Cours 8 Calcul des prédicats : Syntaxe

8.1 Introduction

Le calcul propositionnel reste très limité, et ne permet essentiellement que d'exprimer des opérations booléennes sur des propositions. Si l'on veut pouvoir raisonner sur des assertions mathématiques, il nous faut autoriser des constructions plus riches. Par exemple, on peut vouloir écrire l'énoncé

$$\forall x((\texttt{Premier}(x) \wedge x > 1 + 1) \Rightarrow \texttt{Impair}(x))$$

Le calcul des prédicats(appelé aussi logique du premier ordre) est une extension du calcul propositionnel. Il permet d'introduire des variables appartenant à un domaine arbitraire grâce aux quantificateurs $\forall$ et $\exists$.
Le calcul des prédicats permet de parler de structures comportant *un seul domaine non vide* (contrairement à la logique propositionnelle, ce domaine peut avoir plus de deux valeurs), des fonctions et relations(prédicats) sur ce domaine.
La terminologie *premier ordre* fait référence au fait que les quantifications existentielles et universelles ne sont autorisées que sur les variables. Un énoncé du second ordre, on parle plus généralement d'ordre supérieur, serait un énoncé où l'on autoriserait les quantifications sur les fonctions ou des relations : par exemple, on peut écrire $\neg\exists f(\forall x(f(x) > f(x+1)))$ pour signifier qu'il n'existe pas de suite infiniment décroissante.

Quantificateurs

La formule $\exists x P(x)$ se lit : il existe $x \in D$ tel que $P(x)$ est vraie.

La formule $\forall x P(x)$ se lit : Pour tout $x \in D$, $P(x)$ est vraie.

D est le domaine de définition des prédicats.

Le calcul des prédicats, reste le formalisme le plus courant pour exprimer des propriétés mathématiques.

calcul des prédicats et bases de données

C'est aussi un formalisme très utilisé en informatique pour décrire les objets : par exemple, les langages de requêtes à des bases de données sont essentiellement basés sur ce formalisme, appliqué à des objets finis, qui représentent des données.

L'objectif est alors de définir la logique du premier ordre. Comme pour la logique propositionnelle, on va le faire en parlant d'abord de la *syntaxe*, c'est-à-dire comment on écrit les formules, puis de leur *sémantique*.

8.2 L'alphabet

On distingue dans l'alphabet certains symboles qui sont communs à tous les langages, et certains symboles qui varient d'un langage à l'autre. Les symboles communs à tous les langages sont : les connecteurs, les parenthèses (et) et la virgule , le quantificateur universel $\forall$ et le quantificateur existentiel $\exists$; Un ensemble infini dénombrable de symboles $\mathcal{V}$ de variables.
Les symboles qui peuvent varier d'un langage à l'autre sont capturés par la notion de signature. Une signature fixe les symboles de constantes, les symboles de fonctions et les symboles de relations(prédicats) qui sont autorisés. Formellement :

Signature d'un langage du premier ordre

La signature $\Sigma = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ d'un langage du premier ordre est la donnée :
- d'un premier ensemble $\mathcal{C}$ de symboles, appelés symboles de constantes ;
- d'un second ensemble $\mathcal{F}$ de symboles, appelés symboles de fonctions. A chaque symbole de cet ensemble est associé un entier strictement positif, que l'on appelle son arité.
- d'un troisième ensemble $\mathcal{R}$ de symboles, appelés symboles de relations. A chaque symbole de cet ensemble est associé un entier strictement positif, que l'on appelle son arité.

