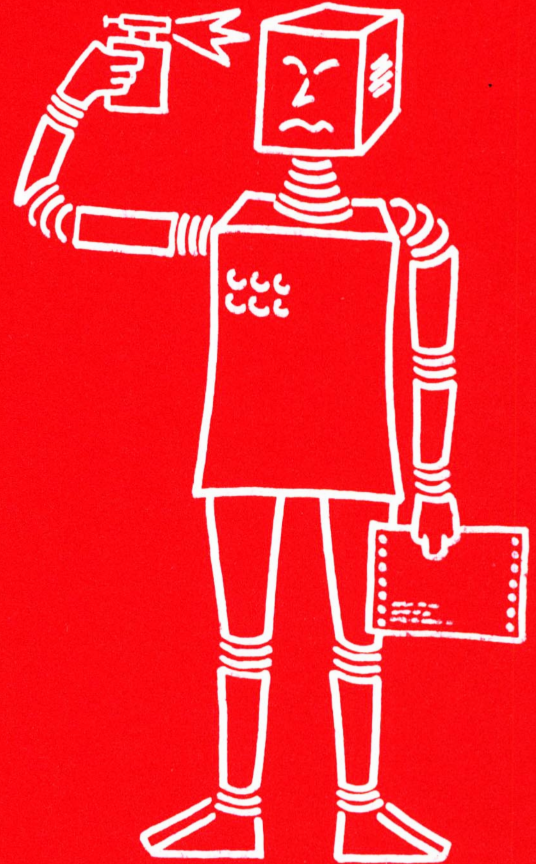
SEKI-PROJEKT

SEKI MEMO

Institut für Informatik III
Universität Bonn
Bertha-von-Suttner-Platz 6
D 5300 Bonn 1, W. Germany

Institut für Informatik I
Universität Karlsruhe
Postfach 6380
D-7500 Karlsruhe 1, W. Germany



Some Basic Notions of First-Order Unification Theory

A. Herold
Institut für Informatik I
Universität Karlsruhe
Postfach 6380
D-7500 Karlsruhe 1

UNIVERSITÄT KARLSRUHE
FAKULTÄT FÜR INFORMATIK
Postfach 63 80, D 7500 Karlsruhe 1

Interner Bericht

Nr. 15/83

SOME BASIC NOTIONS OF FIRST-ORDER UNIFICATION THEORY

A. Herold
Institut für Informatik I
Universität Karlsruhe
Postfach 6380
D-7500 Karlsruhe 1

Abstract

This report does not contain much novel material, but collects the basic notions and the most frequently used lemmata and theorems of first order unification theory. It is restricted to the case of free terms (i.e. no defining equations).

Contents

0. *Introduction*

This report was written during the redesign of the unification module of the Markgraf Karl Refutation Procedure [BES81], [Oh82] an Automatic Theorem Prover developed at the University of Karlsruhe. It collects and summarizes the basic theoretical notions underlying first-order unification theory investigated from an algebraic point of view. So we will not give any algorithms to compute unifiers or matchers. The definitions and lemmata are illustrated by examples and counterexamples.

Substitutions are usually defined as endomorphisms on the free termalgebra in the literature ([CL73], [Hu76], [Lo78] and [Ro65]) and hence substitutions like $\sigma = \{x \leftarrow f(x)\}$ or $\tau = \{x \leftarrow f(y), y \leftarrow g(z), z \leftarrow h(x)\}$ are legal according to this definition.
But actual unification algorithms should not produce such substitutions and in fact all known algorithms don't ([Ro65], [Ro71], [KB70], [Ba73],[Hu76], [MM79], [PW78]).

Example I: Let $s = x$ and $t = f(y)$, then $\sigma = \{x \leftarrow f(x), y \leftarrow x\}$ is a most general unifier of s and t, but a unification algorithm should produce $\tau = \{x \leftarrow f(y)\}$. Note that $\tau = \{x \leftarrow y\} \circ \sigma$ and $\sigma = \{y \leftarrow x\} \circ \tau$

Example II: Let $s = x$ and $t = y$, then $\sigma = \{x \leftarrow y, z \leftarrow x\}$ is a most general unifier of s and t and so are $\tau_1 = \{x \leftarrow y\}$ and $\tau_2 = \{y \leftarrow x\}$. Note that $\sigma = \{x \leftarrow z\} \circ \tau_1$ and $\tau_1 = \{z \leftarrow x\} \circ \sigma$.
For these and other reasons we require an additional property for substitutions: the idempotence, i.e. $\sigma \circ \sigma = \sigma$. Let Σ denote the set of all substitutions and let $\mathcal{S}$ be the set of idempotent substitutions. The restriction to idempotent substitutions has deep consequences; for example $(\Sigma, \circ)$ is a semigroup, but $(\mathcal{S}, \circ)$ is not, since the composition of two idempotent substitutions is i.g. not idempotent. The bijective endomorphisms are not idempotent [Hu76].

In Chapter I we introduce terms, substitutions and renaming substitutions. We state conditions under which the composition of two idempotent substitutions is again idempotent and define a union of substitutions. In Chapter II we define some relations on terms and substitutions and show that the matching problem for substitutions can be solved by a corresponding matching problem for terms.

The existence and some basic properties of idempotent most general unifiers are shown in Chapter III. The last chapter deals with unification of substitutions, where it is shown that the definition of this report is equivalent to the definitions given in [Ch72], [CL73], [CS79] and [Ni80]. Our definition is similar to [Va75]. In addition an example is used to demonstrate that the definition in [Si76] is weaker than ours.

Throughout this paper we use the following standard mathematical notations:

id	the identity function
$f\|_V$	function f restricted to a subset M of its domain
$f(t)\downarrow$	t is in the domain of f
$f(t)\uparrow$	t is not in the domain of f
$\circ$	composition of functions
|	negation, e.g. $x \not\leq y$ means not $x \leq y$
$M \setminus N$	set theoretic difference of M and N
$M \subset N$	M is a subset of N or M is equal N

I. Terms and Substitutions

1. First-Order Terms

Let $\mathfrak{C}$ be the set of *constants* i.e. nullary functionsymbols, $\mathcal{F}$ the set of *functionsymbols*, $\mathfrak{V}$ the set of *variables*, $\Omega = \mathfrak{C} \cup \mathcal{F} \cup \mathfrak{V}$ and $\alpha : \Omega \to \mathbb{N}$ an arity-function with $\alpha(x) = 0$ if $x \in \mathfrak{V} \cup \mathfrak{C}$.

