
La théorie de l'unification

Retour sur la notion de substitution

Définition :

- Une **substitution** est une fonction $\sigma : \mathcal{X} \rightarrow \mathcal{T}_{\Sigma, \mathcal{X}}$.
- 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 $Codom(\sigma) = \{VI(\sigma(x)) \mid x \in Dom(\sigma)\}$.
- Un **renommage** est une substitution **injective** σ t.q. pour tout $x \in Dom(\sigma)$ il y a un $y \in \mathcal{X}$ tel que $\sigma(x) = y$.
- Si le domaine d'une substitution σ est **fini** on note $\sigma = \{x_1 \leftarrow t_1, \dots, x_n \leftarrow t_n\}$ 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 $\sigma(f(t_1, \dots, t_n)) = f(\sigma(t_1), \dots, \sigma(t_n))$.

Composition de 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 : Soit $\sigma = \{x \leftarrow f(y), w \leftarrow g(z, z)\}$ et $\tau = \{y \leftarrow f(a), z \leftarrow g(x, b)\}$. La substitution $\tau \circ \sigma$ est donnée par $\{x \leftarrow f(f(a)), y \leftarrow f(a), w \leftarrow g(g(x, b), g(x, b)), z \leftarrow g(x, b)\}$ et la substitution $\sigma \circ \tau$ est donnée par $\{y \leftarrow f(a), z \leftarrow g(f(y), b), x \leftarrow f(y), w \leftarrow g(z, z)\}$.

Comparer deux substitutions

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

Identifier deux substitutions

Remarque : La relation \leq n'est pas antisymétrique.

Exemple : Soient $\sigma_1 = \{x \leftarrow y\}$ et $\sigma_2 = \{y \leftarrow x\}$. On a $\sigma_1 \leq \sigma_2$ et $\sigma_2 \leq \sigma_1$ et $\sigma_1 \neq \sigma_2$.

Lemme : La relation d'équivalence engendrée par \leq est donnée par:
 $\sigma \sim \sigma'$ ssi \exists un renommage ρ t.q. $\sigma = \rho \circ \sigma'$.

Alors, $\sigma_1 \sim \sigma_2$ dans l'exemple précédent car:
 $\sigma_1 = \sigma_1 \circ \sigma_2$ et $\sigma_2 = \sigma_2 \circ \sigma_1$.

Substitution(s) principale(s)

Soit \mathcal{S} un 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 $\sigma_1 = \{x \leftarrow y\}$,
 $\sigma_2 = \{y \leftarrow x\}$, $\sigma_3 = \{x \leftarrow y, z \leftarrow x\}$, $\sigma_4 = \{x \leftarrow z, y \leftarrow z\}$ et
 $\sigma_5 = \{x \leftarrow a, y \leftarrow a\}$.

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 $\sigma_1 \not\leq \sigma_4$ car $\sigma_1 \neq \{x \leftarrow y, z \leftarrow y\} = \{z \leftarrow y\} \circ \sigma_4$. De même,
 $\sigma_1 \not\leq \sigma_5$ (entre autres).

Unification comme solution d'un système d'équations

Une **équation** est une paire de termes de la forme $s \doteq t$, elle est **unifiable** ssi il existe une substitution σ t.q. $\sigma(s) = \sigma(t)$.
Cette substitution est un **unificateur** ou une **solution** de l'équation $s \doteq t$.

Un **système fini** ou **problème fini d'équations** \mathcal{P} est un ensemble $\{s_1 \doteq t_1, \dots, s_n \doteq t_n\}$ d'équations, il est **unifiable** ssi il existe une substitution qui est unificateur de toutes les équations de \mathcal{P} . Cette substitution est un **unificateur** ou une **solution** de l'ensemble \mathcal{P} .

Définition :

- L'**ensemble de variables** de \mathcal{P} est notée $Var(\mathcal{P})$.
- L'**application d'une substitution** σ à $\mathcal{P} = \{s_1 \doteq t_1, \dots, s_n \doteq t_n\}$ donne le système $\sigma(\mathcal{P}) = \{\sigma(s_1) \doteq \sigma(t_1), \dots, \sigma(s_n) \doteq \sigma(t_n)\}$.

L'unicité

- 1 On identifie deux unificateurs σ et σ' d'un problème \mathcal{P} s'ils ne diffèrent 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} :
 $\sigma_1 = \{x \leftarrow y\}$, $\sigma_2 = \{y \leftarrow x\}$ et $\sigma_3 = \{x \leftarrow y, z \leftarrow w\}$.
Alors $\sigma_1 = \sigma_2$ (car $[\sigma_1]_{\sim} = [\sigma_2]_{\sim}$) et σ_3 n'est plus considéré comme un unificateur de \mathcal{S} .

Modulo ces considérations, l'**unificateur principal 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 $\sigma \sim \sigma'$.

Définition : Un problème d'unification \mathcal{P} est en **forme résolue** ssi il est de la forme $\{x_1 \doteq t_1, \dots, x_n \doteq t_n\}$, où

- ❶ toutes les variables x_i sont distinctes ($i \neq j$ implique $x_i \neq x_j$)
- ❷ aucune x_i n'apparaît dans un t_j ($\forall i \forall j x_i \notin VI(t_j)$)

Notation : Si \mathcal{P} est un système en forme résolue $\{x_1 \doteq t_1, \dots, x_n \doteq t_n\}$ on note $\vec{\mathcal{P}}$ la substitution $\{x_1 \leftarrow t_1, \dots, x_n \leftarrow t_n\}$.

Algorithme d'unification d'un problème \mathcal{P}

- ❶ On démarre avec un problème \mathcal{P}
- ❷ On applique les règles de transformation tant qu'on peut, on obtient un problème \mathcal{S}
- ❸ Si le problème \mathcal{S} est en forme résolue
 - ▶ alors renvoyer $\vec{\mathcal{S}}$
 - ▶ sinon échec

$$\frac{\mathcal{P} \cup \{s \doteq s\}}{\mathcal{P}} \quad (\text{effacer}) \quad \frac{\mathcal{P} \cup \{t \doteq x\} \quad t \notin \mathcal{X}}{\mathcal{P} \cup \{x \doteq t\}} \quad (\text{orienter})$$

$$\frac{\mathcal{P} \cup \{f(s_1, \dots, s_n) \doteq f(t_1, \dots, t_n)\}}{\mathcal{P} \cup \{s_1 \doteq t_1, \dots, s_n \doteq t_n\}} \quad (\text{décomposer})$$

$$\frac{\mathcal{P} \cup \{x \doteq s\} \quad x \in \text{Var}(\mathcal{P}) \quad x \notin VI(s)}{\{x \leftarrow s\}(\mathcal{P}) \cup \{x \doteq s\}} \quad (\text{remplacer})$$

Exemple

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

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

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

Lemme :

- 1 L'algorithme termine.
- 2 Si σ est un unificateur d'un problème $\mathcal{P} = \{x_1 \doteq t_1, \dots, x_n \doteq t_n\}$, alors $\sigma = \sigma \circ \vec{P}$.
- 3 Si une règle transforme un problème \mathcal{P} dans un problème \mathcal{S} , alors les unificateurs de \mathcal{P} et \mathcal{S} sont les mêmes.
- 4 Si \mathcal{P} est en forme résolue, alors \vec{P} est solution du problème \mathcal{P} .

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.