La logique des prédicats

0. Résumé des épisodes précédents

$$A = X \mid \top \mid \bot \mid \neg A \mid A \wedge A \mid A \vee A \mid A \Rightarrow A$$

Qu'est-ce que X ?

Il pleut $\wedge$ il fait jour

Mais pas

$$1 \in \mathbb{N} \wedge 1 \leq 3$$

Décomposer X en $P(t_1, ..., t_n)$

Permet d'introduire des quantificateurs

$$\forall x \ (x + y = y + x)$$

La logique propositionnelle $\longrightarrow$ la logique des prédicats

Aujourd'hui

Définir l'ensemble des propositions

Définir le (sous)-ensemble des propositions démontrables

Définir le (sous)-ensemble des propositions valides dans un modèle

I. Une digression : comment définir un ensemble ou une relation ?

Par une définition explicite :

$$\{x \in \mathbb{N} \mid \exists z \in \mathbb{N}\ x = 2 \times z\}$$

$$\{(x, y) \in \mathbb{N}^2 \mid \exists z \in \mathbb{N}\ x = y \times z\}$$

Mais ça ne suffit pas ...

Par une définition inductive

La notion de définition inductive : le théorème du point fixe

Le premier théorème du point fixe

E, $\leq$ relation d'ordre

u_0, u_1, ... suite croissante
l limite de $(u_i)_i$ si $l = sup\ \{u_0, u_1, ...\}$

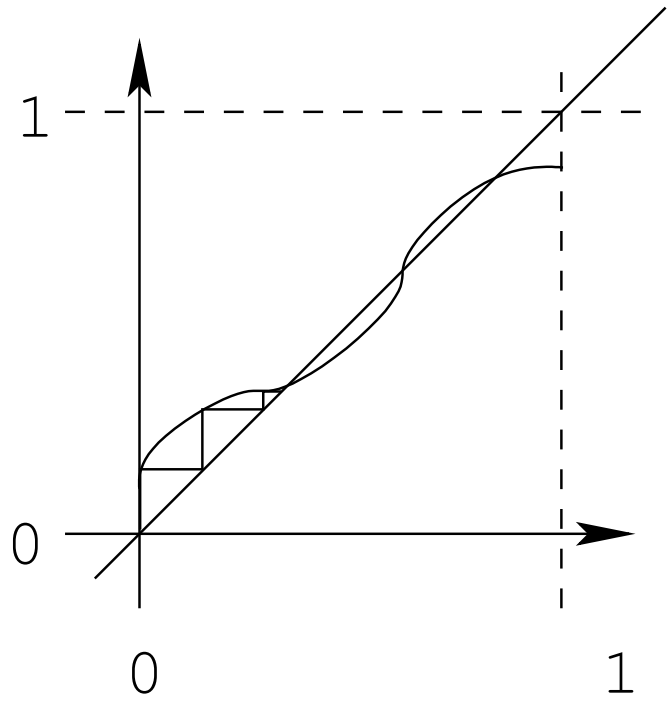
E, $\leq$ faiblement complète si toute suite croissante a une limite
f croissante est continue si $lim_i\ (f\ u_i) = f\ (lim_i\ u_i)$

Théorème : $\leq$ faiblement complète et a un minimum et f continue alors f a un point fixe

Le plus petit point fixe est $lim_i\ (f^i\ m)$

Exemple

$[0,1]$, $\leq$ est faiblement complète



$\mathbb{R}^+$, $\leq$ est-elle faiblement complète ?