We write $\mathfrak{T}(\Omega)$ for the *free termalgebra* over $\mathfrak{C} \cup \mathcal{F}$ with respect to $\mathfrak{V}$(shortly $\mathfrak{T}$) which is given a concrete representation by (i) $\mathfrak{C}, \mathfrak{V} \subset \mathfrak{T}$, (ii) if $f \in \mathcal{F}$, $\alpha(f) = n$ and $t_1, \ldots, t_n \in \mathfrak{T}$ then $f(t_1 \ldots t_n) \in \mathfrak{T}$. This definition permits us to use structural induction.

Let *var* be a mapping which yields the variables of a set of terms:

$$\text{var}: \begin{cases} \mathfrak{P}(\mathfrak{T}) \longrightarrow \mathfrak{P}(\mathfrak{V}) \\ \{t\} \longrightarrow \begin{cases} \phi & \text{if } t = c \in \mathfrak{C} \\ \{x\} & \text{if } t = x \in \mathfrak{V} \\ \bigcup\limits_{i=1}^{n} \text{var}(\{t_i\}) & \text{if } t = f(t_1 \ldots t_n) \end{cases} \\ A \longrightarrow \bigcup\limits_{t \in A} \text{var}(\{t\}) \quad \text{if } A \subset \mathfrak{T} \end{cases}$$

For ease of notation we will write $\text{var}(t_1, \ldots, t_n)$ instead of $\text{var}(\{t_1, \ldots, t_n\})$.

In order to formalize the selection of a subterm in a term we define *subterm selectors* [Wa82]. Let k_α be a natural number, then we call the partial function α

$$\alpha : \begin{cases} \mathfrak{T} \longrightarrow \mathfrak{T} \\ t \longrightarrow \begin{cases} t_{k_\alpha} & \text{if } t = f(t_1 \ldots t_n) \text{ and } n \geq k_\alpha \\ \text{undefined} & \text{else} \end{cases} \end{cases}$$

an *argument selector*. SEL denotes the set of all argument selectors. The identity functions on terms, an argument selector or a finite composition of argument selectors are called subterm selectors or *selectors* for short. Let SEL* denote the set of all selectors. If $t \in \mathfrak{T}$ is in the domain of a selector α we write $\alpha(t)\downarrow$ and $\alpha(t)\uparrow$ if it is not. Selectors are sometimes given a concrete representation called occurences or positions in the literature [Hu80], [PS81], [Ros73], [SS81].

The following lemma which shows that selectors and substitutions (confer next section) commute is easy to prove:

Lemma I.1: Let $t \in \mathfrak{T}$, $\alpha, \beta \in SEL^*$ and $\sigma \in \Sigma$.

(i) If $\alpha(t)\downarrow$, then $\sigma\, \alpha(t) = \alpha(\sigma t)$.

(ii) If $\alpha(t)\downarrow$, $\beta(t)\downarrow$ and $\alpha(t) = \beta(t)$ then $\alpha(\sigma t) = \beta(\sigma t)$.

2. *Substitutions*

A substitution σ is an endomorphism (on the free termalgebra $\mathfrak{T}$), which is the identity on $\mathfrak{C}$, i.e.

$$\sigma: \mathfrak{T} \longrightarrow \mathfrak{T} \quad \text{with}$$

$$\text{(I.1)} \quad \sigma|_{\mathfrak{C}} = id_{\mathfrak{C}}$$

$$\text{(I.2)} \quad \sigma(f(t_1 \ldots t_n)) = f(\sigma(t_1) \ldots \sigma(t_n))$$

The set of all substitutions is Σ and ε is the identity function on Σ called the empty substitution.

In this paper we are interested in a subset of Σ, the set of idempotent substitutions $\mathfrak{S} \subset \Sigma$ with the additional property

$$\text{(I.3)} \quad \sigma \circ \sigma = \sigma$$

where "$\circ$" denotes functional composition. A substitution with $\sigma(x)=f(x)$ for example is not idempotent since

$$\sigma \circ \sigma(x) = f(f(x)) \neq f(x) = \sigma(x) \;.$$

If not explicitly stated otherwise, we mean by substitutions always idempotent substitutions.

The application of a substitution σ to a term t is denoted by σt and to a set W of terms by $\sigma(W) = \{\sigma t \mid t \in W\}$. The following lemma is frequently used throughout this paper:

Lemma I.2: Every substitution $\sigma \in \Sigma$ is uniquely determined by its restriction on $\mathfrak{V}$.

Proof: We use structural induction. Let $t \in \mathfrak{C}$ then $\sigma t = t$ by (I.1). Let $t \in \mathfrak{V}$ then σt is defined by the given restriction. Let $t = f(t_1 \ldots t_n)$ and σt_i is given by the induction hypothesis. Then we have $\sigma t = \sigma f(t_1 \ldots t_n) = f(\sigma t_1 \ldots \sigma t_n)$ by (I.2). □

An immediate consequence is the following

Corollary: Let $\sigma,\tau \in \Sigma$. If $\sigma|_{\mathfrak{V}} = \tau|_{\mathfrak{V}}$ then $\sigma = \tau$.

In order to show that two substitutions are equal we shall often use this corollary by showing that they are equal on the set of variables.