On suppose que $\mathcal{V}, \mathcal{C}, \mathcal{F}, \mathcal{R}$ sont des ensembles disjoints deux à deux. Une formule du premier ordre sera alors un mot sur l'alphabet

$$\mathcal{A}(\Sigma) = \mathcal{V} \cup \mathcal{C} \cup \mathcal{F} \cup \mathcal{R} \cup \{\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, "(", ")", ",", \forall, \exists\}.$$

Exemple 1. *Par exemple, on peut considérer la signature*

$$\Sigma = (\{0,1\}, \{s, +\}, \{\texttt{Impair},\ \texttt{Premier}, =, <\})$$

qui possède les symboles de constante 0 *et* 1*, le symbole de fonction* $+$ *d'arité* 2*, le symbole de fonction* s *d'arité* 1*, les symboles de relations* `Impair` *et* `Premier` *d'arité* 1*, les symboles de relations* $=$ *et* $<$ *d'arité* 2*.*

Exemple 2. *On peut aussi considérer la signature* $\mathcal{L}_2 = (\{c, d\}, \{f, g, h\}, \{R, =\})$ *avec* c, d *deux symboles de constante,* f *un symbole de fonction d'arité* 1*,* g *et* h *deux symboles de fonctions d'arité* 2*,* R *un symbole de relation d'arité* 2*.*

On le note aussi la signature avec le symbole en indiquant si c'est une fonction ou une relation avec son arité. $\mathcal{L}_2$ est noté alors comme : $c^{f0}, d^{f0}, f^{f1}, g^{f2}, h^{f2}, R^{r2}, =^{r2}$.

> Remarquons que les constantes sont considérés comme des fonction d'arité 0. Les variables propositionnels sont des relations(prédicats) d'arité 0.

Exemple 3. *Une signature possible pour la théorie des ensembles est* $\in, =$ *. Notons que toutes les autres opérations sur les ensembles peuvent être définies à partir de ces deux symboles*

Un symbole est surchargé dans une signature, lorsque cette signature comporte deux déclarations distinctes du même symbole. Il est fréquent (en arithmétique) d'utiliser des signatures dans lesquelles le signe moins est surchargé. Il est utilisé simultanément pour obtenir l'opposé d'un nombre ou faire la soustraction de deux nombres

En absence de signature, on prend généralement ces conventions.

Alphabet conventionnelle.

$x, y, z, x_1, x_2, y_1, y_2, \ldots$	les variables	$a, b, c, a_1, a_2, b_1, b_1, \ldots$	les constantes
$P, Q, R, P_1, P_2, Q_1, Q_2, \ldots$	les prédicats	$f, g, h, f_1, f_2, g_1, g_2, \ldots$	les fonctions
$\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$	les connecteurs	$\exists, \forall$	les quantificateurs
$($, $)$	les parenthèses		

On va définir par étapes : d'abord les termes, qui visent à représenter des objets, puis les formules atomiques, qui visent à représenter des relations entre objets, et enfin les formules.

8.3 Les termes

Les termes

Les termes du calcul des prédicats sont définis comme suit

- les variables et les constantes sont des termes.
- Si f est une fonction à n arguments et $t_1, t_2, \ldots, t_n$ sont des termes alors $f(t_1, t_2, \ldots, t_n)$ est un terme.
- L'ensemble des termes est engendré par les deux clauses précédentes.

Exemple 4. $x_1, a_1, f_1(x_1), f_2(x, g(x))$ *et* $h(f(x, y), g(z, x), a)$ *sont des termes.*

8.4 Formule atomique

Formule atomique

Une formule atomique est définie par :

Si P_i est un prédicat (i.e Qu'on interprètera ensuite comme une fonction définie de $D^n \longrightarrow \{0,1\}$) et $t_1, t_2, \ldots, t_n$ sont des termes alors $P_i(t_1, t_2, \ldots, t_n)$ est une formule atomique.

Exemple 5. $P(x, y), P(x, g(x, y), a), Q(a, b), Q(g(f(x, y)), z)$ *sont des formules atomiques.*

8.5 Formules

Définition inductive des formules

1. Chaque formule atomique est une formule (bien formée).
2. Si A et B sont deux formules alors

$$\neg A, \quad (A \wedge B), \quad (A \vee B), \quad (A \Rightarrow B), \quad (A \Leftrightarrow B), \quad \forall x A, \quad \exists x A$$

   sont des formules.
3. L'ensemble de toutes les formules est engendrée par les clauses (1) et (2).