Le deuxième théorème du point fixe

Pour les fonctions croissantes (mais pas forcément continues)

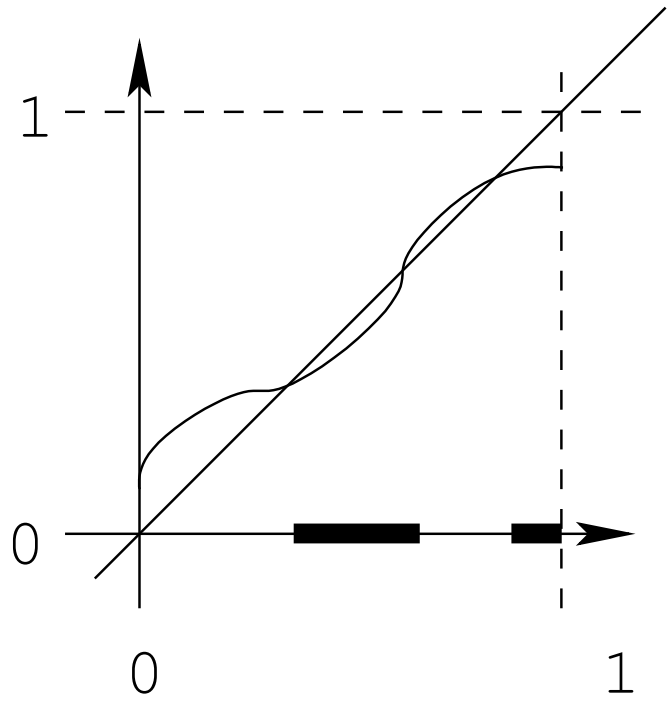
$E, \leq$ fortement complète si tout ensemble a une borne sup
Donc tout ensemble a une borne inf

Théorème : $\leq$ fortement complète et f croissante alors f a un point fixe

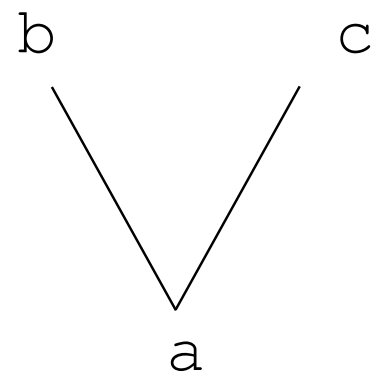
Le plus petit point fixe est $inf\ \{c \mid fc \leq c\}$

Exemple

$[0,1]$, $\leq$ est fortement complète



$\mathbb{R}^+$, $\leq$ est-elle fortement complète ?



Faiblement / fortement complète ?

Un autre exemple

A ensemble quelconque
$\wp(A), \subseteq$ est faiblement et fortement complète

f fonction croissante de $\wp(A)$ dans $\wp(A)$ a un point fixe

Le plus petit point fixe est $\bigcap_{C \mid f\ C \subseteq C} C$
(et aussi $\bigcup_i f^i(\varnothing)$ si f est continue)

Une première définition inductive

$P = 2\mathbb{N}$ est défini par
$0 \in P$ et si $n \in P$ alors $n + 2 \in P$

$$\frac{}{0}$$

$$\frac{n}{n+2}$$

$$\frac{}{0 \in P}$$

$$\frac{n \in P}{n+2 \in P}$$

P n'est pas le seul ensemble qui contient 0 et qui est clos par la fonction $n \mapsto n+2$
Mais c'est le plus petit de ces ensembles

F de $\wp(\mathbb{N})$ dans $\wp(\mathbb{N})$
$F(A) = \{0\} \cup \{x+2 \mid x \in A\}$
F croissante et continue
(A contient 0 et clos par $n \mapsto n+2$) : $F(A) \subseteq A$
P est défini comme le plus petit point fixe de F

Second théorème du point fixe : c'est l'intersection de tous les ensembles qui contiennent 0 et qui sont clos par $n \mapsto n+2$
Premier théorème du point fixe : c'est la réunion de $\varnothing$, $F(\varnothing)$, $F(F(\varnothing))$, ...

Cas général

Un ensemble E
On définit un sous-ensemble B de E
par des fonctions de fermeture (règles) f_1, f_2, ...