For a substitution $\sigma \in \Sigma$ we define the *domain of* σ as

$$DOM(\sigma) = \{x \in \mathfrak{V} \mid \sigma x \neq x\}$$

the *codomain of* σ as

$$COD(\sigma) = \sigma(DOM(\sigma))$$

and the set of *variables introduced by* σ as

$$VCOD(\sigma) = var(COD(\sigma))$$

and the set of *variables of* σ as

$$var(\sigma) = DOM(\sigma) \cup VCOD(\sigma) \quad .$$

If $DOM(\sigma) = \{x_1,...,x_n\}$ is finite, σ can be represented as

$$\sigma = \{x_1 \leftarrow t_1,...,x_n \leftarrow t_n\}$$

with the following meaning $\sigma x_i = t_i$ for $i=1,...,n$ and $\sigma x = x$ else. The *subset* preorder for substitutions $\sigma,\tau \in \mathfrak{S}$ is defined as $\sigma \subset \tau$ iff $DOM(\sigma) \subset DOM(\sigma)$ and $\sigma x = \tau x$ for all $x \in DOM(\sigma)$.

Lemma I.3: For $\sigma \in \mathfrak{S}$: $DOM(\sigma) \cap VCOD(\sigma) = \emptyset$

Proof: Suppose $x \in DOM(\sigma)$ and $x \in VCOD(\sigma)$. Then there exists $y \in DOM(\sigma)$ and $x \in var(\{\sigma y\})$. Let $t = \sigma y$ and $\alpha \in SEL^*$ with $\alpha(t)\downarrow$ and $\alpha(t) = x$. Then $\alpha(\sigma y) = \alpha(t) = x \neq \sigma x = \alpha(\sigma t) = \alpha(\sigma \circ \sigma y)$, hence $\sigma \neq \sigma \circ \sigma$ which is a contradiction. □

Lemma I.4: Let $\sigma \in \mathfrak{S}$. Then for all $t \in \mathfrak{T}$:

$$DOM(\sigma) \cap var(\{\sigma t\}) = \emptyset$$

Proof: (by structural induction on $t \in T$).

Base case: If $t = c \in \mathfrak{C}$ then $var(\{\sigma t\}) = \emptyset$. If $t = x \in \mathfrak{V}$ and $x \in DOM(\sigma)$, then by Lemma I.3 $x \notin VCOD(\sigma)$ and therefore $DOM(\sigma) \cap var(\{\sigma t\}) = \emptyset$. If $x \notin DOM(\sigma)$ then $\sigma x = x$ and hence $DOM(\sigma) \cap var(\{\sigma t\}) = \emptyset$.

Induction Step: Let $t = f(t_1 \ldots t_n)$ and $DOM(\sigma) \cap var(\{\sigma t_i\}) = \emptyset$ for $i=1,\ldots,n$ by the induction hypothesis. Then it is

$$\begin{aligned} DOM(\sigma) \cap var(\{\sigma t\}) &= DOM(\sigma) \cap var(\{\sigma f(t_1 \ldots t_n)\}) \\ &= DOM(\sigma) \cap var(\{f(\sigma t_1 \ldots \sigma t_n)\}) \\ &= DOM(\sigma) \cap \bigcup_{i=1}^{n} var(\{\sigma t_i\}) \\ &= \bigcup_{i=1}^{n} DOM(\sigma) \cap var(\{\sigma t_i\}) \\ &= \emptyset. \end{aligned}$$

□

The next lemma can be used as a different characterization of $\mathfrak{S}$:

<u>*Lemma I.5:*</u> Let $\sigma \in \Sigma$, then

$$\sigma \in \mathfrak{S} \quad \text{iff} \quad DOM(\sigma) \cap VCOD(\sigma) = \emptyset \quad .$$

Proof: (i) The forward direction is shown in Lemma I.3.

(ii) Let $DOM(\sigma) \cap VCOD(\sigma) = \emptyset$. Now for every $x \in \mathfrak{V}$ $var(\sigma x) \subset VCOD(\sigma)$ and hence $var(\sigma x) \cap DOM(\sigma) = \emptyset$. But then $\sigma\sigma x = \sigma x$ and hence $\sigma \in \mathfrak{S}$, since σ is defined by its restriction on $\mathfrak{V}$.

The following lemma is concerned with sets of terms and substitutions

<u>*Lemma I.6:*</u> Let C, D be sets of terms and $\sigma \in \mathfrak{S}$:

(i) If $C \subset D$ then $\sigma(C) \subset \sigma(D)$;

(ii) $\sigma(C) \cup \sigma(D) = \sigma(C \cup D)$;

(iii) $\sigma(C) \cap \sigma(D) = \sigma(C \cap D)$;

(iv) $\sigma(C) \setminus \sigma(D) \subset \sigma(C \setminus D)$;

(v) If $\sigma(C) \setminus \sigma(D) \neq \sigma(C \setminus D)$, then there exist $t \in C$, $s \in D$ with $t \neq s$ and $\sigma t = \sigma s$.

Proof: (i) Let $s \in \sigma(C)$ then $s = \sigma t$ with $t \in C$. Since $t \in D$, $s = \sigma t \in \sigma(D)$.

(ii), (iii) and (iv) are easily shown.

(v) We have $\sigma(C) \setminus \sigma(D) \subsetneqq \sigma(C \setminus D)$ by (iv) and the assumption, i.e. $\sigma(C \setminus D) \setminus (\sigma(C) \setminus \sigma(D)) \neq \emptyset$.

Hence there exists $r \in \sigma(C \setminus D)$ and $r \notin \sigma(C) \setminus \sigma(D)$ and $t \in C \setminus D$ with $r = \sigma t$. Since $r \in \sigma(C)$ and $r \notin \sigma(C) \setminus \sigma(D)$ we have $r \in \sigma(D)$, i.e. there exists $s \in D$ with $r = \sigma s$ and $s \neq t$ by $t \notin D$.

□

The composition of substitutions is the usual functional composition, but unfortunately the composition of two idempotent substitutions is in general not idempotent.

Lemma I.7: The composition $\sigma \circ \tau$ of two substitutions $\sigma, \tau \in \mathcal{S}$ satisfies condition (I.1) and (I.2) but in general not condition (I.3).

