UE ELFE - Expression Logique et Fonctionnelle ... Évidemment

Notes de cours

Unification

Pour une signature donnée Σ, on rappelle qu'un **terme** $t \in \mathcal{T}_\Sigma$ est
- soit une *constante* c (i.e. un atome ou un nombre) ;
- soit une *variable* $x, y, \dots$;
- soit un *terme complexe* de la forme $f(t_1, \dots, t_n)$ où $t_1, \dots, t_n$ sont eux-même des termes de $\mathcal{T}_\Sigma$ et f un symbole de fonction d'arité n.

On note $\mathit{var}(t_1, \dots, t_n)$ l'ensemble des variables apparaîssant dans tous les termes $t_1, \dots, t_n$.

1 Introduction

2 Substitutions

Une **substitution** (notée σ ou μ) est une fonction de l'ensemble des variables dans les termes : $\sigma : \mathit{Var} \to \mathcal{T}_\Sigma$. On ne s'intéresse qu'à des substitutions qui sont presque partout l'identité (i.e. telles qu'on a $\sigma(x) \neq x$ pour un ensemble fini de variables x, ensemble qu'on appelera, abusivement, *domaine* de σ, noté $\mathit{dom}(\sigma)$). On note

$$\sigma \quad = \quad \{t_1/x_1, \dots, t_n/x_n\} \tag{1}$$

pour la substitution qui substitue t_i à la variable x_i pour tout $i \in \{1, \dots, n\}$ et de domaine $\{x_1, \dots, x_n\}$ [1].
On étend le domaine d'application des substitutions aux termes t ainsi :
- si $t = x$ alors $\sigma(t) = \sigma(x)$,
- si t est une constante alors $\sigma(t) = t$,
- si $t = f(t_1, \dots, t_m)$ alors $\sigma(t) = f(\sigma(t_1), \dots \sigma(t_m))$.

On appliquera aussi les substitutions à des ensembles de termes résultant naturellement en l'ensemble des termes auxquels la substitution a était appliquée :

$$\sigma(\{t_1, \dots, t_n\}) = \{\sigma(t_1), \dots, \sigma(t_n)\}$$

On note ϵ la **substitution vide** ou **identité** i.e., telle que $\mathit{dom}(\epsilon) = \emptyset$.
On ne considérera que des substitutions qui ne réintroduisent pas une variable substituée ; c'est-à-dire, pour σ définie en (1), $\{x_1, \dots, x_n\} \cap \mathit{var}(t_1, \dots, t_n) = \emptyset$. L'algorithme d'unification itératif donné ci-dessous vérifie cette propriété par un test dit d'*occurs-check*. Notez qu'elle est équivalente à la propriété d'idempotence de l'opération de composition sur les substitutions (i.e. $\sigma\sigma = \sigma$).
Soient les substutitions

$$\sigma = \{t_1/x_1, \dots, t_n/x_n\} \quad \text{et} \quad \mu = \{u_1/y_1, \dots, u_m/y_m\}$$

avec $\{x_1, \dots, x_n\} \cap \{y_1, \dots, y_m\} = \emptyset$ et $\{y_1, \dots, y_m\} \cap \mathit{var}(t_1, \dots, t_n) = \emptyset$, leur composition est donc :

$$\sigma\mu = \{t_1/x_1, \dots, t_n/x_n, \sigma(u_1)/y_1, \dots, \sigma(u_m)/y_m\}$$

On peut définir un **préordre $\leq$ sur les substitutions** ainsi : $\mu \leq \sigma$ si et seulement s'il existe une substitution μ' telle que $\sigma = \mu'\mu$. On dit alors que μ est **plus générale** que σ ou que σ est une **instance** de μ.

---

1. on autorisera aussi la notation $\{x_1 \mapsto t_1, \dots, x_n \mapsto t_n\}$.

3 Algorithme d'unification

Le problème d'**unification** peut être énoncé ainsi :

étant donné un système d'équations entre termes,
existe-t-il une substitution qui satisfait les équations ?

