Logique et Représentation des connaissances

LOGIQUE DES PREDICATS

Odile PAPINI,

Université de la Méditerranée.

odile.papini@esil.univmed.fr

http://odile.papini.perso.esil.univmed.fr/index.html

Plan

Partie I : La logique des prédicats

- Introduction
- langage, syntaxe
- système formel
- sémantique

Partie II : Raisonnement en logique des prédicats

- formes prénexes, formes normales, formes de Skolem
- interprétation de Herbrand
- résolution

Introduction : limites de la logique des prédicats

Exemple de raisonnement

Tout homme est mortel,

Socrate est un homme,

donc Socrate est mortel.

en logique propositionnelle

p : "Tout homme est mortel",

q : "Socrate est un homme",

r : "donc Socrate est mortel".

$$p \wedge q \rightarrow r$$

“**Pour tout x**, si **x est un homme** alors **x est mortel**”,

“Socrate est un homme”,

“donc Socrate est mortel”.

x est un homme est représenté par **H(x)**

x est mortel est représenté par **M(x)**

$$\forall x\ (H(x) \to M(x))\ \wedge\ H(Socrate) \to M(Socrate)$$

Le langage de la logique des prédicats : $\mathcal{L}_{Pr}$

Vocabulaire

un ensemble infini dénombrable de symboles de prédicats ou **prédicats**

un ensemble infini dénombrable de symboles **fonctionnels**

un ensemble infini dénombrables de **variables**

les connecteurs : $\neg, \quad , \wedge, \quad \vee, \quad \rightarrow, \quad \leftrightarrow$

les **quantificateurs** $\forall, \quad \exists$

Définitions

terme

- x une variable , f un symbole fonctionnel est un terme
- si $t_1, \cdots, t_n$ sont des termes alors $f(t_1, \cdots, t_n)$ est un terme

atome

- si $t_1, \cdots, t_n$ sont des termes et P est un prédicat alors $P(t_1, \cdots, t_n)$ est un atome

formule

- un atome est une formule
- si A et B sont des formules alors $\neg A, \quad A \wedge B, \quad A \vee B, \quad A \rightarrow B, \quad A \leftrightarrow B$ sont des formules
- si A est une formule et x une variable alors $\forall x \quad A, \quad \exists x \quad A$ sont des formules

Portée des quantificateurs

atome ou formule à laquelle la quantification s'applique

variables liées

variables sous la portée de quantificateurs

si A est une formule,

l'ensemble $\mathbf{Varlie(}A\mathbf{)}$ des variables liées de A est défini par :

- si A est un atome alors $\mathbf{Varlie(}A\mathbf{)} = \emptyset$
- si A est de la forme $B \to C$ alors $\mathbf{Varlie(}A\mathbf{)} = \mathbf{Varlie(}B\mathbf{)} \cup \mathbf{Varlie(}C\mathbf{)}$
- si A est de la forme $\neg\, B$ alors $\mathbf{Varlie(}A\mathbf{)} = \mathbf{Varlie(}B\mathbf{)}$
- si A est de la forme $\forall x\ B$ alors $\mathbf{Varlie(}A\mathbf{)} = \mathbf{Varlie(}B\mathbf{)} \cup \{x\}$

variables libres

variables qui ne sont pas sous la portée de quantificateurs

si A est une formule, $\mathbf{Var}(A)$ est l'ensemble des variables de$A$,

l'ensemble $\mathbf{Varlib}(A)$ des variables libres de A est défini par :

- si A est un atome alors $\mathbf{Varlib}(A) = \mathbf{Var}(A)$
- si A est de la forme $B \rightarrow C$ alors $\mathbf{Varlib}(A) = \mathbf{Varlib}(B) \cup \mathbf{Varlib}(C)$
- si A est de la forme $\neg\, B$ alors $\mathbf{Varlib}(A) = \mathbf{Varlib}(B)$
- si A est de la forme $\forall x\ B$ alors $\mathbf{Varlib}(A) = \mathbf{Varlib}(B)$ - $\{x\}$

une formule sans variable libre est dite **close** ou **fermée**

Substitutions

soit $A(x)$ une formule contenant x comme **variable libre**

soit t un terme

$A(t)$: obtenue en **remplaçant les occurences libres** de x par t dans $A(x)$

Si x **ou** t **apparaissent comme variables liées dans la formule** $A(x)$ **alors**

renommer ces occurences

Système formel de la logique des prédicats

les axiomes

soit A, B, C des formules, x une variable et t un terme, D une formule n'ayant pas x pour variable libre