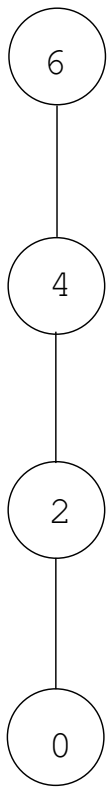
$F(A) = \bigcup_i \{f_i(a_1, ..., a_{n_i}) \mid a_1, ..., a_{n_i} \in A\}$
F croissante et continue
B est le plus petit point fixe de F

La notion de dérivation

$x \in B$ si $x \in F^k(\varnothing)$ pour un certain k

c.-à-d. s'il existe i, $y_1, ..., y_n \in F^{k-1}(\varnothing)$ tq $x = f_i(y_1, ..., y_n)$

Par récurrence sur k si $x \in B$ alors il existe un arbre dont les nœuds sont étiquetés par des éléments de E et les enfants d'un nœud x sont $y_1, ..., y_n$ tq il existe i tq $x = f_i(y_1, ..., y_n)$

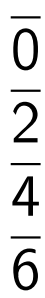


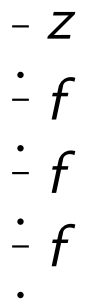
$$\overline{0}$$
$$\overline{2}$$
$$\overline{4}$$
$$\overline{6}$$

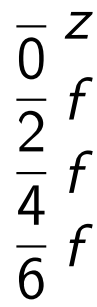
$$\overline{0 \in P}$$
$$\overline{2 \in P}$$
$$\overline{4 \in P}$$
$$\overline{6 \in P}$$

Étiquettes

Un arbre dont les nœuds sont étiquetés par des éléments de E ou par le nom de la règle ou par les deux