Proof: Let $\sigma, \tau \in \mathcal{S}$. Then with $c \in \mathfrak{C}$ we have $\sigma \circ \tau\ c = \sigma(\tau c) = \sigma c = c$, therefore (I.1); and for $t = f(t_1 \ldots t_n)$

$$
\begin{aligned}
(\sigma \circ \tau)t &= \sigma(\tau t) = \sigma(\tau\ f(t_1 \ldots t_n)) \\
&= \sigma\ f(\tau t_1 \ldots \tau t_n) \\
&= f(\sigma(\tau t_1) \ldots \sigma(\tau t_n)) \\
&= f((\sigma \circ \tau)t_1 \ldots (\sigma \circ \tau)t_n), \text{ therefore (I.2).}
\end{aligned}
$$

A counterexample for condition (I.3): let $\sigma = \{x | f(y\ z)\}, \tau = \{y | b\}$. Then $\sigma \circ \tau x = f(y\ z)$ and $(\sigma \circ \tau) \circ (\sigma \circ \tau) x = f(b\ z) \neq f(y\ z)$.

□

Lemma I.7 shows, that $(\mathcal{S}, \circ)$ is not a semigroup (since it is not closed under $\circ$), whereas $(\Sigma, \circ)$ is a semigroup. For ease of notation we will often omit the "$\circ$"-symbol, i.e. we write $\sigma\tau$ for $\sigma \circ \tau$.

So we are looking for conditions such that the composition of two idempotent substitutions is again idempotent. To this end we define for $y \in \mathfrak{V}$ the set $DOM(\sigma, y) = \{x \in DOM(\sigma) \mid y \in var(\sigma x)\} \subset DOM(\sigma)\}$ of all variables $x \in DOM(\sigma)$ such that y is a variable in σx.

Lemma I.8: Let $\sigma,\tau \in \mathfrak{S}$. Then $\sigma\tau \in \mathfrak{S}$ iff for all $y \in VCOD(\sigma) \cap DOM(\tau)$.

(i) $\tau y \notin \mathfrak{V}$ implies $DOM(\sigma,y) \subset DOM(\tau)$ and

(ii) $\tau y \in \mathfrak{V}$ (i.e. $\tau y = z$) implies $\sigma z = y$ or $DOM(\sigma,y) \subset DOM(\tau)$.

Proof: (A) First we show by contradiction that $\sigma\tau \in \mathfrak{S}$ implies (i) and (ii). Suppose there exists $y \in VCOD(\sigma) \cap DOM(\tau)$ such that $\tau y \notin \mathfrak{V}$ and $DOM(\sigma,y) \not\subset DOM(\tau)$ or $\tau y \in \mathfrak{V}$ (i.e. $\tau y = z$) and $\sigma z \neq y$ and $DOM(\sigma,y) \not\subset DOM(\tau)$.

Let $\tau y \notin \mathfrak{V}$ and $x \in DOM(\sigma,y)$ and $x \notin DOM(\tau)$, then $\sigma\tau x = \sigma x = t$ with $\sigma x = t \neq x$ and $y \in var(t)$, i.e. $y \in VCOD(\sigma\tau)$, and $\sigma\tau y \notin \mathfrak{V}$ since $\tau y \notin \mathfrak{V}$, i.e. $y \in DOM(\sigma,\tau)$. Summarizing we get $y \in DOM(\sigma\tau) \cap VCOD(\sigma\tau)$ which is a contradiction to $\sigma\tau \in \mathfrak{S}$ by Lemma I.5.

Now let $\tau y = z \in \mathfrak{V}$, $\sigma z \neq y$ and $x \in DOM(\sigma,y)$ and $x \notin DOM(\tau)$. By the same argument as above $y \in VCOD(\sigma\tau)$ and since $\sigma\tau y = \sigma z \neq y$ $y \in DOM(\sigma\tau)$ which again is a contradiction to $\sigma\tau \in \mathfrak{S}$.

(B) The other direction is shown by structural induction on τx for an arbitrary $x \in \mathfrak{V}$.

Base case: I. $\tau x = c \in \mathfrak{C}$: $\sigma\tau\sigma\tau x = \sigma\tau\sigma c = \sigma\tau c = \sigma\tau\tau x = \sigma\tau x$, since $\tau \in \mathfrak{S}$.

II. $\tau x = u \in \mathfrak{V}$:
If $u \notin DOM(\sigma)$ then $\sigma\tau\sigma\tau x = \sigma\tau\sigma u = \sigma\tau u = \sigma\tau\tau x = \sigma\tau x$.
Now let $u \in DOM(\sigma)$.

Case 1: $DOM(\tau) \cap var(\sigma u) = \phi$, i.e. $\tau\sigma u = \sigma u$ and hence $\sigma\tau\sigma\tau x = \sigma\tau\sigma u = \sigma\sigma u = \sigma u = \sigma\tau x$.

Case 2: $DOM(\tau) \cap var(\sigma u) = W \neq \phi$. We again have to distinguish three cases:

Case 2.1: There exists $y \in W$ with $\tau y \notin \mathfrak{V}$. By condition (i) it is $DOM(\sigma,y) \subset DOM(\tau)$ and hence $u \in DOM(\tau)$, since $y \in var(\sigma u)$. Now if $u = x$, then $\tau x = x$ contradicts $x \in DOM(\tau)$; and if $u \neq x$, then $u \in VCOD(\tau)$ which by Lemma I.5 is a contradiction to $\tau \in \mathfrak{S}$.

Case 2.2: For all $y \in W$ $\tau y \in \mathfrak{V}$ and there exists $v \in W$ such that $\tau v = z$ and $\sigma z \neq v$. Hence by (ii) it is $DOM(\sigma,v) \subset DOM(\tau)$ and as above it is $u \in DOM(\tau)$, which leads to the same contradiction as in case 2.1.

Case 2.3: For all $y \in W$ $\tau y \in V$ and $\tau y = z$ and $\sigma z = y$. We show that $\sigma\tau\sigma u = \sigma u$ by for all $w \in var(\sigma u)$ $\sigma\tau w = w$. That this is a sufficient argument can be easily proved by induction.

Now let $w \in var(\sigma u)$. If $w \notin DOM(\tau)$ then $\sigma\tau w = \sigma w = w$ since $DOM(\sigma) \cap VCOD(\sigma) = \emptyset$ and $w \in VCOD(\sigma)$. If $w \in DOM(\tau)$ then $w \in W$ and hence $\sigma\tau w = w$, since by assumption $\tau w = z$ and $\sigma z = w$. Summarizing we get $\sigma\tau\sigma\tau x = \sigma\tau\sigma u = \sigma u = \sigma\tau x$, which was to be shown.

