de domaine E, l'interprétation $[t](I)$ du terme t dans I est une application de

$$[t](I) \; : \; \begin{array}{ccc} E^X & \rightarrow & E \\ \delta & \mapsto & [t](I)(\delta) \end{array}$$

définie par la façon suivante :

- si t est une variable x alors $[x](I)(\delta) = \delta(x)$
- si t est une constante c alors $[c](I)(\delta) = I(c)$
- si t est de la forme $f(t_1, \ldots, t_n)$ alors $[f(t_1, \ldots, t_n)](I)(\delta) = I(f)([t_1](I)(\delta), \ldots, [t_n](I)(\delta))$

Remarques

L'interprétation I sert à évaluer les symboles de fonctions (et les constantes), alors que la valuation δ sert à évaluer les variables.
La valeur de $[t](I)(\delta)$ ne dépend que de la valeur de δ sur les variables de t (pas de la valeur de δ sur les autres variables).

Example

Soit le terme $t = f(s(x),\ a)$, on considère l'interprétation I telle que $|I| = \mathbb{N}$ et $I(f) = \times$, $I(s) = x \mapsto x + 1$ et $I(a) = 2$ et la valuation δ telle que $\delta(x) = 10$, on a
$[t](I)(\delta) = (10 + 1) \times 2 = 22$

Definition

Soit I une interprétation et E le domaine de I ($|I| = E$), soit δ une valuation des variables et soit ϕ une formule du premier ordre, la valeur de la formule ϕ dans l'interprétation I par rapport à la valuation δ notée $[\phi](I)(\delta)$ est un élément de $\{0,\ 1\}$ défini inductivement de la façon suivante :

- $[r(t_1, \ldots, t_n)](I)(\delta) = I(r)([t_1](I)(\delta), \ldots, [t_n](I)(\delta))$
- $[t_1 = t_2](I)(\delta) = ([t_1](I)(\delta) = [t_2](I)(\delta))$
- $[\neg\phi](I)(\delta) = \overline{[\phi](I)(\delta)}$
- $[\phi_1 \vee \phi_2](I)(\delta) = [\phi_1](I)(\delta) + [\phi_2](I)(\delta)$
- $[\phi_1 \wedge \phi_2](I)(\delta) = [\phi_1](I)(\delta)\ .\ [\phi_2](I)(\delta)$
- $[\phi_1 \Rightarrow \phi_2](I)(\delta) = \overline{[\phi_1](I)(\delta)} + [\phi_2](I)(\delta)$
- $[\forall x \phi](I)(\delta) =$ si pour tout élément e de E, $[\phi](I)(\delta[x := e]) = 1$ alors 1 sinon 0
- $[\exists x \phi](I)(\delta) =$ s'il existe un élément e de E, tel que $[\phi](I)(\delta[x := e]) = 1$ alors 1 sinon 0

Remarques

- l'interprétation d'un atome est analogue à celle d'un terme, mais sa valeur est un élément de $\{0,\ 1\}$
- les connecteurs logiques $\neg$, $\vee$, $\wedge$ et $\Rightarrow$ sont interprétés pas les fonctions booléennes $\bar{\ }$, $+$, $.$ et $\Rightarrow$
- les quantificateurs $\forall$ et $\exists$ sont interprétés selon le sens courant dans le méta-langage par "pour tout" et "il existe"

Example

Soit le langage $\mathcal{L}$ défini par $F = \{a\}$ avec arité(a)=0 et $R = \{r\}$ avec arité(r)=2,
Soit l'interprétation I définie par : $|I| = \mathbb{N}$, $I(r) =<$ et $I(a) = 0$
Soit la valuation définie par $\delta(x) = 2$, $\delta(y) = 1$
On a
$[r(x,\ y)](I)(\delta) = 0$
$[r(y,\ x)](I)(\delta) = 1$
$[\forall z\ r(z,\ y)](I)(\delta) = 0$
$[\exists z\ r(z,\ y)](I)(\delta) = 1$
$[\exists z\ r(z,\ a)](I)(\delta) = 0$
$[\exists z \forall u\ r(z,\ u)](I)(\delta) = 0$
$[\exists z \forall u\ r(u,\ z)](I)(\delta) = 0$
$[\forall z \exists u\ r(z,\ u)](I)(\delta) = 1$
$[\exists z \forall u\ (r(z,\ u)\ \vee\ u = a)](I)(\delta) = 1$
démonstration :
(il existe un élément e_1 de $\mathbb{N}$ tel que $[\forall u\ (r(z,\ u)\ \vee\ u = a)](I)(\delta[z := e_1])$