Exemple

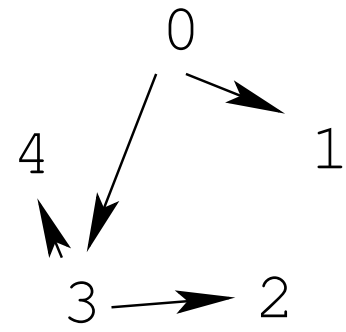
$E = \{a, b\}^*$

$$\overline{b}$$

$$\frac{X}{a\ X\ a}$$

Exemple

$E = \{0, 1, 2, 3, 4\}$



$$\frac{}{x \; C \; x}$$

$$\frac{}{x \; C \; y} \text{ si } x \; R \; y$$

$$\frac{x \; C \; y \quad y \; C \; z}{x \; C \; z}$$

$$\frac{0 \; C \; 3 \quad 3 \; C \; 2}{0 \; C \; 2}$$

La notion de fermeture réflexive-transitive

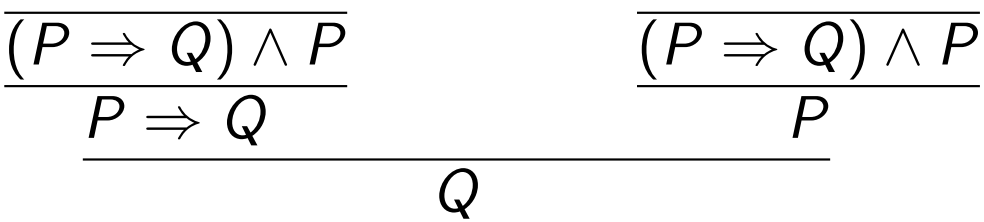
Exemple

$$\frac{A \wedge B}{A}$$

$$\frac{A \wedge B}{B}$$

$$\frac{A \Rightarrow B \quad A}{B}$$

$$\frac{}{(P \Rightarrow Q) \wedge P}$$

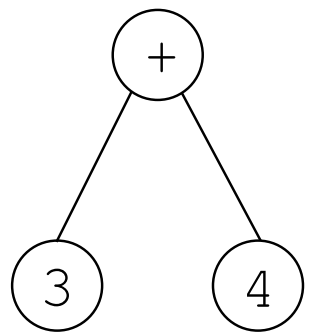

(P ⇒ Q) ∧ P
P ⇒ Q
(P ⇒ Q) ∧ P
P
Q

II. La notion de langage en général

On oublie la contrainte de linéarité du langage

On ne s'intéresse pas à savoir si on écrit $3 + 4$, $+(3, 4)$ ou $34+$

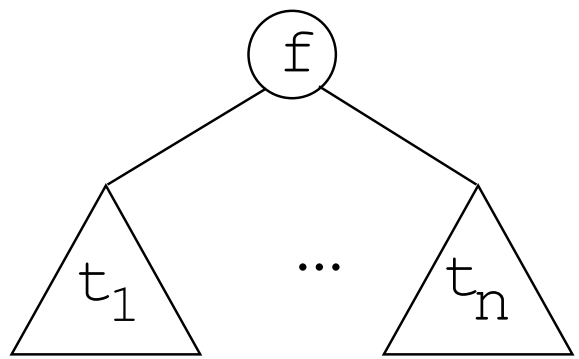
Les expressions sont des arbres



Les langages sans variables

Un langage (sans variables) est un ensemble de symboles, chacun muni d'un nombre entier appelé son arité ou nombre d'arguments
L'ensemble des expressions du langage est l'ensemble d'arbres défini inductivement par la règle

$$\frac{t_1 \quad t_n}{f(t_1, \ldots, t_n)} \text{ si } f \text{ est un symbole d'arité } n$$



Exemple

Une constante (c.-à-d. symbole d'arité nulle) 0
Un symbole unaire S
Deux symboles binaires $+$, $\times$
Deux symboles unaires *pair*, *impair*
Un symbole binaire $\Rightarrow$

