La théorie de l'unification

Les Σ-algèbres

Σ : Ensemble de symboles de fonction ayant une arité $n \in \mathbb{N}$.
$\mathcal{X}$: Ensemble de variables.

$\mathcal{T}(\mathcal{X}, \Sigma)$: Ensemble de termes sur $\mathcal{X}$ et Σ :

$$\frac{x \in \mathcal{X}}{x \in \mathcal{T}(\mathcal{X}, \Sigma)} \qquad \frac{t_1, \ldots, t_n \in \mathcal{T}(\mathcal{X}, \Sigma) \quad f \text{ est d'arité } n \in \Sigma}{f(t_1, \ldots, t_n) \in \mathcal{T}(\mathcal{X}, \Sigma)}$$

On écrit $Var(t)$ l'ensemble de toutes les variables de t. Un terme t est clos si $Var(t) = \emptyset$.

Les substitutions

Definition

- Une substitution est une fonction $\sigma : \mathcal{X} \to \mathcal{T}(\Sigma, \mathcal{X})$.
- Le domaine d'une substitution σ est l'ensemble $Dom(\sigma) = \{x \in \mathcal{X} \mid \sigma(x) \neq x\}$.
- Le codomaine d'une substitution σ est l'ensemble $Codom(\sigma) = \{Var(\sigma(x)) \mid x \in Dom(\sigma)\}$.
- Un renommage est une substitution injective σ t.q. $\sigma(x) = y$ $\forall x \in Dom(\sigma)$.
  Exemple : $\sigma = \{x/y, y/w\}$ est un renommage. Toute permutation est un renommage mais pas l'inverse comme le montre l'exemple.
- Si le domaine d'une substitution σ est fini on note $\sigma = \{x_1/t_1, \ldots, x_n/t_n\}$ si $\sigma(x_i) = t_i$ et $x_i \in Dom(\sigma)$.
- L'application d'une substitution à un terme est l'extension de σ aux termes donnée par $\sigma(f(t_1, \ldots, t_n)) = f(\sigma(t_1), \ldots, \sigma(t_n))$.

Comparer deux substitutions

Soient σ et τ deux substitution. La composition de σ avec τ est donnée par $(\sigma \circ \tau)(x) = \sigma(\tau(x))$.

Exemple : $\{y/b, z/h(c)\} \circ \{x/f(y), y/z\} = \{x/f(b), y/h(c), z/h(c)\}$

La substitution σ est une instance de la substitution τ (ou τ est plus générale que σ), ce que l'on écrit $\sigma \leq \tau$, ss'il existe une substitution ρ t.q. pour toute variable $x \in \mathcal{X}$, $\sigma(x) = (\rho \circ \tau)(x)$.

Exemple : $\{x/f(y), y/z\}$ est plus générale que $\{x/f(b), y/h(c), z/h(c)\}$

Identifier deux substitutions

Remarque : La relation $\leq$ n'est pas antysimmetrique.

Exemple : Soient $\sigma_1 = \{x/y\}$ et $\sigma_2 = \{y/x\}$.
On a $\sigma_1 \leq \sigma_2$ car $\sigma_1 = \{x/y\} \circ \sigma_2$.
On a $\sigma_2 \leq \sigma_1$ car $\sigma_2 = \{y/x\} \circ \sigma_1$.
Mais $\sigma_1 \neq \sigma_2$.

Exemple : Soient $\sigma_1 = \{x/y\}$ et $\sigma_3 = \{x/y, z/w, w/z\}$.
On a $\sigma_1 \leq \sigma_3$ car $\sigma_1 = \{z/w, w/z\} \circ \sigma_3$.
On a $\sigma_3 \leq \sigma_1$ car $\sigma_3 = \{z/w, w/z\} \circ \sigma_1$.
Mais $\sigma_1 \neq \sigma_3$.

Lemma

$\sigma \sim \sigma'$ ssi $\exists$ un renommage ρ t.q. $\sigma = \rho \circ \sigma'$.

Donc $\sigma_1 \sim \sigma_2 \sim \sigma_3$ dans l'exemple précédent.

Substitution(s) principale(s)

Soit $\mathcal{S}$ en ensemble de substitutions et $\tau \in \mathcal{S}$. On dit que τ est principale ssi toute substitution $\sigma \in \mathcal{S}$ est une instance de τ.

Exemple : Soit $\mathcal{S} = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5\}$, où $\sigma_1 = \{x/y\}$, $\sigma_2 = \{y/x\}$, $\sigma_3 = \{x/y, w/z, z/w\}$, $\sigma_4 = \{x/u, y/u\}$ et $\sigma_5 = \{x/a, y/a\}$.

Alors σ_1, σ_2 et σ_3 sont principales pour $\mathcal{S}$.
En effet,
$\sigma_2, \sigma_3, \sigma_4, \sigma_5 \leq \sigma_1$ et $\sigma_1, \sigma_3, \sigma_4, \sigma_5 \leq \sigma_2$ et $\sigma_1, \sigma_2, \sigma_4, \sigma_5 \leq \sigma_3$
Mais $\sigma_1 \not\leq \sigma_4$ et $\sigma_1 \not\leq \sigma_5$ (entre autres).

Unification comme solution d'un système d'équations

Definition

Deux termes A et B sont unifiables ss'il existe une substitution σ t.q. $\sigma(A) = \sigma(B)$ (σ est donc un unificateur de A et B).

Une équation est une paire de termes de la forme $A \doteq B$. On dit qu'elle est unifiable ssi les termes A et B le sont.
Un système/problème d'équations E est un ensemble d'équations. On dit qu'il est unifiable ssi il existe une substitution qui est unificateur de toutes les équations de E. Cette substitution est appelée solution de E.
On s'intéresse aux systèmes d'équations finis.