ssi il existe un élément e_1 de $\mathbb{N}$ tel que
pour tout élément e_2 de $\mathbb{N}$, $[r(z,\ u)\ \vee\ u = a](I)(\delta[z := e_1, u := e_2])$

ssi il existe un élément e_1 de $\mathbb{N}$ tel que
pour tout élément e_2 de $\mathbb{N}$
$[\ r(z,\ u)](I)(\delta[z := e_1, u := e_2]) + [u = a](I)(\delta[z := e_1, u := e_2])$

en choisissant $e_1 = 0$
on obtient : pour tout e_2 de $\mathbb{N}$, $(0 < e_2\ +\ e_2 = 0)$ qui est vrai CQFD)

Proposition

Pour I fixé $[\phi](I)(\delta)$ ne dépend que de la valeur de δ pour les variables libres de ϕ (pas d'autres variables)

Cela signifie que pour I est donnée, si δ_1 et δ_2 coïncident sur les variables libres de ϕ alors $[\phi](I)[\delta_1) = [\phi](I)(\delta_2)$.
On peut démontrer ce résultat par récurrence structurelle sur ϕ.

Si ϕ ne possède pas de variables libres (c'est-à-dire est une formule close) $[\phi](I)(\delta)$ ne dépend pas d'un quelconque δ.

Pour les formules ϕ closes on peut donc noter $[\phi](I)$ la valeur de vérité de ϕ dans l'interprétation I.

Definition (Modèle)

Soient $\mathcal{L}$ un langage du premier ordre, I une interprétation de $\mathcal{L}$, ϕ une formule de $\mathcal{L}$ et $\mathcal{A}$ un ensemble de formules de $\mathcal{L}$,

- I est un modèle de ϕ si et seulement si **pour toute valuation** δ des variables $[\phi](I)(\delta) = 1$
- I est un modèle de $\mathcal{A}$ si et seulement si I est un modèle de chacune des formules de $\mathcal{A}$
- $\mathcal{A}$ est contradictoire si et seulement si $\mathcal{A}$ n'a pas de modèle

Proposition

Soient ϕ une formule, $\mathcal{A}$ un ensemble de formules et I une interprétation.

- I est un modèle de ϕ si et seulement si I est un modèle de la clôture universelle de ϕ
- I est un modèle de $\mathcal{A}$ si et seulement si I est un modèle des clôtures universelles des formules de $\mathcal{A}$

Démonstration

Soient une formule ϕ comportant une seule variable libre x, et I un modèle de ϕ, on a les équivalences suivantes :
I un modèle de ϕ
ssi pour toute valuation $\delta : X \longrightarrow |I|$, $[\phi](I)(\delta) = 1$
ssi pour toute valuation δ, pour tout élément e de $|I|$, $[\phi](I)(\delta[x := e]) = 1$
ssi pour toute valuation δ $[\forall x\phi](I)(\delta) = 1$
ssi $[\forall x\phi](I) = 1$
ssi I est un modèle de ϕ
On peut finir la démonstration en faisant un raisonnement par récurrence sur le nombre de quantificateurs $\forall$ de ϕ

Remarques

- quand dans une formule les variables sont libres il est préférable de les considérer comme des variables universellement quantifiées