$$\mathit{impair}(S(S(S(0)))) \Rightarrow \mathit{pair}(S(S(S(S(0)))))$$

Si un nombre est impair alors son successeur est pair

$$\forall x \ (\mathit{impair}(x) \Rightarrow \mathit{pair}(S(x)))$$

Des variables
Des symboles qui lient des variables

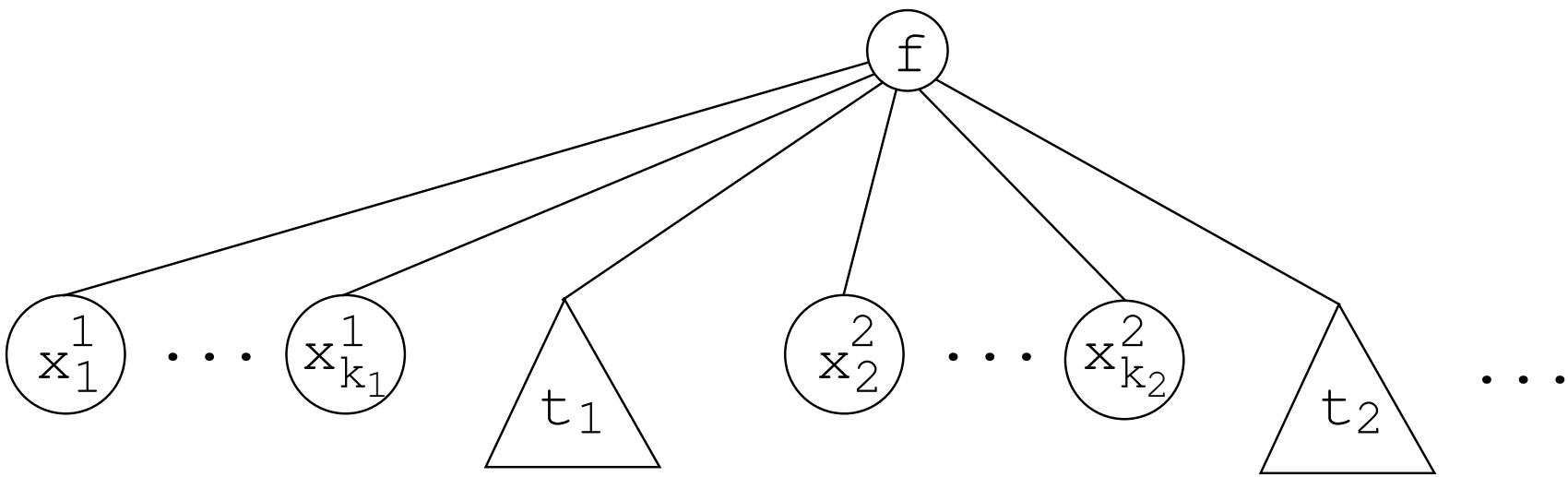
Les langages avec variables

L'arité d'un symbole est un n-uplet $(k_1, ..., k_n)$
le symbole a n arguments, il lie k_1 variables dans le premier, ..., k_n variables dans le $n^{\text{ème}}$
Exemple : $\forall$ a l'arité (1)
Un ensemble de symboles et un ensemble infini de variables

Les expressions sont définies inductivement par les règles :

- les variables sont des expressions,
- si f est un symbole d'arité $(1, 3)$, t et u sont des expressions, w, x, y, z sont des variables alors $f(w\ t, x\ y\ z\ u)$ est une expression (à généraliser)

$f(x_1^1 ... x_{k_1}^1\, t_1, ..., x_1^n ... x_{k_n}^n\, t_n)$ est l'arbre



Les variables et les variables libres

- $Var(x) = \{x\}$,
- $Var(f(x_1^1 ... x_{k_1}^1 \, t_1, ..., x_1^n ... x_{k_n}^n \, t_n))$
  $= Var(t_1) \cup \{x_1^1, ..., x_{k_1}^1\} \cup \cdots \cup Var(t_n) \cup \{x_n^n, ..., x_{k_n}^n\}$.

$Var(\forall x \ (x = x))$?

- $VL(x) = \{x\}$,
- $VL(f(x_1^1 ... x_{k_1}^1 \, t_1, ..., x_1^n ... x_{k_n}^n \, t_n))$
  $= (VL(t_1) \setminus \{x_1^1, ..., x_{k_1}^1\}) \cup \cdots \cup (VL(t_n) \setminus \{x_n^n, ..., x_{k_n}^n\})$

$VL(\forall x \ (x = x))$?

Les langages à plusieurs sortes d'objets

0, S, $+$, $\times$, $pair$, $impair$, $\Rightarrow$, $\forall$

On veut distinguer 0, $S(0)$, $S(x)$, ... termes
de $pair(0)$, $impair(0)$, $\forall x\ (pair(x))$, ... propositions

Mais aussi peut-être les termes de vecteurs, les termes de scalaires, ...

Les langages à plusieurs sortes d'objets

Un ensemble de sortes $\{Terme, Prop\}$ plus généralement $\mathcal{S}$

L'arité d'un symbole est un $n+1$-uplet de sortes $(s_1, ..., s_n, s')$

Si t_1 terme de sorte s_1, t_2 terme de sorte s_2, ..., t_n terme de sorte s_n et f d'arité $(s_1, ..., s_n, s')$ alors $f(t_1, ..., t_n)$ de sorte s'

Plusieurs sortes d'objets + lieurs

$$((s_1^1, ..., s_{k_1}^1, s'^1), ..., (s_1^n, ..., s_{k_n}^n, s'^n), s'')$$

Exemple $\forall$ d'arité $((\mathit{Terme}, \mathit{Prop}), \mathit{Prop})$

III. La notion de langage de la logique des prédicats

Ensemble $\mathcal{S}$ de sortes de termes et une sorte de plus *Prop*

Seulement deux symboles lieurs $\forall$ et $\exists$

Les symboles se divisent en

- les symboles de fonction f d'arité $(s_1, ..., s_n, s')$
- les symboles de prédicat P d'arité $(s_1, ..., s_n, Prop)$ (notée $(s_1, ..., s_n)$)
- les symboles communs à tous les langages $\top$, $\bot$, $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\forall$, $\exists$

$$\forall x \ (pair(x) \Rightarrow impair(S(x)))$$

IV. La notion de proposition démontrable

Une première (et presque bonne) idée

Un sous-ensemble de l'ensemble des propositions inductivement défini par des règles de déduction

$$\frac{A \wedge B}{A}$$

$$\frac{A \wedge B}{B}$$

$$\frac{A \Rightarrow B \quad A}{B}$$

Règles zéro-aires : axiomes

Mais ...