Exemple : $f(x, g(x, a))$ et $f(f(a), y)$ sont unifiables avec l'unificateur $\{x/f(a), y/g(f(a), a)\}$.
$f(x, g(x, a))$ et $f(f(a), f(b, a))$ ne le sont pas.

L'unicité

1. On identifie deux unificateurs σ et σ' d'un problème $\mathcal{P}$ s'ils ne différent que par des renommage de variables, c'est à dire, si $\sigma \sim \sigma'$.
2. On considère uniquement comme unificateurs de $\mathcal{P}$ les substitutions σ t.q. $Dom(\sigma) \subseteq Var(\mathcal{P})$.

Exemple : Soit $\mathcal{S} = \{x \doteq y\}$. Prenons trois unificateurs principaux de $\mathcal{S}$: $\sigma_1 = \{x/y\}$,$\sigma_2 = \{y/x\}$ et $\sigma_3 = \{x/y, z/w, w/z\}$.
Alors $\sigma_1 = \sigma_2$ (car $[\sigma_1]_\sim = [\sigma_2]_\sim$) et σ_3 n'est plus considéré comme un unificateurs de $\mathcal{S}$.

L'unicité

Module ces considérations, l'unificateur pricipal d'un problème $\mathcal{P}$ est unique modulo renommage, c'est à dire :
Si σ et σ' sont deux unificateurs principaux de $\mathcal{P}$, alors $\sigma \sim \sigma'$.

Les formes résolues

Definition

Un système d'équations E est en forme résolue ssi il est de la forme $\{\alpha_1 \doteq t_1, \ldots, \alpha_n \doteq t_n\}$, où

- toutes les variables α_i sont distinctes ($i \neq j$ implique $\alpha_i \neq \alpha_j$)
- aucune α_i n'apparaît dans un t_j ($\forall i, \alpha_i \notin \bigcup_{1 \leq j \leq n} Var(t_j)$)

Notation : Si E est un système en forme résolue $\{\alpha_1 \doteq t_1, \ldots, \alpha_n \doteq t_n\}$ on note $\vec{E}$ la substitution $\{\alpha_1/t_1, \ldots, \alpha_n/t_n\}$.

Les règles de transformation

$$\frac{E \cup \{s \doteq s\}}{E} \quad \text{(effacer)} \qquad \frac{E \cup \{t \doteq \alpha\} \quad t \notin \mathcal{X}}{E \cup \{\alpha \doteq t\}} \quad \text{(orienter)}$$

$$\frac{E \cup \{f(s_1, \ldots, s_n) \doteq f(t_1, \ldots, t_n)\}}{E \cup \{s_1 \doteq t_1, \ldots, s_n \doteq t_n\}} \quad \text{(décomposer)}$$

$$\frac{E \cup \{\alpha \doteq s\} \quad \alpha \in Var(E) \quad \alpha \notin Var(s)}{E\{\alpha/s\} \cup \{\alpha \doteq s\}} \quad \text{(remplacer)}$$

Algorithme d'unification d'un système E

1. On démarre avec un système E
2. On applique les règles de transformation tant qu'on peut, on obtient un problème P
3. Si le système P est en forme résolue
    - alors renvoyer $\vec{P}$.
    - sinon échec

Exemple

Soit $\mathcal{P} = \{f(x, h(b), c) \doteq f(g(y), y, c)\}$.

$$\cfrac{\cfrac{\cfrac{\cfrac{f(x, h(b), c) \doteq f(g(y), y, c)}{x \doteq g(y), h(b) \doteq y, c \doteq c}\,\mathsf{d}}{x \doteq g(y), h(b) \doteq y}\,\mathsf{e}}{x \doteq g(y), y \doteq h(b)}\,\mathsf{o}}{x \doteq g(h(b)), y \doteq h(b)}\,\mathsf{r}$$

L'unificateur principal de P est $\sigma = \{x/g(h(b)), y/h(b)\}$.
Ainsi, $\sigma f(x, h(b), c) = f(g(h(b)), h(b), c) = \sigma f(g(y), y, c)$.

Vers la correction et la complétude de l'algorithme

Lemma

1. *L'algorithme termine.*
2. *Si σ est un unificateur d'une forme résolue P, alors $\sigma = \sigma\vec{P}$.*
3. *Si une règle transforme un problème P dans un problème S, alors les solutions de P et S sont les mêmes.*
4. *Si E est en forme résolue, alors $\vec{E}$ est solution du problème E.*

Démonstration de la terminaison

Une variable est non résolue dans un problème E si elle y apparaît plus d'une fois. La terminaison de l'algorithme d'unification est par récurrence sur le triplet $\langle n1, n2, n3 \rangle$ muni de l'ordre lexicographique, où

n1 :	nb de variables pas résolues
n2 :	taille du problème
n3 :	nb d'équations de la forme $t = x$

En effet,

	n1	n2	n3
Remplacement	$>$		
Effacer	$\geq$	$>$	
Decomposer	$=$	$>$	
Orienter	$=$	$=$	$>$

Correction et complétude de l'algorithme

Théorème : (Correction) Si l'algorithme trouve une substitution $\vec{S}$ pour le problème P, alors P est unifiable et $\vec{S}$ est un unificateur principal de P.
Autrement dit,
Si P n'est pas unifiable, l'algorithme échoue.

Théorème : (Complétude) Si le système P est unifiable, alors l'algorithme calcule l'unificateur principal de P.
Autrement dit,
Si l'algorithme échoue, alors le système P n'est pas unifiable.