Unification

3.7.1 Definition (Unifier)

Two terms s and t of the same sort are said to be *unifiable* if there exists a well-sorted substitution σ so that $s\sigma = t\sigma$, the substitution σ is then called a well-sorted *unifier* of s and t.

The unifier σ is called *most general unifier*, written $\sigma = mgu(s, t)$, if any other well-sorted unifier τ of s and t it can be represented as $\tau = \sigma\tau'$, for some well-sorted substitution τ'.

A state of the naive standard unification calculus is a set of equations E or $\bot$, where $\bot$ denotes that no unifier exists. The set E is also called a *unification problem*.

The start state for checking whether two terms s, t, $\mathrm{sort}(s) = \mathrm{sort}(t)$, (or two non-equational atoms A, B) are unifiable is the set $E = \{s = t\}$ ($E = \{A = B\}$). A variable x is *solved* in E if $E = \{x = t\} \uplus E'$, $x \not\in \mathrm{vars}(t)$ and $x \not\in \mathrm{vars}(E)$.

A variable $x \in \mathrm{vars}(E)$ is called *solved* in E if $E = E' \uplus \{x = t\}$ and $x \not\in \mathrm{vars}(t)$ and $x \not\in \mathrm{vars}(E')$.

Standard (naive) Unification

Tautology $E \uplus \{t = t\} \;\Rightarrow_{\text{SU}}\; E$

Decomposition $E \uplus \{f(s_1, \ldots, s_n) = f(t_1, \ldots, t_n)\} \;\Rightarrow_{\text{SU}}$
$E \cup \{s_1 = t_1, \ldots, s_n = t_n\}$

Clash $E \uplus \{f(s_1, \ldots, s_n) = g(s_1, \ldots, s_m)\} \;\Rightarrow_{\text{SU}}\; \perp$
if $f \neq g$

Substitution $\quad E \uplus \{x = t\} \;\Rightarrow_{\mathrm{SU}}\; E\{x \mapsto t\} \cup \{x = t\}$

if $x \in \mathit{vars}(E)$ and $x \notin \mathit{vars}(t)$

Occurs Check $\quad E \uplus \{x = t\} \;\Rightarrow_{\mathrm{SU}}\; \perp$

if $x \neq t$ and $x \in \mathit{vars}(t)$

Orient $\quad E \uplus \{t = x\} \;\Rightarrow_{\mathrm{SU}}\; E \cup \{x = t\}$

if $t \notin \mathcal{X}$

3.7.2 Theorem (Soundness, Completeness and Termination of $\Rightarrow_{\mathrm{SU}}$)

If s, t are two terms with $\mathrm{sort}(s) = \mathrm{sort}(t)$ then

1. if $\{s = t\} \Rightarrow^*_{\mathrm{SU}} E$ then any equation $(s' = t') \in E$ is well-sorted, i.e., $\mathrm{sort}(s') = \mathrm{sort}(t')$.
2. $\Rightarrow_{\mathrm{SU}}$ terminates on $\{s = t\}$.
3. if $\{s = t\} \Rightarrow^*_{\mathrm{SU}} E$ then σ is a unifier (mgu) of E iff σ is a unifier (mgu) of $\{s = t\}$.
4. if $\{s = t\} \Rightarrow^*_{\mathrm{SU}} \perp$ then s and t are not unifiable.
5. if $\{s = t\} \Rightarrow^*_{\mathrm{SU}} \{x_1 = t_1, \ldots, x_n = t_n\}$ and this is a normal form, then $\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$ is an mgu of s, t.

Size of Unification Problems

Any normal form of the unification problem E given by

$$\{f(x_1, g(x_1, x_1), x_3, \ldots, g(x_n, x_n)) = f(g(x_0, x_0), x_2, g(x_2, x_2), \ldots, x_{n+1})\}$$

with respect to $\Rightarrow_{\mathrm{SU}}$ is exponentially larger than E.

Polynomial Unification

The second calculus, polynomial unification, prevents the problem of exponential growth by introducing an implicit representation for the mgu.

For this calculus the size of a normal form is always polynomial in the size of the input unification problem.

Tautology $E \uplus \{t = t\} \;\Rightarrow_{\text{PU}}\; E$

Decomposition $E \uplus \{f(s_1, \ldots, s_n) = f(t_1, \ldots, t_n)\} \;\Rightarrow_{\text{PU}}$
$E \uplus \{s_1 = t_1, \ldots, s_n = t_n\}$

Clash $E \uplus \{f(t_1, \ldots, t_n) = g(s_1, \ldots, s_m)\} \;\Rightarrow_{\text{PU}}\; \bot$
if $f \neq g$

Occurs Check $E \uplus \{x = t\} \Rightarrow_{\mathrm{PU}} \bot$
if $x \neq t$ and $x \in \mathit{vars}(t)$

Orient $E \uplus \{t = x\} \Rightarrow_{\mathrm{PU}} E \uplus \{x = t\}$
if $t \notin \mathcal{X}$

Substitution $E \uplus \{x = y\} \Rightarrow_{\mathrm{PU}} E\{x \mapsto y\} \uplus \{x = y\}$
if $x \in \mathit{vars}(E)$ and $x \neq y$

Cycle $E \uplus \{x_1 = t_1, \ldots, x_n = t_n\} \ \Rightarrow_{\mathrm{PU}} \ \bot$

if there are positions p_i with $t_i|_{p_i} = x_{i+1}, t_n|_{p_n} = x_1$ and some $p_i \neq \epsilon$

Merge $E \uplus \{x = t, x = s\} \ \Rightarrow_{\mathrm{PU}} \ E \uplus \{x = t, t = s\}$

if $t, s \notin \mathcal{X}$ and $|t| \leq |s|$

3.7.4 Theorem (Soundness, Completeness and Termination of $\Rightarrow_{\text{PU}}$)

If s, t are two terms with $\text{sort}(s) = \text{sort}(t)$ then

1. if $\{s = t\} \Rightarrow^*_{\text{PU}} E$ then any equation $(s' = t') \in E$ is well-sorted, i.e., $\text{sort}(s') = \text{sort}(t')$.
2. $\Rightarrow_{\text{PU}}$ terminates on $\{s = t\}$.
3. if $\{s = t\} \Rightarrow^*_{\text{PU}} E$ then σ is a unifier (mgu) of E iff σ is a unifier (mgu) of $\{s = t\}$.
4. if $\{s = t\} \Rightarrow^*_{\text{PU}} \bot$ then s and t are not unifiable.

3.7.5 Theorem (Normal Forms Generated by $\Rightarrow_{\text{PU}}$)

Let $\{s = t\} \Rightarrow^{*}_{\text{PU}} \{x_1 = t_1, \ldots, x_n = t_n\}$ be a normal form. Then

1. $x_i \neq x_j$ for all $i \neq j$ and without loss of generality $x_i \notin \text{vars}(t_{i+k})$ for all i, k, $1 \leq i < n$, $i + k \leq n$.
2. the substitution $\{x_1 \mapsto t_1\}\{x_2 \mapsto t_2\} \ldots \{x_n \mapsto t_n\}$ is an mgu of $s = t$.