Pour démontrer $A \Rightarrow B$: supposons A et démontrons B

Non seulement la proposition à démontrer varie, mais aussi l'ensemble d'hypothèses

Un séquent $\Gamma \vdash A$ formé d'un ensemble d'hypothèses Γ et d'une conclusion A

Les règles

Un sous-ensemble de l'ensemble des séquents inductivement défini par des règles de déduction

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}$$

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}$$

$$\frac{\Gamma \vdash A \Rightarrow B \quad \Gamma \vdash A}{\Gamma \vdash B}$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \Rightarrow B}$$

La classification des règles

La plupart des règles concernent un symbole (connecteur ou quantificateur) unique

$$\frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \wedge\text{-intro}$$

classification des règles en fonction du symbole concerné

Conclusion ou prémisse : intro / élim
intro / élim : fabriquer / utiliser

Exceptions : axiome, tiers exclu, négation
Négation : symbole composite : $\neg A$ peut être défini comme $A \Rightarrow \bot$

Les règles une par une

axiome : la notion de contexte, raisonnement hypothético-déductif
$\top$: pas d'élim
$\bot$: pas d'intro
$\wedge$: ras
$\vee$: intros triviales, élim démonstration par cas
$\Rightarrow$: intro : la notion de contexte, raisonnement hypothético-déductif
$\neg$: lien avec $\Rightarrow$, les deux formes de raisonnement par l'absurde, forme radicale de raisonnement hypothético-déductif

Les règles une par une

$\forall$: « soit x un objet », notion de généricité (x n'apparaît pas (libre) dans Γ), substitution

$\exists$: substitution, « $\exists x\ P$, appelons le y », y générique

tiers exclu : en déduction naturelle : un cheveu sur la soupe (mais pas dans d'autres systèmes)

La substitution

$\forall$-élim et $\exists$-intro : une opération annexe : la substitution $(t/x)u$
L'opération qui donne son sens au mot variable
Les langages de la logique des prédicats et tous les autres langages

Définition simple pour les langages sans symboles lieurs de var.

- $(t/x)(f(u_1, ..., u_n)) = f((t/x)u_1, ..., (t/x)u_n)$
- $(t/x)x = t$
- $(t/x)y = y$ si $x \neq y$

Dans les langages avec des symboles lieurs de variables

$(4/x)(\forall x\ P(x)) = \forall x\ P(4)$ ou $\forall x\ P(x)$?

Règle 1 : ne substituer que les variables libres

Première tentative :

- $\langle t/x\rangle(\forall y\ A) = \forall y\ (\langle t/x\rangle A)$ si $x \neq y$
- $\langle t/x\rangle(\forall x\ A) = \forall x\ A$

Mais ce n'est pas suffisant

$\langle 4/y \rangle(\forall x\ P(x+y)) = \forall x\ P(x+4)$

$\langle z/y \rangle(\forall x\ P(x+y)) = \forall x\ P(x+z)$

$\langle x/y \rangle(\forall x\ P(x+y)) = \forall x\ P(x+x)$

L'occurrence libre de x a été capturée

Règle 2 : éviter les captures de variables

$(x/y)(\forall x\ P(x+y)) = \forall w\ P(w+x)$

Renommer la variable liée x en w

Pourquoi w plutôt que v ?

C'est équivalent (variable liée = variable muette)

Équivalence alphabétique (α-équivalence)

L'équivalence alphabétique

- $\forall x\ A \sim \forall y\ B$ si pour toute variable z qui n'apparaît ni dans $\forall x\ A$ ni dans $\forall y\ B$ on a $\langle z/x\rangle A \sim \langle z/y\rangle B$

Exemple : $\forall x\ P(x + w)$ et $\forall y\ P(y + w)$ sont équivalents

Désormais on ne raisonne plus que sur des classes d'expressions modulo équivalence alphabétique

La substitution (enfin ...)