- Soient le langage défini par $V = \{x,\ y,\ z\}$, $F = \{f,\ a\}$, $arité(f) = 2$, $arité(a) = 0$ et les formules :
  $\phi_1 \ : \ \forall x \forall y \forall z \ f(x,\ f(y,\ z)) = f(f(x,\ y),\ z)$
  $\phi_2 \ : \ \forall x \ (f(a,\ x) = x \wedge f(x,\ a) = x)$
  $\phi_3 \ : \ \forall x \exists y \ (f(x,\ y) = a \wedge f(y,\ x) = a)$

  Soient les interprétations suivantes :
  I_1 définie par $|I_1| = \mathbb{N}$, $I_1(a) = 0$, $I_1(f) = +$ et
  I_2 définie par $|I_2| = \mathbb{Z}$, $I_2(a) = 0$, $I_2(f) = +$

  Soit I une interprétation,
  si I est un modèle de ϕ_1, ϕ_1 exprime que $I(f)$ est une opération associative,
  si I est un modèle de ϕ_2, ϕ_2 exprime que $I(a)$ l'élément neutre de l'opération $I(f)$,
  si I est un modèle de ϕ_3, ϕ_3 exprime que tout élément de $|I|$ admet un symétrique pour $I(f)$,

  I_1 est un modèle des formules ϕ_1 et ϕ_2,
  I_2 est un modèle des formules ϕ_1 et ϕ_2 et ϕ_3

Soient le langage défini par $V = \{x,\ y,\ z\}$, $R = \{r\}$, $arit\acute{e}(r) = 2$, et les formules :

$\phi_1\ :\ \forall x \forall y\ (r(x,\ y) \Rightarrow \exists z\ r(x,\ z) \wedge r(z,\ y))$

$\phi_2\ :\ \exists x \forall y\ (r(x,\ y) \vee x = y)$

I_1 définie par $|I_1| = \mathbb{N}$ et $I_1(r) =<$,

I_2 définie par $|I_2| = \mathbb{R}$, $I_2(r) =<$

I_1 n'est pas un modèle de ϕ_1

I_1 est un modèle de ϕ_2

I_2 est un modèle de ϕ_1

I_2 n'est pas un modèle de ϕ_2

Definition

Soit ϕ une formule et $\mathcal{A}$ un ensemble de formules closes

- ϕ est un théorème (de la logique du premier ordre) si et seulement si **toute interprétation** est un modèle de ϕ, on le note $\models \phi$
- on dit que deux formules ϕ et ψ sont équivalentes si et seulement si $\phi \Leftrightarrow \psi$ est un théorème de la logique du premier ordre, on a $\models \phi \Leftrightarrow \psi$
- on dit que ϕ se déduit sémantiquement de $\mathcal{A}$ si et seulement si tout modèle de $\mathcal{A}$ est un modèle de ϕ, on le note $\mathcal{A} \models \phi$

Example

Soit $\mathcal{L}$ le langage défini par $V = \{x,\ y,\ z,\ u,\ x',\ y'\}$, $F = \{*,\ e\}$ tel que $arit\acute{e}(e) = 0$ et $arit\acute{e}(*) = 2$

Soit la formule
$\phi :\ \forall x \forall y \forall z\ [(x * y) * z = x * (y * z) \wedge e * x = x \wedge x * e = x \wedge \exists x'\ (x' * x = e \wedge x * x' = e)]$

Soit $\psi\ :\ \forall u\ [\forall x(u * x = x \wedge x * u = x) \Rightarrow u = e]$

Chaque modèle de ϕ est un groupe, ψ exprime l'unicité de l'élément neutre, il est bien connu que $\phi \Rightarrow \psi$ est un théorème.

Exemples de théorèmes

$p(c) \Rightarrow \exists x\ p(x)$

$\exists x \forall y\ r(x,\ y) \Rightarrow \forall y \exists x\ r(x,\ y)$

$\forall x \forall y\ r(x,\ y) \Leftrightarrow \forall y \forall x\ r(x,\ y)$

$\exists x \exists y\ r(x,\ y) \Leftrightarrow \exists y \exists x\ r(x,\ y)$

$\forall x\ (p(x) \wedge q(x)) \Leftrightarrow (\forall x\ p(x)) \wedge (\forall x\ q(x))$

$\exists x\ (p(x) \vee q(x)) \Leftrightarrow (\exists x\ p(x)) \vee (\exists x\ q(x))$

$(\forall x\ p(x)) \vee (\forall x\ q(x)) \Rightarrow \forall x\ (p(x) \vee q(x))$

$\exists x\ (p(x) \wedge q(x)) \Rightarrow (\exists x\ p(x)) \wedge (\exists x\ q(x))$

On mettra en évidence dans la suite des moyens de démontrer ces théorèmes.