En d'autres termes, soit un système d'équations $E = \{t_1 = u_1, \ldots, t_n = u_n\}$, on cherche la substitution σ telle que pour tout $i \in \{1, \ldots, n\}$ on a $\sigma(t_i) = \sigma(u_i)$. Une telle substitution est appelée **unificateur**. Si celle-ci existe, on dit que l'ensemble E est **unifiable**.
Quand E est unifiable alors E admet potentiellement une infinité d'unificateurs :

Proposition 1. *Si μ est un unificateur d'un système E et $\mu \leq \sigma$ alors σ est aussi un unificateur de E.*

On dit qu'un système E est en **forme résolue** s'il est de la forme

$$E = \{x_1 = t_1, \ldots, x_n = t_n\}$$

où les x_i sont deux à deux distincts et $\{x_1, \ldots, x_n\} \cap \mathit{var}(t_1, \ldots, t_n) = \emptyset$.

Proposition 2. *Si E est le système en forme résolue $\{x_1 = t_1, \ldots, x_n = t_n\}$, alors il est unifiable et la substitution $\mu = \{t_1/x_1, \ldots, t_n/x_n\}$ est un plus petit unificateur pour $\leq$.*

Pour un système d'équations E donné, le plus petit unificateur est aussi appelé **unificateur le plus général** (noté *upg*). Dans notre cas, l'*upg* est unique (au renommage[2] prêt des variables car $\leq$ n'est qu'un préordre). Plus formellement, μ est un *upg* pour E si et seulement si pour tout unificateur σ de E on a $\mu \leq \sigma$.
On dit que deux systèmes d'équations E_1 et E_2 sont **équivalents** (noté $E_1 \sim E_2$) si et seulement s'ils admettent exactement les même unificateurs.

L'**algorithme d'unification** d'un système d'équation E consiste à ramener E à une forme résolue E' équivalente pour laquelle on connaît immédiatement l'*upg* (qui est donc également celui de E). L'algorithme est une suite d'étapes qui *réduit* le système en un système équivalent et plus proche de la forme résolue, ou bien détecte que le système n'a pas de solution. Une étape de réduction est noté $E \to E'$ ou bien $E \to\perp$ s'il est détecté que E n'a pas de solution. On note $\to^*$ la clôture transitive de $\to$. Résoudre E c'est donc appliquer une suite de transformations $E \to^* E'$ tel que E' est en forme résolue ou bien $E \to^*\perp$. Les étapes de réductions autorisées, formalisant l'algorithme, sont les suivantes :

$$\begin{array}{rl}
\text{(effacement)} & \{t = t\} \uplus E \to E \\
\text{(décomposition)} & \{f(t_1, \ldots, t_n) = f(u_1, \ldots, u_n)\} \uplus E \to \{t_1 = u_1, \ldots, t_n = u_n\} \uplus E \\
\text{(élimination)} & \{x = t\} \uplus E \to \{t/x\}(E) \uplus \{x = t\} \quad \text{si } t \neq x \text{ et } x \notin \mathit{var}(t) \\
\text{(inversion)} & \{t = x\} \uplus E \to \{x = t\} \uplus E \quad \text{si } t \notin \mathit{Var} \\
(\text{conflit}_1) & \{f(t_1, \ldots, t_n) = g(u_1, \ldots, u_m)\} \uplus E \to\perp \quad \text{si } f \neq g \text{ ou } n \neq m \\
(\text{conflit}_2) & \{a = b\} \uplus E \to\perp \quad \text{si } a \neq b \\
\text{(occurs check)} & \{x = t\} \uplus E \to\perp \quad \text{si } x \in \mathit{var}(t)
\end{array}$$

Théorème 1. *L'algorithme d'unification a les propriétes suivantes :*
- *la relation $\to$ est bien fondée ;*
- *si $E \to^* E' \not\to$, alors E' est sous forme résolue équivalente à E avec $\mathit{var}(E') = \mathit{var}(E)$, et E' définit un upg de E.*
- *si $E \to^*\perp$, alors E' n'est pas unifiable.*

Corollaire 1. *Le problème d'unification est décidable.*

2. un renommage est une substitution injective qui substitue des variables aux variables.