A1) $(A \rightarrow (B \rightarrow A))$

A2) $((A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)))$

A3) $((\neg A \rightarrow \neg B) \rightarrow (B \rightarrow A))$

A4) $(\forall x\ A(x) \rightarrow A(t))$

A5) $((D \rightarrow B) \rightarrow (D \rightarrow \forall x\ B)$

règles de déduction

modus ponens

$$\frac{\vdash A, \vdash A \to B}{\vdash B}$$

généralisation

$$\frac{\vdash A}{\vdash \forall x\, A}$$

Déduction

Soit B une formule de $\mathcal{L}_{Pr}$ et $H_1, \cdots, H_n$ **des hypothèses**

Une **déduction** de B à partir des hypothèses $H_1, \cdots, H_n$

$$H_1, \cdots, H_n \vdash B$$

est une suite de formules $F_1, \cdots, F_i, F_n$ telle que :

$F_n = B$

et F_i, $1 \leq i < n$ est :

- soit une des hypothèses $H_1, \cdots, H_n$
- soit un axiome
- soit obtenue par l'application de règles de déduction à partir de formules F_j, $j < i$

proposition :

$$\forall A \in \mathcal{L}_{Pr} \quad \vdash (A \longrightarrow A)$$

proposition :

$\forall A_1, \cdots, A_{n-1} \in \mathcal{L}_{Pr}$

$$\text{si} \quad A_1, \cdots, A_{n-1} \vdash (A_n \longrightarrow B) \quad \text{alors} \quad A_1, \cdots, A_n \vdash B$$

proposition (théorème de déduction):

Soient $A_1, \cdots, A_n$ des formules closes de $\mathcal{L}_{Pr}$

$$\text{si} \quad A_1, \cdots, A_n \vdash B \quad \text{alors} \quad A_1, \cdots, A_{n-1} \vdash (A_n \longrightarrow B)$$

sémantique de la logique des prédicats

interprétation : $I = (D, I_c, I_v)$ où

• D ensemble non vide, domaine d'interprétation

• I_c la fonction :

$$\begin{array}{ccc} D^n & \rightarrow & D \\ f & \rightarrow & I_c(f) \end{array} \qquad \qquad \begin{array}{ccc} D^m & \rightarrow & \{0,1\} \\ P & \rightarrow & I_c(P) \end{array}$$

• I_v la fonction :

$$\begin{array}{ccc} Var & \rightarrow & D \\ x & \rightarrow & I_v(x) \end{array}$$

interprétation d'une formule de la logique des prédicats

A une formule de $\mathcal{L}_{Pr}$, association d'une valeur de vérité $I(A)$ à A

- si x est une variable libre alors $I(x) = I_v(x)$
- $I(f(t_1, \cdots, t_n)) = (I_c(f))(I(t_1), \cdots, I(t_n))$
- $I(P(t_1, \cdots, t_m)) = (I_c(P))(I(t_1), \cdots, I(t_m))$
- si A et B sont des formules alors $\neg A$, $A \wedge B$, $A \vee B$, $A \rightarrow B$, $A \leftrightarrow B$ s'interprètent comme dans la logique propositionnel
- si A est une formule et x une variable alors $I(\forall x\ A) = 1$ si $I_{x/d}(A) = 1$ pour tout élément $d \in D$
- si A est une formule et x une variable alors $I(\exists x\ A) = 1$ si $I_{x/d}(A) = 1$ pour au moins un élément $d \in D$

quelques définitions

Soient $A \in \mathcal{L}_{Pr}$, $B \in \mathcal{L}_{Pr}$ et $\mathcal{F} \subset \mathcal{L}_{Pr}$

A est une **tautologie** , $\models A$, si pour toute interprétation I, $I(A) = 1$

B est une **conséquence** de A si pour toute interprétation I, $I(A) = 1$ alors $I(B) = 1$, on écrit $A \models B$

B est une **conséquence** de $\mathcal{F}$ si pour toute interprétation I, tq $\forall A \in \mathcal{F}$, $I(A) = 1$ alors $I(B) = 1$, on écrit $\mathcal{F} \models B$

A est **satisfaisable** s'il existe une interprétation I tq $I(A) = 1$

$\mathcal{F}$ est **satisfaisable** s'il existe une interprétation I tq $\forall A \in \mathcal{F}$, $I(A) = 1$

A est **insatisfaisable** ou **incohérente** si pour toute interprétation I, $I(A) = 0$