Pour montrer qu'une formule ϕ n'est pas un théorème, on met en évidence une interprétation I qui n'est pas un modèle de ϕ

Contre-exemples de théorèmes

- $\forall x \exists y\ r(x,\ y) \Rightarrow \exists y \forall x\ r(x,\ y)$
  Soit l'interprétation I telle que $|I| = \mathbb{N}$ et $I(r) =<$
  on a $[\forall x \exists y\ r(x,\ y)](I) = 1$ et $[\exists y \forall x\ r(x,\ y)](I) = 0$
  donc $[\forall x \exists y\ r(x,\ y) \Rightarrow \exists y \forall x\ r(x,\ y)](I) = 0$
  I n'est pas un modèle de la formule CQFD
- $(\exists x\ p(x)) \wedge (\exists x q(x)) \Rightarrow \exists x\ (p(x) \wedge q(x))$
  cette formule n'est pas un théorème
  soit l'interprétation I suivante : $|I| = \mathbb{N}$, $I(p) = x \mapsto x < 1$, $I(q) = x \mapsto x > 1$
  on a $[\exists x\ p(x)](I) = 1$, $[\exists x q(x)](I) = 1$ donc $[(\exists x\ p(x)) \wedge (\exists x q(x))](I) = 1$
  mais $[\exists x\ (p(x) \wedge q(x))](I) = 0$
  donc I n'est pas un modèle de la formule.

Résultats (suite)

Proposition

Soient ϕ, ψ deux formules closes et $\mathcal{A}$ un ensemble de formules,

$$\mathcal{A} \cup \{\phi\} \models \psi \text{ si et seulement si } \mathcal{A} \models \phi \Rightarrow \psi$$

Definition

Soit le système formel suivant $\mathcal{S} = (For(\mathcal{L}),\ \mathcal{A},\ \mathcal{R})$ tel que

- $For(\mathcal{L})$ est l'ensemble des formules de la logique du premier ordre construites sur le langage $\mathcal{L}$
- $\mathcal{A} = \{\mathcal{V},\ \mathcal{F}\} \cup \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3 \cup \mathcal{A}_4 \cup \mathcal{A}_5$

$$\begin{aligned}
\mathcal{A}_1 &= \{\alpha \Rightarrow (\beta \Rightarrow \alpha)\ ;\ \alpha, \beta \in For(\mathcal{L})\} \\
\mathcal{A}_2 &= \{(\alpha \Rightarrow (\beta \Rightarrow \gamma)) \Rightarrow ((\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \gamma))\ ;\ \alpha, \beta, \gamma \in For(\mathcal{L})\} \\
\mathcal{A}_3 &= \{\alpha \Rightarrow (\neg\alpha \Rightarrow \beta)\ ;\ \alpha, \beta \in For(\mathcal{L})\} \\
\mathcal{A}_4 &= \{(\alpha \Rightarrow \beta) \Rightarrow ((\neg\alpha \Rightarrow \beta) \Rightarrow \beta)\ ;\ \alpha, \beta \in For(\mathcal{L})\} \\
\mathcal{A}_5 &= \{\alpha[x := a] \Rightarrow \exists x\alpha\ ;\ \alpha \in For(\mathcal{L})\ et\ a \in F\}
\end{aligned}$$

- $\mathcal{R} = \{modus\ ponens,\ \exists\text{-}introduction\}$

$$\frac{\alpha,\ \alpha \Rightarrow \beta}{\beta}(m.p.) \qquad \alpha, \beta \in For(\mathcal{L})$$

$$\frac{\alpha \Rightarrow \beta}{\exists x\alpha \Rightarrow \beta}(\exists\text{-}intro) \qquad x \text{ variable non libre dans } \beta$$