- $(t/x)(\forall y\ A) = \forall z\ (t/x)\langle z/y\rangle A$ où z est une variable quelconque différente de x et y et qui n'apparaît ni dans t ni dans A

Généraliser tout cela à la substitution simultanée $t_1/x_1, ..., t_n/x_n$ et à un langage quelconque

Un empilement de notions : substitution avec captures $\rightarrow$ équivalence alphabétique $\rightarrow$ classes d'expressions $\rightarrow$ substitution

De nombreuses erreurs dans les livres

De nombreuses erreurs dans les systèmes de calcul symbolique (langages de programmation, systèmes de calcul formel, systèmes de traitement de démonstrations, ...)

V. Comment démontrer qu'une proposition n'est pas démontrable ? La notion de modèle

La notion de modèle

Un langage $\mathcal{L} = (\mathcal{S}, \mathcal{F}, \mathcal{P})$
Un modèle de ce langage est formé de

- pour chaque s, un ensemble non vide $\mathcal{M}_s$
- un ensemble non vide $\mathcal{B}$, un sous-ensemble $\mathcal{B}^+$
- pour chaque f d'arité $(s_1, ..., s_n, s')$, une fonction $\hat{f}$ de $\mathcal{M}_{s_1} \times ... \times \mathcal{M}_{s_n}$ dans $\mathcal{M}_{s'}$
- pour chaque P d'arité $(s_1, ..., s_n)$, une fonction $\hat{P}$ de $\mathcal{M}_{s_1} \times ... \times \mathcal{M}_{s_n}$ dans $\mathcal{B}$
- $\hat{\top}$, $\hat{\bot}$, $\hat{\neg}$, $\hat{\wedge}$, $\hat{\vee}$, $\hat{\Rightarrow}$, $\hat{\forall}$, $\hat{\exists}$

Un langage et un modèle de ce langage

Une fonction $[\![\ ]\!]$ qui associe

- à chaque terme t un élément $[\![t]\!]$ de $\mathcal{M}_s$ (s sorte de t)
- à chaque proposition A, un élément $[\![A]\!]$ de $\mathcal{B}$

Morphisme :
$[\![f(t_1, ..., t_n)]\!] = \hat{f}([\![t_1]\!], ..., [\![t_n]\!])$
$[\![P(t_1, ..., t_n)]\!] = \hat{P}([\![t_1]\!], ..., [\![t_n]\!])$
$[\![A \wedge B]\!] = \hat{\wedge}([\![A]\!], [\![B]\!])$ ...

Combien de fonctions possibles ?

Une seule

si on se limite aux termes et propositions sans variables
Mais plusieurs si on a des variables
La fonction $[\![\ ]\!]$ complètement définie par sa valeur sur les variables
(idem morphisme d'e.v. défini par son image sur une base)

Valuation : fonction de domaine fini qui associe aux variables
$x_1, ..., x_n$ de sortes $s_1, ..., s_n$ des éléments $a_1, ..., a_n$ de $\mathcal{M}_{s_1}$, ..., $\mathcal{M}_{s_n}$
$\phi = (x_1 = a_1, ..., x_n = a_n)$

$[\![\ ]\!]_\phi$
$[\![x]\!]_\phi = \phi(x)$, $[\![f(t_1, ..., t_n)]\!]_\phi = \hat{f}([\![t_1]\!]_\phi, ..., [\![t_n]\!]_\phi)$, ...
Toutes les expressions dont les variables sont dans le domaine de ϕ

$[\![\forall x\ A]\!]_\phi$?

$[\![\forall x\ A]\!]_\phi = \hat{\forall}([\![A]\!]_\phi)$

$VL(A) \subseteq VL(\forall x\ A) \cup \{x\}$
On considère l'ensemble de toutes les valeurs $[\![A]\!]_{\phi,x=a}$
Et c'est à cet ensemble qu'on applique $\hat{\forall}$ ou $\tilde{\exists}$
Fonctions de $\wp^+(\mathcal{B})$ dans $\mathcal{B}$

Pour une proposition A sans variables $[\![A]\!]_\phi$ indépendant de ϕ
La notion de valuation inutile

Pour une proposition A close également
Mais la notion de valuation nécessaire pour les sous-expressions

À quoi sert $\mathcal{B}^+$?