Induction step: $\tau x = f(t_1 \ldots t_n)$:

$$
\begin{aligned}
\sigma\tau\sigma\tau x &= \sigma\tau\sigma\tau\tau x && \text{(by } \tau \in \mathfrak{S}\text{)}\\
&= \sigma\tau\sigma\tau\ f(t_1 \ldots t_n)\\
&= f(\sigma\tau\sigma\tau t_1 \ldots \sigma\tau\sigma\tau t_n)\\
&= f(\sigma\tau t_1 \ldots \sigma\tau t_n) && \text{(by induction hypothesis)}\\
&= \sigma\tau\ f(t_1 \ldots t_n)\\
&= \sigma\tau\ \tau x && \text{(by } \tau \in \mathfrak{S}\text{)}\\
&= \sigma\tau\ x
\end{aligned}
$$

□

Since the condition of Lemma I.8 is very technical, we shall often use a sufficient condition for $\sigma\tau \in \mathfrak{S}$ which is easier to check.

Corollary: Let $\sigma,\tau \in \mathfrak{S}$. If

(I.4) $DOM(\tau) \cap VCOD(\sigma) = \emptyset$

then $\sigma\tau \in \mathfrak{S}$.

The following two technical lemmata will often be used in the sequel.

Lemma I.9: Let $\sigma,\tau \in \mathfrak{S}$ and $\sigma \circ \tau \in \mathfrak{S}$ then

(i) $DOM(\sigma \circ \tau) \subset DOM(\sigma) \cup DOM(\tau)$

(ii) $DOM(\sigma) \subset DOM(\sigma \circ \tau)$

Proof: (i) Let $x \in DOM(\sigma \circ \tau)$, i.e. $\sigma \circ \tau x \neq x$. If $x \in DOM(\tau)$ we are finished and if $x \notin DOM(\tau)$ then $\sigma\tau x = \sigma x \neq x$, i.e. $x \in DOM(\sigma)$.

(ii) Let $x \in DOM(\sigma)$, i.e. $\sigma x \neq x$. If $\tau x = x$ then clearly $x \in DOM(\sigma \circ \tau)$. If $\tau x \neq x$ suppose $\sigma\tau x = x$. But this implies $x \in VCOD(\sigma)$ which is a contradiction to $\sigma \in \mathfrak{S}$.

□

Lemma I.10: Let $\sigma, \tau, \theta \in \Sigma$. If $\sigma = \tau$ then $\sigma\theta = \tau\theta$ and $\theta\sigma = \theta\tau$.

Proof: For all $x \in \mathfrak{V}$ $(\theta \circ \sigma)x = \theta(\sigma x) = \theta(\tau x) = (\theta \circ \tau)x$. Let $\theta x = t$, then $\sigma t = \tau t$ and therefore $(\sigma\theta)x = (\tau\theta)x$.

□

3. *Renaming Substitutions*

A substitution $\rho \in \mathcal{S}$ is called a *renaming substitution* with respect to $\emptyset \neq V \subset \mathfrak{V}$ iff the following conditions are satisfied.

(I.5) $\rho(\mathfrak{V}) \subset \mathfrak{V}$ (equivalently $COD(\rho) \subset \mathfrak{V}$)

(I.6) ρ is injective on V,
i.e. for all $x,y \in V$ $x \neq y$ implies $\rho x \neq \rho y$

(I.7) $DOM(\rho) = V$

We write REN(V) for the set of all renaming substitutions with respect to V. For example $\rho = \{x_1 \leftarrow y_1,\ x_2 \leftarrow y_2,\ x_3 \leftarrow y_3\}$ is in $REN(\{x_1,x_2,x_3\})$ whereas $\tilde{\rho} = \{x_1 \leftarrow y_1,\ x_2 \leftarrow y_1,\ x_3 \leftarrow y_3\}$ is not.

Let $\rho \in REN(V)$. We define the *converse* ρ^c of the renaming substitution ρ by

$$\rho^c\, x = y \quad \text{iff} \quad \rho y = x.$$

The next lemma shows that ρ^c is a renaming substitution.

<u>*Lemma 1.11:*</u> If $\rho \in REN(V)$ then $\rho^c \in REN(\rho(V))$.

Proof: Condition (I.5) and (I.7) are immediate comsequences of the definition. It remains to be shown that ρ is injective i.e. for all $x,y \in \rho(V)$ $\rho^c\, x = \rho^c\, y$ implies $x = y$. Let $x,y \in \rho(V)$, i.e. $x = \rho u$ and $y = \rho v$ with $u,v \in V$. Therefore we have $\rho^c\, x = \rho^c \rho u = u$, $\rho^c y = \rho^c \rho v = v$ and $u = v$, hence $x = \rho u = \rho v = y$.

□

<u>*Lemma 1.12:*</u> Let $\rho \in REN(V)$. Then

$$\rho\rho^c = \rho \quad \text{and} \quad \rho^c\rho = \rho^c.$$

Proof (by cases):

Case 1: $x \notin var(\rho)$, i.e. $\rho x = x$ and $\rho^c x = x$ and therefore $\rho\rho^c x = x$ and $\rho^c \rho x = x$.

Case 2: $x \in COD(\rho)$ then $x \notin V$ by the idempotence of ρ and therefore $\rho x = x$ and $\rho^c x = y$ with $\rho y = x$. Now $\rho \circ \rho^c x = \rho y = x = \rho x$ and $\rho^c \circ \rho x = \rho^c x$.

Case 3: $x \in DOM(\rho)$, then $x \notin \rho V$ by the idempotence of ρ and therefore $\rho^c x = x$ and $\rho x = y$ and $\rho^c y = x$. Now $\rho\rho^c x = \rho x$ and $\rho^c \rho x = \rho^c y = x = \rho^c x$.