Remarques

- Le système précédent est du à Hilbert, il est obtenu par extension à partir du système formel vu dans le cours concernant le calcul des propositions
- Les schémas d'axiomes de $\mathcal{A}_5$ signifie que si α est vérifié pour *a*, alors il existe effectivement un "objet" *x* vérifiant α
- La règle $\exists$-*introduction* signifie que si on prouve β en utilisant une hypothèse α contenant la variable libre *x*, alors l'existence d'une valeur de cette variable satisfaisant α permet de démontrer β

Proposition

Soient le système formel $\mathcal{S}$ défini précédemment, on a les propriétés suivantes :

- pour toute formule ϕ et tout ensemble de formules $\mathcal{A}$,

  $$\mathcal{A} \vdash \alpha \text{ implique } \mathcal{A} \models \alpha$$

  le système formel $\mathcal{S}$ est valide
- pour toute formule ϕ et tout ensemble de formules $\mathcal{A}$,

  $$\mathcal{A} \vdash \alpha \text{ est équivalent à } \mathcal{A} \models \alpha$$

  le système formel $\mathcal{S}$ est complet

- système de déduction naturelle (Gentzen)
- système de calcul des séquents
- système basé sur les clauses et la résolution

Tous ces systèmes possèdent les propriétés de validité et de complétude,
on étudiera plus précisément le système basé sur la résolution et les clauses (comme on l'a vu pour le cas propositionnel)

- initiation à la démonstration automatique
- objectifs :
  -comment montrer qu'une formule α est un théorème de la logique du premier ordre ($\models \alpha$)
  -comment montrer qu'une formule se déduit sémantiquement d'un ensemble de formules $\mathcal{A} \models \alpha$
- Plan
  - interprétations de Herbrand
  - mise sous forme clausale des formules de la logique du premier ordre
  - définition du principe de résolution pour les clauses de la logique du premier ordre
  - utilisation du principe de résolution

Problème

α une formule de la logique du premier ordre

α est un théorème de la logique du premier ordre

ssi

toute interprétation I est un modèle de α

ssi

pour toute interprétation I, pour toute valuation δ, $[\alpha](I)(\delta) = 1$

Remarque : Pour montrer qu'une formule est un théorème, il n'est pas nécessaire de considérer des interprétations arbitraires, on considérera des interprétations particulières, les interprétations de Herbrand.

"Comment créer une interprétation basée sur la syntaxe"

Definition (Univers de Herbrand)

Soit F un ensemble de symboles fonctionnels, l'univers de Herbrand est l'ensemble des termes clos (sans variable) construit sur F, on le note $T(F)$.
Remarque : si F ne comporte pas de symbole de constante, $T(F) = \emptyset$, dans ce cas on ajoute un symbole de constante quelconque, de façon à ce que $T(F) \neq \emptyset$

Example

- Soit $F = \{s,\ a\}$ avec arité(s)=1 et arité(a)=0, l'univers de Herbrand construit sur F est

  $$T(F) = \{a,\ s(a),\ s(s(a)),\ s(s(s(a))),\ \ldots\}$$

- Soit $F = \{f,\ s,\ a\}$ avec arité(f)=2, arité(s)=1 et arité(a)=0. l'univers de Herbrand construit sur F est

  $$T(F) = \{a,\ s(a),\ s(s(a)), \ldots, s(f(a,\ a)), \ldots, f(a,\ a),\ f(a,\ s(a)),\ f(s(a),\ a), \ldots, f(f(a,\ a),\ a), \ldots\}$$

Definition (Base de Herbrand)

Soient F un ensemble de symboles fonctionnels et R un ensemble de symboles de relation
La base de Herbrand construite sur F et R est l'ensemble des atomes de la forme $r(t_1, \ldots, t_n)$ où les termes t_i sont des termes clos.

(La base de Herbrand est donc l'ensemble des atomes clos construits sur F et R).