Remarque. On peut utiliser des symboles spéciaux pour désigner des prédicats $=, \neq, \leq, \geq, >, <$ ou des fonctions particulières $\{+, -, *, /\}$.
En abréviation on peut écrire, $+(x, *(y, z))$ en $x + y * z$ et $\leq (*(3, 5), +(y, 5))$ en $3 + x \leq y + 5$.

8.6 Priorités :

Définition 1 (Formule à priorité)**.** *Une formule à priorité est :*

1. *Une formule atomique.*
2. *Si A est une formule à priorité alors $\neg A$ est une formule à priorité.*
3. *Si A et B sont des formules à priorité alors $A\alpha B$ est une formule à priorité.*
4. *Si A est une formule à priorité et si x une variable alors $\forall x A$ et $\exists x A$ sont des formules à priorité.*
5. *Si A est une formule à priorité alors (A) est une formule à priorité.*

On règle les problèmes de priorités en attribuant aux symboles du tableau suivant des priorités décroissantes du haut vers le bas du tableau :

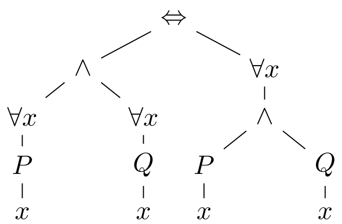
Opérations(fonctions)	
$-+$ unaire	
$*, /$	associatif gauche
$+, -$ binaire	associatif gauche
Relations(Prédicats)	
$=, \neq, <, \leq, >, \geq$	
Négation, Quantificateur	
$\neg, \forall, \exists$	
Connectives binaires	
$\wedge$	associatif gauche
$\vee$	associatif gauche
$\Rightarrow$	associatif droit
$\Leftrightarrow$	associatif gauche

Exemple 1. La formule (à priorité)

$$\forall x P(x) \wedge \forall x Q(x) \Leftrightarrow \forall x (P(x) \wedge Q(x))$$

peut être vue comme une abréviation de la formule

$$((\forall x P(x) \wedge \forall x Q(x)) \Leftrightarrow \forall x (P(x) \wedge Q(x)))$$



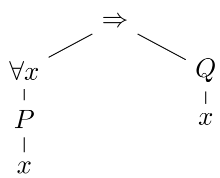
Exemple 2. La formule (à priorité)

$$\forall x \forall y \forall z \; x \leq y \wedge y \leq z \Rightarrow x \leq z$$

peut être vue comme une abréviation de la formule

$$\forall x \forall y \forall z ((\leq (x, y) \wedge \leq (y, z)) \Rightarrow \leq (x, z))$$

Exemple 3. Notez que, puisque la priorité du $\forall$ est plus forte que celle de $\Rightarrow$, dans la formule $\forall x P(x) \Rightarrow Q(x)$, l'opérande gauche de l'implication est $\forall x P(x)$. La structure de la formule sera ainsi représentée par la formule à priorité $(\forall x P(x)) \Rightarrow Q(x)$ ou par l'arbre suivant :



Exemples
• L'énocé $\forall x((Premier(x) \wedge x > 1 + 1) \Rightarrow Impair(x))$ est une formule sur la signature Σ de l'exemple 8.1.
• $\exists x(s(x) = 1 + 0 \vee \forall y \; x + y > s(x))$ aussi.
Exemples de formules sur la signature $\mathcal{L}_2$:
• $\forall x \forall y \forall z((R(x, y) \wedge R(y, z) \Rightarrow R(x, z))$
• $\forall x \exists y(g(x, y) = c \wedge g(y, x) = c)$; • $\forall x \neg f(x) = c$; • $\forall x \exists y \neg f(x) = c$.
Par contre l'expression $\exists x R(f(x))$ n'est pas une formule sur la signature $\mathcal{L}_2$!