□

Remark: There does not exist an inverse of a renaming substitution, since such an inverse must be injective which contradicts the idempotence. For example let $\rho = \{x \leftarrow y\}$ then $\rho x = \rho y$ but $x \neq y$ and let $\rho = \{x \leftarrow y, y \leftarrow x\}$ then ρ is injective but not idempotent.

If we refer to non-idempotent substitutions we define a variable renaming as an injective substitution with $COD(\rho) \subset \mathfrak{V}$ and $|DOM(\rho)| < \infty$. Then ρ is a bijective substitution called permutation in [Hu76]. For example $\rho = \{x \leftarrow y, y \leftarrow z, z \leftarrow x\}$ is a permutation.

4. *Union of Substitutions*

Since we can represent substitutions as sets (of pairs) the question arises if the union of two substitutions is again a substitution.

Two substitutions $\sigma,\tau \in \mathfrak{S}$ are called union-compatible, iff for all $x \in D = DOM(\sigma) \cap DOM(\tau)$ $\sigma x = \tau x$. Let

$$UC = \{(\sigma,\tau) \mid \sigma,\tau \in \mathfrak{S} \text{ and } \sigma,\tau \text{ are union-compatible}\}$$

be the set of pairs of union-compatible substitutions then we define

$$\sqcup : \begin{cases} UC & \longrightarrow \quad \mathfrak{T} \longrightarrow \mathfrak{T} \\ (\sigma,\tau) & \longrightarrow \quad \sigma\sqcup\tau : \mathfrak{T} \longrightarrow \mathfrak{T} \end{cases}$$

with $\sigma\sqcup\tau \; x = \sigma x$ if $x \in DOM(\sigma)$

and $\sigma\sqcup\tau \; x = \tau x$ else.

The following lemma which shows that our definitions are justified (i.e. $\sigma\sqcup\tau \in \Sigma$)is easy to prove

<u>*Lemma I.13:*</u> Let $(\sigma,\tau) \in UC$ then

(i) $\sigma\sqcup\tau : \mathfrak{T} \longrightarrow \mathfrak{T}$ is in Σ

(ii) $DOM(\sigma\sqcup\tau) = DOM(\sigma) \cup DOM(\tau)$

We are now looking for conditions under which the union of two substitutions is idempotent.

<u>*Lemma I.14:*</u> If for $(\sigma,\tau) \in UC$ the following conditions hold

(i) $\forall x \in DOM(\sigma)$ $\quad \tau\sigma x = \sigma x$

(ii) $\forall x \in DOM(\tau)$ $\quad \sigma\tau x = \tau x$

then $\sigma\sqcup\tau \in \mathfrak{S}$.

Proof: By Lemma I.13 it remains to be shown that $\sigma\sqcup\tau$ is idempotent. The proof is by structural induction on $(\sigma\sqcup\tau)x$.

Base case: I. Let $(\sigma\sqcup\tau)x = c \in \mathfrak{C}$, then $(\sigma\sqcup\tau)(\sigma\sqcup\tau)x = (\sigma\sqcup\tau)c$ $c = (\sigma\sqcup\tau)x$.

II. Let $(\sigma\sqcup\tau)x = y \in \mathfrak{V}$ then there are two cases:

Case 1: $y = x$: $(\sigma \sqcup \tau)(\sigma \sqcup \tau)x = (\sigma \sqcup \tau)x$.

Case 2: $y \neq x$: w.l.o.g. suppose $x \in DOM(\sigma)$, i.e. $\sigma \sqcup \tau x = \sigma x = y$. Since $\sigma \in \mathfrak{S}$ and $DOM(\sigma) \cap VCOD(\sigma) = \emptyset$ we have $y \notin DOM(\sigma)$ and hence $(\sigma \sqcup \tau)(\sigma \sqcup \tau)x = (\sigma \sqcup \tau)y = \tau y = \tau\sigma x$ and with (i) $\tau\sigma x = \sigma x = (\sigma \sqcup \tau)x$.

Induction Step: Let $(\sigma \sqcup \tau)x = f(t_1 \ldots t_n)$ and again w.l.o.g. let $x \in DOM(\sigma)$, i.e. $(\sigma \sqcup \tau)x = \sigma x$.

$$\begin{aligned}
(\sigma \sqcup \tau)(\sigma \sqcup \tau)x &= (\sigma \sqcup \tau)\sigma x \\
&= (\sigma \sqcup \tau)(\sigma \sqcup \tau)\sigma x && \text{(see Lemma I.15)} \\
&= (\sigma \sqcup \tau)(\sigma \sqcup \tau)\ f(t_1 \ldots t_n) \\
&= f((\sigma \sqcup \tau)(\sigma \sqcup \tau)t_1 \ldots (\sigma \sqcup \tau)(\sigma \sqcup \tau)t_n) \\
&= f((\sigma \sqcup \tau)t_1 \ldots (\sigma \sqcup \tau)t_n) && \text{(by induction hypothesis)} \\
&= (\sigma \sqcup \tau)f(t_1 \ldots t_n) \\
&= (\sigma \sqcup \tau)\sigma x \\
&= \sigma x && \text{(see Lemma I.15)} \\
&= (\sigma \sqcup \tau)x
\end{aligned}$$

□

Remark: The following conditions are equivalent to (i) and (ii) in Lemma I.14.

(i) $DOM(\tau) \cap VCOD(\sigma) = \emptyset$

(ii) $DOM(\sigma) \cap VCOD(\tau) = \emptyset$

To complete the proof of Lemma I.14 we have to show

Lemma I.15: Under the hypothesis of Lemma I.14 the following two equations hold

$$\forall x \in DOM(\sigma) \quad (\sigma \sqcup \tau)\sigma x = \sigma x$$
$$\forall x \in DOM(\tau) \quad (\sigma \sqcup \tau)\tau x = \tau x \quad .$$

Proof: We only show the first equation using again structural induction on σx.

Base case: I. $\sigma x = c \in \mathfrak{C}$, then $(\sigma \sqcup \tau)\sigma x = (\sigma \sqcup \tau)c = c = \sigma x$.

II. $\sigma x = y \in \mathfrak{V}$. Again we have two cases:

Case 1: $x = y$, then $x \notin DOM(\sigma)$.