$\mathcal{F}$ est **insatisfaisable** si pour toute interprétation I, $\exists A \in \mathcal{F}$ tq $I(A) = 0$

proposition : $\mathcal{F} \subset \mathcal{L}_{Pr}$ ensemble de formules closes, B formule close
$\mathcal{F} \models B$ ssi $\mathcal{F} \cup \{\neg B\}$ est insatisfaisable

quelques propriétés

$$(\forall x\ A \wedge \forall x\ B) \equiv \forall x\ (A \wedge B)$$

$$(\forall x\ A \vee \forall x\ B) \models \forall x\ (A \vee B)$$

$$\forall x\ (A \to B) \models (\forall x\ A \to \forall x\ B)$$

$$\forall x\ (A \leftrightarrow B) \models (\forall x\ A \leftrightarrow \forall x\ B)$$

$$\exists x\ (A \vee B) \equiv (\exists x\ A \vee \exists x\ B)$$

$$\exists x\ (A \wedge B) \models (\exists x\ A \wedge \exists x\ B)$$

$$\exists x\ (A \to B) \equiv (\exists x\ A \to \exists x\ B)$$

$$\forall x\ \neg A \equiv \neg \exists x\ A$$

$$\forall x\ \forall y\ A \equiv \forall y\ \forall x\ A$$

$$\exists x\ \exists y\ A \equiv \exists y\ \exists x\ A$$

$$\exists x\ \forall y\ A \models \forall y\ \exists x\ A$$

quelques théorèmes

théorème (d'adéquation) : $\forall A \in \mathcal{L}_{Pr}$ si $\vdash A$ alors $\models A$
(les formules qui sont des théorèmes sont des tautologies)

théorème (de complétude) : $\forall A \in \mathcal{L}_{Pr}$ si $\models A$ alors $\vdash A$

théorème (de compacité) : Soit $\mathcal{F}$ un ensemble de formules de $\mathcal{L}_{Pr}$.
Si toute famille finie $\mathcal{F}' \subset \mathcal{F}$ est satisfaisable alors $\mathcal{F}$ est aussi satisfaisable.

théorème (de finitude) : Soit $\mathcal{F}$ un ensemble de formules de $\mathcal{L}_{Pr}$. Soit $B \in \mathcal{L}_{Pr}$
si $\mathcal{F} \models B$ alors $\exists \mathcal{F}' \subset \mathcal{F}$ fini tq $\mathcal{F}' \models B$

théorème (de complétude généralisé) : Soit $\mathcal{F} \subset \mathcal{L}_{Pr}$ et $B \in \mathcal{L}_{Pr}$,
$\mathcal{F} \models B$ ssi $\mathcal{F} \vdash B$

proposition : Soit $\mathcal{F} \subset \mathcal{L}_{Pr}$ et B une tautologie, $\mathcal{F} \vdash \neg B$ si $\mathcal{F}$ n'a pas de modèle.

quelques résultats de décidabilité et d'indécidabilité

La logique des prédicats est indécidable

Il n'existe aucun programme qui pour une formule $A \in \mathcal{L}_{Pr}$ indique en un temps fini si A est une tautologie

Toute théorie axiomatique égalitaire ayant :

- un nombre fini de symboles, un nombre fini de constantes
- un seul symbole fonctionnel unaire f
- un nombre fini de prédicats unaires et le prédicat binaire égalité
- n'ayant pas d'axiomes non logiques

est décidable

formes prénexes, formes normales

$$Q_1\, x_1\ \cdots\ Q_n\, x_n\, M$$

proposition : pour toute formule A il existe une forme prénexe équivalente à A

algorithme

- élimination des connecteurs d'implication et d'équivalence
- renommage des variables (plus de variable libre et liée en même temps)
- suppression des quantificateurs inutiles
- transfert du connecteur de négation immédiatement devant les atomes
- transfert des quantificateurs en tête des formules

extension du vocabulaire à la logique des prédicats

littéral : un atome ou la négation d'un atome

clause : disjonction de littéraux

cube : conjonction de littéraux

forme conjonctive normale : forme prénexe dont la matrice M est une conjonction de clauses

forme disjonctive normale : forme prénexe dont la matrice M est une disjonction de cubes

formes de Skolem

proposition : S_A forme de skolem de A, A est satisfaisable ssi S_A est satisfaisable

transformation de A en forme de Skolem S_A

- transformer A en forme prénexe : $Q_1\, x_1\ \cdots\ Q_n\, x_n\, M$
- transformer M en forme conjonctive normale M'
- skolémiser M' :

  1) associer à toute variable quantifiée existentiellement le terme constitué par un symbole fonctionnel ayant pour arguments la liste des variables quantifiées universellement qui précèdent la variable

