Substitutions

Recall that a substitution is a mapping $\theta$ from some set of variables to the set of terms (not necessarily ground). We represent substitutions by a tuple

$$\theta = [X_1 \mapsto t_1, \ldots, X_n \mapsto t_n]$$

which indicates that each variable $X_i$ is mapped to the term $t_i$ by the substitution $\theta$. 
Given a formula or term $A$, the result of applying the substitution to $A$, written $A\theta$ is obtained by simultaneously replacing each free occurrence in $A$ of a variable $X$ in the domain of $\theta$ by the term $\theta(X)$.

Examples:
1. $\text{father}(X)[X \mapsto \text{mother}(Y), Y \mapsto \text{joe}] = \text{father}(\text{mother}(Y))$
2. $\text{parent}(\text{father}(X), Y)[X \mapsto \text{mother}(Y), Y \mapsto \text{joe}] = \text{parent}(\text{father}(\text{mother}(Y)), \text{joe})$

The operation of applying a substitution can be repeated. Suppose $\theta_1$ and $\theta_2$ are two substitutions. We write $A\theta_1\theta_2$ for $(A\theta_1)\theta_2$, i.e., the result of first applying $\theta_1$ and then $\theta_2$.

Example:
If $\theta_1 = [X \mapsto f(X), Y \mapsto g(a, T), Z \mapsto b]$ and $\theta_2 = [X \mapsto a, Y \mapsto b, V \mapsto X]$ then
$$p(U, V, W, X, Y, Z)\theta_1\theta_2 = p(U, V, W, f(X), g(a, T), b)\theta_2$$
$$= p(U, X, W, f(a), g(a, T), b)$$
The composition of two substitutions

- $\theta_1 = [X_1 \mapsto t_1, \ldots, X_m \mapsto t_m]$
- $\theta_2 = [Y_1 \mapsto s_1, \ldots, Y_n \mapsto s_n]$

(where the variables $X_i$ are not necessarily distinct from the variables $Y_j$) is the substitution $\theta_1 \circ \theta_2$ obtained from the tuple

$[X_1 \mapsto t_1 \theta_2, \; X_2 \mapsto t_2 \theta_2, \ldots, X_m \mapsto t_m \theta_2, \; Y_1 \mapsto s_1, \; \ldots, Y_n \mapsto s_n]$

by deleting each pair “$Y_i \mapsto s_i$” for which $Y_i$ is one of the variables $X_1, \ldots, X_m$.

Example:

If $\theta_1 = [X \mapsto f(X), \; Y \mapsto g(a, T), \; Z \mapsto b]$ and $\theta_2 = [X \mapsto a, \; Y \mapsto b, \; V \mapsto X]$ then

$$\theta_1 \circ \theta_2 = [X \mapsto f(a), \; Y \mapsto g(a, T), \; Z \mapsto b, \; V \mapsto X]$$
**Proposition:** For any atom or formula $A$, we have

$$A(\theta_1 \circ \theta_2) = A\theta_1 \theta_2$$

i.e. the result of applying the composition $\theta_1 \circ \theta_2$ is the same as the result of applying first $\theta_1$ and then $\theta_2$ to $A$.

**Example:**

For the substitutions $\theta_1$ and $\theta_2$ in the example above we determined that $\theta_1 \circ \theta_2 = [X \mapsto f(a), \ Y \mapsto g(a, T), \ Z \mapsto b, \ V \mapsto X]$. Note that

$$p(U, V, W, X, Y, Z)[X \mapsto f(a), \ Y \mapsto g(a, T), \ Z \mapsto b, \ V \mapsto X]$$

$$= p(U, X, W, f(a), g(a, T), b)$$

which is the same as the result of $p(U, V, W, X, Y, Z)\theta_1 \theta_2$ obtained above.
Intuitively, the definition of the composition is designed to make the proposition true. The definition does the right thing for this because any occurrence of a variable $X_i$ in $A$ will be mapped first to $t_i$ by $\theta_1$. Then we apply $\theta_2$ to the result, producing the subterm $t_i\theta_2$. For occurrences of $Y_i$ in $A$ there are two cases:

1. if $Y_i$ is $X_j$ then $Y_i$ is replaced by $t_j$ when applying $\theta_1$: we then apply $\theta_2$ to the result, producing $t_j\theta_2$;
2. if $Y_i$ is not one of the variables $X_j$ then each occurrence of $Y_i$ in $A$ is unchanged by applying $\theta_1$. Then when we apply $\theta_2$, the occurrence of $Y_i$ will be replaced by $s_i$.

Remarks:

1. Substituting a variable for itself changes nothing, so
   
   \[
   [X \mapsto X, \ Y_1 \mapsto t_1, \ldots, Y_n \mapsto t_n] = [Y_1 \mapsto t_1, \ldots, Y_n \mapsto t_n]
   \]

2. The empty substitution $\epsilon = []$ changes nothing, so
   
   $\epsilon \circ \theta = \theta \circ \epsilon = \theta$

   for all substitutions $\theta$.