Example

Soient $F = \{s,\ a\}$ tels que arité(s)=1 et arité(a)=0, et $R = \{r\}$ tel que arité(r)=2,

l'univers de Herbrand est ici l'ensemble des termes clos $\{a,\ s(a),\ s(s(a)),\ \ldots, s(s(s(a))), \ldots,\}$

la base de Herbrand construite sur F et R est
$\{r(a,\ a),\ r(a,\ s(a)),\ r(s(a),\ a),\ r(s(a), s(a)), \ldots\}$, dans ce cas c'est l'ensemble de tous les atomes clos que l'on peut construire avec le symbole de relation r et les termes clos.

Definition

Soit un langage défini par F et R, une interprétation de Herbrand H est définie comme suit :

- $|H| = T(F)$ chaque symbole de fonctions $f \in F$ est interprété de façon canonique, i.e. $H(f) = f$
- les symboles de relations sont interprétés différemment selon les cas, on définit H par un sous-ensemble E de la base de Herbrand du langage (voir définition suivante)

Definition

Soient $\mathcal{L}$ un langage, H une interprétation de Herbrand définie par E (un sous-ensemble de la base de Herbrand).
La valeur des termes, atomes, formules dans H est définie de la façon suivante :

- interprétation des termes clos : $H(f)(t_1, \ldots, t_n) = f(t_1, \ldots, t_n)$
- interprétation des termes quelconques : si σ est une valuation, $[t](H)(\sigma) = \sigma(t)$ (ici une valuation est une application qui à chaque variable fait correspondre un terme clos (c'est à dire une substitution close)
- interprétation des atomes :

  $[r(t_1, \ldots, t_n)](H)(\sigma) = 1$ si et seulement si $r(\sigma(t_1), \ldots, \sigma(t_n)) \in E$
  (E est l'ensemble des atomes vrais dans l'interprétation)
- l'interprétation des formules en découle (c'est la même que dans le cas général)

Théorème de Herbrand

Un ensemble $\mathcal{A}$ de formules sans quantificateur admet un modèle si et seulement si $\mathcal{A}$ admet un modèle de Herbrand.

Démonstration

Si l'ensemble de formules admet un modèle de Herbrand, il admet évidemment un modèle.

Réciproque : supposons que $\mathcal{A}$ admette un modèle I
Pour définir l'interprétation H il suffit définir un sous-ensemble E de la base de Herbrand.
Posons $E = \{r(t_1, \ldots, t_n)\ ;\ t_i\ clos\ et\ [r(t_1, \ldots, t_n)](I)(\delta) = 1\}$
(remarquons que le second membre ne dépend pas de la valuation δ car les termes t_i sont clos).
E définit une interprétation de Herbrand répondant à la question.

Example

Soit le langage $\mathcal{L}$ défini par les ensembles suivants : $F = \{s,\ a\}$, $R = \{p\}$ et les formules $\phi_1\ :\ p(a)$ et $\phi_2\ :\ p(x) \Rightarrow p(s(s(x)))$
Soit l'interprétation I définie par $|I| = \{0,\ 1\}$ et $I(a) = 0$ et $I(s)$ définie par $x \mapsto (x+1)\ mod\ 2$ et $I(p)$ défini par $I(p)(x) = 1$ ssi $x = 0$
I est un modèle de $\{\phi_1,\ \phi_2\}$
L'univers de Herbrand de $\mathcal{L}$ est $\mathcal{H} = \{a,\ s(a),\ s(s(a)),\ s(s(s(a))),\ \ldots\}$
La base de Herbrand de $\mathcal{L}$ est $\mathcal{B} = \{p(a),\ p(s(a)),\ p(s(s(a))),\ p(s(s(s(a))),\ \ldots\}$
Posons $E = \{A\ ;\ A \in \mathcal{B}\ et\ [A](I)(\delta) = 1\}$, on trouve $E = \{p(a),\ p(s(s(a))), \ldots\}$ qui définit un modèle de Herbrand de $\{\phi_1,\ \phi_2\}$