  2) remplacer chaque occurence de variable quantifiée existentiellement par le terme défini en 1)

  3) supprimer les quantificateurs existentiels

théorème de Herbrand

on associe à une formule conjonctive normale F l'ensemble C des clauses correspondantes

univers de Herbrand associé à un ensemble de clauses C :

ensemble de tous les termes sans variable construit à partir du vocabulaire de C

système de Herbrand SH_C associé à C :

ensemble des clauses obtenues à partir de C en remplaçant les variables par des éléments de l'univers de Herbrand

théorème de Herbrand :

C est satisfaisable ssi SH_C est satisfaisable

Raisonnement déductif

définition (rappel)

$\mathcal{F}$ un ensemble de formules de $\mathcal{L}_{Pr}$ et I une interprétation,

$$\mathcal{F} \text{ est insatisfaisable ou incohérent ssi } \forall I \text{ telle que } \exists F \in \mathcal{F},\ I(F) = 0.$$

proposition

$F \in \mathcal{L}_{Pr}$, $\quad \mathcal{F} \models F$ ssi $\mathcal{F} \cup \{\neg F\}$ est insatisfaisable ou incohérent.

on ramène le problème de conséquence logique à celui de la cohérence ou de la satisfaisabilité.

réfutation : on cherche à montrer l'incohérence

Résolution en calcul des prédicats

résolution R et **saturation S** (instanciation dans l'univers de Herbrand H_C) semi-commutatives :

$$R(S(H_C)) \subseteq S(R(H_C))$$

plus généralement :

$$R^n(S(H_C)) \subseteq S(R^n(H_C))$$

l'unification

substitution : Var : ensemble des variables, T : ensemble des termes

σ fonction de Var dans T, tq l'ensemble $\{x \in Var, \sigma(x) \neq x\}$ est fini

instance : t un terme, l : un littéral,

- instance de t (resp. de l), le terme noté $\sigma(t)$ (resp. le littéral $\sigma(l)$), obtenu en remplaçant toutes les occurences des variables x par $\sigma(x)$.
- t_1 et t_2 des termes, t_2 est une instance de t_1 s'il existe une substitution σ telle que $t_2 = \sigma(t_1)$

unificateur :

2 littéraux l et l' sont **unifiables** s'il existe une substitution σ telle que $\sigma(l) = \sigma(l')$. La substitution est appelée l'unificateur.

unificateur principal: σ' est plus général que σ, s'il existe σ" tq $\sigma = \sigma" \circ \sigma'$

il existe un unificateur plus général que tous les autres : **unificateur principal**

```
algorithme unification (t1, t2 : des termes)
début
si l'un des termes (t1 ou t2) est une variable x, (l'autre est t) alors
  si x = t alors
     possible ← vrai
     σ ← ∅
  sinon
     si x apparaît dans t alors
       possible ← faux
     sinon
       possible ← vrai
       σ ← (σ(x) = t)
     finsi
  fin si
sinon (t1 = f(x1, ···, xn) et t2 = g(y1, ···, ym))
  si f ≠ g ou n ≠ m alors
     possible ← faux
  sinon
     i ← 0
     possible ← vrai
     σ ← ∅
     tant que i < n et possible faire
       i ← i + 1
       (possible, σ') ← unification(σ(xi), σ(yi))
       si possible alors
         σ ← σ' ∘ σ finsi
     fin tant que
  fin si
fin si
fin
```

proposition (résolution)

S : un ensemble de clauses, c_1, $c_2 \in S$,

l_1 apparaît dans c_1 et $\neg l_2$ apparaît dans c_2

θ une substitution de renommage tq $\theta(c_1)$ et c_2 n'ont aucune variable libre en commun

soit σ_p l'unificateur principal de $\theta(c_1)$ et c_2

$S \models S \cup \{r\}$ et $S \cup \{r\} \models S$

avec $r = \sigma(\theta(c_1 \backslash \{l_1\}) \vee (c_2 \backslash \{\neg l_2\}))$ appelée résolvante

Algorithme de résolution

<u>`début`</u>

<u>`tant que`</u> $\Box \notin S$ <u>`faire`</u>

`choisir` l_1, l_2, c_1, c_2 `tels que` $l \in c_1$ `et` $l_2 \in c_2$ `et` l_1, l_2 `unifiables`

`calculer la résolvante` r
`à partir de l'unificateur principal`

`remplacer` S `par` $S \cup \{r\}$

<u>`fin tant que`</u>

<u>`fin`</u>