3. Composition of substitutions is associative:
   
   $\theta_1 \circ (\theta_2 \circ \theta_3) = (\theta_1 \circ \theta_2) \circ \theta_3$
We can now address the question of how to “match” literals when doing “resolution with variables”.

Suppose we want to match a ground instance of the clause
\[ p(X, f(X), a). \]
and a ground instance of the goal
\[ \bot : \neg p(g(Y), Z, T). \]
so that we can produce the resolvent \( \bot \).

Then we need to find values to substitute for the variables \( X, Y, Z \) that make these two atoms equal. We need to have
\[ X = g(Y), \quad f(X) = Z, \quad a = T. \]

One substitution that satisfies these constraints is
\[ \theta = [X \mapsto g(a), \ Y \mapsto a, \ Z \mapsto f(g(a)), \ T \mapsto a], \]
for
\[ p(X, f(X), a)\theta = p(g(a), f(g(a)), a) = p(g(Y), Z, T)\theta \]

But there is a “more general” solution,
\[ \theta = [X \mapsto g(U), \ Y \mapsto U, \ Z \mapsto f(g(U)), \ T \mapsto a], \]
for which
\[ p(X, f(X), a)\theta = p(g(U), f(g(U)), a) = p(g(Y), Z, T)\theta \]
(Here \( U \) is a variable, which possibly denotes \( a \).)
Sometimes there is no way to substitute values so as to match two atoms.

Examples:

1. There is no substitution \( \theta \) such that

\[
p(f(X), g(a))\theta = p(Z, Z)\theta
\]

for we would need to have \( Z\theta = f(X\theta) = g(a) \) whereas the latter two terms cannot be equal, whatever \( \theta \) is.

2. There is no substitution \( \theta \) such that

\[
p(f(X), X)\theta = p(Z, Z)\theta
\]

for we would need to have \( Z\theta = f(X\theta) = X\theta \) whereas the latter two terms cannot be equal, whatever \( \theta \) is, since \( f(X\theta) \) will have three more symbols ("f", "(" and ")") than \( X\theta \).

**Definition:** Let \( S = \{E_1, \ldots, E_n\} \) be a set of expressions (i.e., terms or atoms). A *unifier* for \( S \) is a substitution \( \theta \) such that

\[
E_1\theta = E_2\theta = \ldots = E_n\theta
\]

If a unifier exists, we say that \( S \) is *unifiable*.

Examples:

1. \( \{p(X), q(Y)\} \) is *not* unifiable
2. \( \{p(f(g(X), Y)), p(f(h(X), Z))\} \) is *not* unifiable
3. \( \{f(X), X\} \) is *not* unifiable
The following are all unifiers for $S = \{p(X), \ p(Y)\}$

1. $[X \mapsto f(a), \ Y \mapsto f(a)]$
2. $[X \mapsto Y]$
3. $[X \mapsto T, \ Y \mapsto T]$

**Definition:** A substitution $\theta_1$ is as at least as general as a substitution $\theta_2$ if there exists a substitution $\sigma$ such that $\theta_2 = \theta_1 \circ \sigma$.

**Definition:** A most general unifier (m.g.u.) of a set $S$ of expressions is a unifier $\theta$ of $S$ that is at least as general as any other unifier.
Example: Consider $S = \{ p(X), p(Y) \}$.

$\theta_1 = [X \mapsto Y]$ can be shown to be a most general unifier. Let’s consider some other unifiers...

1. $\theta_2 = [X \mapsto a, \ Y \mapsto a]$ is a unifier and if we take $\sigma = [Y \mapsto a]$ then

   
   $\theta_1 \circ \sigma = [X \mapsto a, \ Y \mapsto a] = \theta_2$

   
   so $\theta_1$ is at least as general as $\theta_2$.

2. $\theta_2 = [Y \mapsto X]$ is a unifier and if we take $\sigma = [Y \mapsto X]$ then

   
   $\theta_1 \circ \sigma = [X \mapsto X, \ Y \mapsto X] = [Y \mapsto X] = \theta_2$

   
   so $\theta_1$ is at least as general as $\theta_2$.

Computing a most general unifier:

Definition: The disagreement set of a set $S$ of expressions is obtained by locating the leftmost position at which not all expressions in $S$ have the same symbol, and placing in $D$ all the subexpressions starting at that position.

Example: if

\[
S = \{ q(f(X), g(X), a), \\
q(f(X), g(X), b), \\
q(f(X), g(h(T)), a) \}
\]

then $D = \{ X, h(T) \}$
Unification algorithm

1. Put \( k = 0 \) and \( \theta_0 = \epsilon \) (the empty substitution)

2. If \( S\theta_k \) is a singleton set
   then halt: \( \theta_k \) is the m.g.u. of \( S \).

3. Compute the disagreement set \( D_k \) of \( S\theta_k \)

4. If there exist terms \( v \) and \( t \) in \( D_k \) such that
   (a) \( v \) is a variable
   (b) \( v \) does not occur in \( t \) ("occurs check")
   then
     (a) put \( \theta_{k+1} := \theta_k \circ [v \mapsto t] \)
     (b) \( k := k + 1 \)
     (c) go to to step 2
   else halt: \( S \) is not unifiable