8.7 Théorème de Lecture unique :

Comme pour le calcul propositionnel, on peut toujours décomposer une formule, et ce de façon **unique.**

Proposition 1 (Décomposition / Lecture unique)**.**
Soit F une formule. Alors F est d'une, et exactement d'une, des formes suivantes :
$\boxed{1.}$ *une formule atomique ;*
$\boxed{2.}$ $\neg G$*, où G est une formule ;*
$\boxed{3.}$ $(G\alpha H)$ *où G et H sont des formules et α un connecteur binaire ;*
$\boxed{4.}$ $\forall xG$ *où G est une formule et x une variable ;*
$\boxed{5.}$ $\exists xG$ *où G est une formule et x une variable.*
De plus dans le premier cas, il y a une unique façon de "lire" la formule atomique. Dans chacun des autres cas, il y a unicité de la formule G et de la formule H et du connecteur.

8.8 Variables libres et liées

Le sens de la formule $x + 2 = 4$ dépend de x : la formule n'est vraie (en arithmétique) que si $x = 2$. La variable x est libre dans la formule.

Par contre toujours en arithmétique $\forall x(x + 2 = 4)$ est une formule fausse et la formule $\forall x\forall y(x + y = y + x)$ est une formule vraie indépendamment des valeurs de x et y. Ces deux formules n'ont pas de variables libres. Précisons ces notions.

Occurrences libres et liées

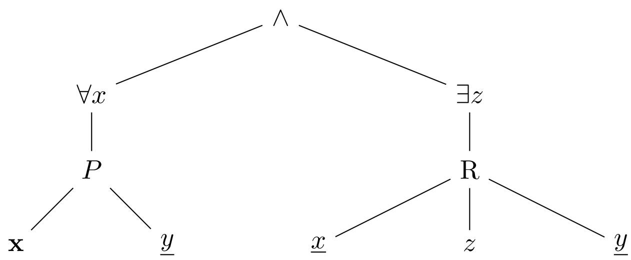
Soit x une variable et A une formule.
Dans une formule $\forall xA$ ou $\exists xA$, la portée de la liaison pour x est A.
Une occurrence de x dans une formule est libre si elle n'est pas dans la portée d'une liaison pour x.

Pour voir les occurrences des variables, on dessine les structures des formules en faisant apparaître $\forall x$ et $\exists x$ comme des sommets de ces structures.

Dessinons la structure de la formule

$$\forall xP(x, y) \wedge \exists zR(x, z, y),$$

puisque la priorité de $\forall$ est supérieure à celle du $\wedge$, cette structure est :



Sur ce type de structure, une occurrence liée de la variable x c'est une occurrence en dessous d'un sommet $\forall x$ ou $\exists x$. Une occurrence de x qui n'est pas sous un tel sommet est libre. Revenons à l'exemple : l'occurrence en gras de x est liée, l'occurrence soulignée de x est libre. L'occurrence de z est liée.

Variables libres et liées

$\boxed{1.}$ x est une variable libre d'une formule *ssi* il y a une occurrence libre de x dans la formule.
$\boxed{2.}$ Une formule sans variable libre est aussi appelée une formule fermée.

Revenons à l'exemple, la formule a comme variables libres x et y. Tandis que z est une variable liée.

Variable liée : variable muette En d'autres termes, x est une variable muette dans le sens où la valeur de vérité de $\forall xF$ ou $\exists xF$ aura vocation, lorsqu'on parlera de la sémantique des formules, à ne pas dépendre de x

Remarque 1. *On a le même phénomène en mathématique avec le symbole intégrale : dans l'expression $\int_a^b f(t)dt$, t est une variable muette (liée). En particulier $\int_a^b f(u)du$ est exactement la même intégrale.*

Exercices.
$\boxed{1.}$ Exprimer la définition de continuité de f en une formule du logique du premier ordre.
$\boxed{2.}$ En utilisant le prédicat égalité, exprimer la formule exprimant il existe un unique élément vérifiant P.

---

1. **Olivier Bournez** : *Fondements de l'informatique Logique, modèles, et calculs. Cours INF423 de l'Ecole Polytechnique* Chap 5.
2. **Michel Levy** : *Logique et démonstration automatique - Introduction à la logique propositionnelle et à la logique du premier ordre.*