Case 2: $x \neq y$. As in the proof of the above lemma we have $y \notin DOM(\sigma)$ and therefore $(\sigma \sqcup \tau)\sigma x = (\sigma \sqcup \tau)y = \tau y = \tau\sigma x$. With condition (i) of Lemma I.14 we get $\tau\sigma x = \sigma x$.

Induction step: Let $\sigma x = f(t_1 \ldots t_n)$ then

$$
\begin{aligned}
(\sigma \sqcup \tau)\sigma x &= (\sigma \sqcup \tau)\sigma\sigma x \\
&= (\sigma \sqcup \tau)\sigma\ f(t_1 \ldots t_n) \\
&= f((\sigma \sqcup \tau)\sigma\ t_1 \ldots (\sigma \sqcup \tau)\sigma\ t_n) \\
&= f(\sigma\ t_1 \ldots \sigma\ t_n) \qquad \text{(by induction hypothesis)} \\
&= \sigma\sigma x \\
&= \sigma x
\end{aligned}
$$

□

The following two lemmata state a connection between union and composition of substitutions.

<u>*Lemma I.16*</u>: For $\sigma \in \$$ and $\lambda \subset \sigma$, $\lambda \neq \sigma$ there exists $\lambda' \in SUB$ with $\lambda' \subset \sigma$, $(\lambda, \lambda') \in UC$ and $\sigma = \lambda \sqcup \lambda' = \lambda \circ \lambda' = \lambda' \circ \lambda$.

Proof: Define λ' with $\lambda' x = \sigma x$ for $x \in DOM(\sigma) \setminus DOM(\lambda)$ and $\lambda' x = x$ elsewhere. Since $\lambda' \subset \sigma$ we have $\lambda' \in \$$ and $(\lambda, \lambda') \in UC$. The equation follows from the definition of λ'.

□

<u>*Lemma I.17:*</u> If the conditions of Lemma I.14 are satisfied, we have

$$\sigma \sqcup \tau = \sigma\tau = \tau\sigma \quad .$$

Proof: It is sufficient to show $\sigma \sqcup \tau = \sigma\tau$, because $\sigma \sqcup \tau = \tau \sqcup \sigma = \tau\sigma$.

If $x \in DOM(\sigma) \cap DOM(\tau)$ then we have

$$\sigma \sqcup \tau x = \sigma x = \tau x = \sigma\sigma x = \sigma\tau x \quad .$$

If $x \in DOM(\sigma) \setminus DOM(\tau)$, then it is

$$\sigma \sqcup \tau x = \sigma x = \sigma\tau x \text{ , since } \tau x = x.$$

If $x \in DOM(\tau) \setminus DOM(\sigma)$, then by (ii) of Lemma I.14

$$\sigma \sqcup \tau x = \tau x = \sigma\tau x \quad .$$

If $x \notin DOM(\sigma) \cup DOM(\tau)$, then

$$\sigma \sqcup \tau x = x = \sigma x = \sigma\tau x \quad .$$

□

II. Matching

In the following we only investigate *finite* sets of terms and therefore the restriction to substitutions with a finite domain, i.e. $\sigma \in \mathcal{S}$ with $|DOM(\sigma)|<\infty$ is sufficient.

1. Instances of Terms

The *subsumption relation* $\leq$ turned out to be a very important relation on terms [Hu76], [Pl70], [Re70].

Def. II.1: Let $s,t \in \mathcal{T}$. We define

$$s \leq t \quad \text{iff} \quad \exists\lambda \in \mathcal{S} \qquad s = \lambda t$$

$$s \equiv t \quad \text{iff} \quad s \leq t \quad \text{and} \quad t \leq s \;.$$

We say ρ is an *instance* of t, t *subsumes* s or t is *more general than* s if $s \leq t$ and s is *equivalent* to t if $s \equiv t$. For example if $s = g(a\ b)$ and $t = g(x\ y)$ then we have $s \leq t$ with $\lambda = \{x \leftarrow a,\ y \leftarrow b\}$. If $s = g(x)$ and $t = x$ then $s \nleq t$ since $\{x \leftarrow g(x)\} \notin \mathcal{S}$. If $s = f(x\ y)$ and $t = f(u\ v)$ then $s \equiv t$ since $s \leq t$ and $t \leq s$ with $\lambda_1 = \{u \leftarrow x,\ v \leftarrow y\}$ and $\lambda_2 = \{x \leftarrow u,\ y \leftarrow v\}$. If $s = f(x\ y)$ and $t = f(y\ x)$ then $s \not\equiv t$. Since $s \nleq t$ and $t \nleq s$ ($\lambda = \{x \leftarrow y,\ y \leftarrow x\} \notin \mathcal{S}$). This demonstrates that the condition $\lambda \in \mathcal{S}$ restricts the set of pairs which are in the "$\leq$"-resp. "$\equiv$"-relation. But particularly in the last case this restriction is absolutely essential since it prevents the free functionsymbol f to satisfy the commutative axiom $f(x\ y) \equiv f(y\ x)$.

Lemma II.1: $\leq$ is a reflexive, but not a transitive relation.

Proof: With $\lambda = \varepsilon$ the reflexivity is trivial. A counterexample for the transitivity let $r = g(y)$, $s = x$ and $t = y$ then $r \leq s$ with $\{x \leftarrow g(y)\}$ and $s \leq t$ with $\{y \leftarrow x\}$ but $r \nleq t$.

□

The next lemma shows that $\equiv$ is not an equivalence relation.

Lemma II.2: $\equiv$ is a reflexive, symmetric but not a transitive relation.

Proof: Since $s \leq s$, $\equiv$ is reflexive and since $s \leq t$ and $t \leq s$ is equivalent to $t \leq s$ and $s \leq t$, $\equiv$ is symmetric. The counter-example for the transitivity is as follows. Let $r = f(x\ y)$, $s = f(u\ v)$, $t = f(y\ x)$ then $r \equiv s$ and $s \equiv t$ but $r \not\equiv t$.

□

The next two lemmata show the connection between the equivalence relation $\equiv$ and renaming substitutions.

