Proof. If N is unsatisfiable, saturation via the tableau calculus generates a closed tableau. So there is an i such that $N \Rightarrow^i_{\text{TAB}} N'$ and N' is closed. Every closed branch is the result of finitely many tableau rule applications on finitely many clauses $\{C_1, \ldots, C_n\} \subseteq N$. Let M be the union of all these finite clause sets, so $M \subseteq N$. Tableau is sound, so M is a finite, unsatisfiable subset of N. $\square$

3.7 Unification

Definition 3.7.1 (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'.

Obviously, two terms of different sort cannot be made equal by well-sorted instantiation. Since well-sortedness is preserved by all rules of the unification calculus, we assume from now an that all equations, terms, and substitutions are well-sorted.

The first calculus is the naive standard unification calculus that is typically found in the (old) literature on automated reasoning [21]. 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, $\operatorname{sort}(s) = \operatorname{sort}(t)$, (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 $E = \{x = t\} \uplus E'$, $x \notin \operatorname{vars}(t)$ and $x \notin \operatorname{vars}(E)$.

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

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

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

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

Substitution $\quad E \uplus \{x = t\} \;\Rightarrow_{\text{SU}}\; E\{x \mapsto t\} \cup \{x = t\}$
if $x \in vars(E)$ and $x \notin vars(t)$

Occurs Check $\quad E \uplus \{x = t\} \;\Rightarrow_{\text{SU}}\; \bot$

if $x \neq t$ and $x \in vars(t)$

Orient $\quad E \uplus \{t = x\} \;\Rightarrow_{\text{SU}}\; E \cup \{x = t\}$
if $t \notin \mathcal{X}$

Theorem 3.7.2 (Soundness, Completeness and Termination of $\Rightarrow_{\text{SU}}$)**.** If s, t are two terms with $\text{sort}(s) = \text{sort}(t)$ then

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

Proof. 1. by induction on the length of the derivation and a case analysis for the different rules.
2. for a state $E = \{s_1 = t_1, \ldots, s_n = t_n\}$ take the measure $\mu(E) := (n, M, k)$ where n is the number of unsolved variables, M the multiset of all term depths of the s_i, t_i and k the number of equations $t = x$ in E where t is not a variable. The state $\bot$ is mapped to $(0, \emptyset, 0)$. Then the lexicographic combination of $>$ on the naturals and its multiset extension shows that any rule application decrements the measure.
3. by induction on the length of the derivation and a case analysis for the different rules. Clearly, for any state where Clash, or Occurs Check generate $\bot$ the respective equation is not unifiable.
4. a direct consequence of 3.
5. if $E = \{x_1 = t_1, \ldots, x_n = t_n\}$ is a normal form, then for all $x_i = t_i$ we have $x_i \notin \text{vars}(t_i)$ and $x_i \notin \text{vars}(E \setminus \{x_i = t_i\})$, so $\{x_1 = t_1, \ldots, x_n = t_n\}\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\} = \{t_1 = t_1, \ldots, t_n = t_n\}$ and hence $\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$ is an mgu of $\{x_1 = t_1, \ldots, x_n = t_n\}$. By 3. it is also an mgu of s, t. ☐

Example 3.7.3 (Size of Standard 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 $\Rightarrow_{\text{SU}}$ is exponentially larger than E.

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

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

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

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

Cycle $E \uplus \{x_1 = t_1, \ldots, x_n = t_n\} \;\Rightarrow_{\text{PU}}\; \bot$
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 $E \uplus \{x = t, x = s\} \;\Rightarrow_{\text{PU}}\; E \uplus \{x = t, t = s\}$
if $t, s \notin \mathcal{X}$ and $|t| \leq |s|$

Theorem 3.7.4 (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.

Theorem 3.7.5 (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 $s = t$.

Proof. 1. If $x_i = x_j$ for some $i \neq j$ then Merge is applicable. If $x_i \in \text{vars}(t_i)$ for some i then Occurs Check is applicable. If the x_i cannot be ordered in the described way, then either Substitution or Cycle is applicable.
2. Since $x_i \notin \text{vars}(t_{i+k})$ the composition yields the mgu. □

Lemma 3.7.6 (Size of Unifiers)**.** Let $\{s = t\}$ be a unification problem between two non-variable terms. Then

1. if s and t are linear then for any unifier σ and any term $r \in \text{codom}(\sigma)$, $|r| < |s|$ and $|r| < |t|$ as well as $\text{depth}(r) < \text{depth}(s)$ and $\text{depth}(r) < \text{depth}(t)$,
2. if s is shallow and linear, then the mgu σ of s and t is also a matcher from s to t, i.e., $s\sigma = t$

Proof. Both parts follow directly from the structure of the terms s, t: if they are both linear then the substitution rule is never applied. If s is shallow and linear, it has the form $f(x_1, \ldots, x_n)$, all x_i different, then the unifier is $\sigma = \{x_i \mapsto t|_i \mid 1 \leq i \leq n\}$. □

3.8 First-Order Free-Variable Tableau

An important disadvantage of standard first-order tableau is that the γ ground term instances need to be guessed. The main complexity in proving a formula to be valid lies in this guessing as for otherwise tableau terminates with a proof. Guessing useless ground terms may result in infinite branches. A natural idea is to guess ground terms that can eventually be used to close a branch. Of course, it is not known which ground term will close a branch. Therefore, it would be great to postpone the γ instantiations. This is the idea of free-variable first-order tableau. Instead of guessing a ground term for a γ formula, free-variable tableau introduces a fresh variable. Then a branch can be closed if two complementary literals have a common ground instance, i.e., their atoms are unifiable. The instantiation is delayed until a branch is closed for two literals via unification. As a consequence, for δ formulas no longer constants are introduced but shallow, so called *Skolem* terms in the formerly universally quantified variables that had the δ formula in their scope.

The new calculus needs to keep track of scopes of variables, so I move from a state as a set of pairs of a sequence and a set of constants, see standard first-order tableau Section 3.6, to a set of sequences of pairs (M_i, X_i) where X_i is a set of variables.

Definition 3.8.1 (Direct Free-Variable Tableau Descendant)**.** Given a γ- or δ-formula ϕ, Figure 3.2 shows its direct descendants.

The notion of closedness, Section 3.6, transfers exactly from standard to free-variable tableau. For α- and β-formulas the definition of an *open* formula remains unchanged as well. A γ- or δ-formula is called *open* in (M, X) if no direct descendant is contained in M. Note that instantiation of a tableau may remove direct descendants of γ- or δ-formulas by substituting terms for variables. Then a branch, pair (M, X), sequence M, is *open* if it is not closed and there is an open formula in M or there is pair of unifiable, complementary literals in M.