Un langage $\mathcal{L} = (\mathcal{S}, \mathcal{F}, \mathcal{P})$
Un modèle de ce langage est formé de

- pour chaque s, un ensemble non vide $\mathcal{M}_s$
- un ensemble non vide $\mathcal{B}$, un sous-ensemble $\mathcal{B}^+$
- pour chaque f d'arité $(s_1, ..., s_n, s')$, une fonction $\hat{f}$ de $\mathcal{M}_{s_1} \times ... \times \mathcal{M}_{s_n}$ dans $\mathcal{M}_{s'}$
- pour chaque P d'arité $(s_1, ..., s_n)$, une fonction $\hat{P}$ de $\mathcal{M}_{s_1} \times ... \times \mathcal{M}_{s_n}$ dans $\mathcal{B}$
- $\hat{\top}$, $\hat{\bot}$, $\hat{\neg}$, $\hat{\wedge}$ $\hat{\vee}$, $\hat{\Rightarrow}$, $\hat{\forall}$, $\hat{\exists}$

À quoi sert $\mathcal{B}^+$?

Une proposition close A est valide dans un modèle si $[\![A]\!] \in \mathcal{B}^+$

Une proposition A qui a des variables libres $x_1, ..., x_n$ est valide si $\forall x_1 \ ... \ \forall x_n \ A$ est valide

Un séquent $A_1, ..., A_n \vdash B$ est valide si la proposition $(A_1 \wedge ... \wedge A_n) \Rightarrow B$ est valide

Un cas particulier : les modèles bivalués

$\mathcal{B} = \{0, 1\}$
$\mathcal{B}^+ = \{1\}$

$\hat{\top} = 1$, $\hat{\bot} = 0$, $\hat{\neg}$, $\hat{\wedge}$, $\hat{\vee}$, $\hat{\Rightarrow}$ ...
Que sont $\hat{\forall}$ et $\hat{\exists}$?
Désormais : tous les modèles sont bivalués

Un exemple

Langage : une seule sorte, une constante c, deux prédicats unaires P et Q

$\mathcal{M} = \{\pi, e\}$
$\hat{c} = \pi$,
$\hat{P}$ est la fonction qui associe 0 à π et 1 à e
$\hat{Q}$ est la fonction qui associe 1 à π et 1 à e

Est-ce que $P(c)$ est valide ? $Q(c)$? $P(c) \vee Q(c)$? $\forall x\ P(x)$?
$\exists x\ P(x)$? $\forall x\ Q(x)$? $\exists x\ Q(x)$?

Un autre exemple

Langage : une sorte, symbole de fonction binaire $+$, symbole de prédicat binaire $=$

($\mathbb{N}$, addition sur $\mathbb{N}$, égalité sur $\mathbb{N}$) $\forall x \forall y \exists z \ (x + z = y)$ est-elle valide ?
Même question pour $\mathbb{Z}$ muni de l'addition et de l'égalité sur $\mathbb{Z}$?

La proposition $\forall x \forall y \ (x + y = y + x)$ est-elle valide dans ces modèles ? Un modèle dans lequel elle n'est pas valide ?

Quel rapport avec la question

Comment démontrer qu'une proposition n'est pas démontrable ?

Le théorème de correction (démontré la semaine prochaine)

Si $\Gamma \vdash A$ démontrable, alors $\Gamma \vdash A$ valide dans tous les modèles

On contrapose : s'il existe un modèle où $\Gamma \vdash A$ n'est valide, alors $\Gamma \vdash A$ n'est pas démontrable

Exercice : $P(a) \vee P(b) \vdash P(a)$ non démontrable

La prochaine fois : Le théorème de correction et sa réciproque