Lemma II.3: For $s,t \in \mathfrak{T}$ if $s \equiv t$ then there exists a $\rho \in REN(V)$ with $V \subset var(s)$ such that $\rho s = t$.

Proof: With $s \equiv t$, i.e. $s \leq t$ and $t \leq s$ there exist $\hat{\sigma},\hat{\tau} \in \mathfrak{S}$ by definition such that $s = \hat{\tau}t$ and $t = \hat{\sigma}s$. Let $\sigma = \hat{\sigma}|_{var(s)}$ and $\tau = \hat{\tau}|_{var(t)}$, i.e. $DOM(\sigma) \subset var(s)$ and $DOM(\tau) \subset var(t)$.

We prove that σ is a renaming substitution. Since already $\sigma \in \mathfrak{S}$ by assumption we only have to show (a) $COD(\sigma) \subset \mathfrak{V}$ and (b) the injectivity-condition (I.6).

For (a) suppose there exists x $DOM(\sigma)$ with $\sigma x = r$ and $r \notin \mathfrak{V}$. Let $\alpha \in SEL^*$ with $\alpha(s)\downarrow$ and $\alpha(s) = x$. Then we have $\alpha(\sigma s) = r$ and $\alpha(\tau\sigma s) = \tau r \notin \mathfrak{V}$, but $\tau\sigma s = \sigma t = s$ and therefore $\alpha(\tau\sigma s) = \alpha(s) \in \mathfrak{V}$, which is a contradiction.

For (b) let $x,y \in DOM(\sigma) \subset var(s)$ and $\sigma x = \sigma y$. Then there exist $\alpha,\beta \in SEL^*$ with $\alpha(s)\downarrow$, $\beta(s)\downarrow$, $\alpha(s) = x$ and $\beta(s) = y$ and $\alpha(\sigma s) = \sigma x = \sigma y = \beta(\sigma s)$ and by Lemma I.1 (ii) $\alpha(\tau\sigma s) = \beta(\tau\sigma s)$. Since $\tau\sigma s = s$ we have $\alpha(\tau\sigma s) = x$ and $\beta(\tau\sigma s) = y$ and hence $x = y$.

□

Lemma II.4: For $s,t \in \mathfrak{T}$ if there exists $\rho \in REN(V)$ with $\rho s = t$, $V \subset var(s)$ and $\rho(V) \cap var(s) = \emptyset$, then $s \equiv t$.

Proof: Since $t \leq s$ by assumption we have to show $s \leq t$. By assumption $\rho(V) \cap var(s) = \emptyset$ and therefore $DOM(\rho) \cap var(s) = \emptyset$. Then

$$s = \rho^c s = \rho^c \rho s = \rho^c t \quad ,$$

i.e. $s \leq t$ since $\rho^c = \rho^c\rho$, and hence $s \equiv t$.

□

Summarizing:

Proposition II.1: Let $s,t \in \mathfrak{T}$, then $s \equiv t$ iff s and t are equal up to a renaming substitution.

2. *Instances of Substitutions*

We define a restricted equality of substitutions.

<u>Def. II.2:</u> Let $W \subset \mathfrak{V}$ and $\sigma, \tau \in \mathcal{S}$, then

$$\sigma = \tau[W] \text{ iff } \sigma x = \tau x \text{ for all } x \in W.$$

For example let $\sigma = \{x \leftarrow f(a\ b),\ y \leftarrow c\}$ and $\tau = \{x \leftarrow f(a\ b)\}$ then $\sigma = \tau[\{x\}]$. Next we extend the subsumption relation and the equivalence relation on terms to substitutions.

<u>Def. II.3:</u> Let $W \subset \mathfrak{V}$ and $\sigma, \tau \in \mathcal{S}$:

(i) $\sigma \le \tau$ iff $\exists \lambda \in \mathcal{S}$ $\sigma = \lambda\tau$; we say σ is a (*strong*) *instance* of τ or τ is *more general than* σ or τ *subsumes* σ.

(ii) $\sigma \le \tau[W]$ iff $\exists \lambda \in \text{SUB}$ $\sigma = \lambda\tau[W]$; we say σ is a (*weak*) *instance* of τ or τ is *more general than* σ with respect to W (w.r.t.W).

(iii) $\sigma \equiv \tau$ iff $\sigma \le \tau$ and $\tau \le \sigma$; we say σ and τ are *equivalent*.

(iv) $\sigma \equiv \tau[W]$ iff $\sigma \le \tau[W]$ and $\tau \le \sigma[W]$; we say σ and τ are *equivalent* w.r.t.W.

We give an example for each relation:

(i) $\sigma = \{x \leftarrow f(a\ b),\ u \leftarrow a\}$, $\tau = \{x \leftarrow f(u\ b)\}$, then $\sigma \le \tau$ with $\lambda = \{u \leftarrow a\}$

(ii) $\sigma = \{x \leftarrow f(a\ b)\}$ $\tau = \{x \leftarrow f(u\ b)\}$, then $\sigma = \tau[\{x\}]$ with $\lambda = \{u \leftarrow a\}$

(iii) $\sigma = \{x \leftarrow y\}$, $\tau = \{y \leftarrow x\} = \sigma^C$, then $\sigma \equiv \tau$ because $\sigma^C\sigma = \sigma^C = \tau$ and $\sigma = \sigma\sigma^C = \sigma\tau$

(iv) $\sigma = \{x \leftarrow f(u\ b)\}$, $\tau = \{x \leftarrow f(v,b)\}$, then $\sigma \equiv \tau[\{x\}]$ with $\lambda = \{u \leftarrow v\}$ and $\lambda' = \{v \leftarrow u\} = \lambda^C$

<u>*Lemma II.5:*</u> Let $W \subset \mathfrak{V}$

(i) $\le [W]$ is a reflexive, but not a transitive relation.

(iv) $\equiv [W]$ is a reflexive and symmetric, but not a transitive relation.

Proof: We only give the counterexamples to the transitivity, since reflexivity and symmetry are obvious.

$\leq$ [W]: $\rho = \{x \leftarrow f(g(y)b)\}$, $\sigma = \{x \leftarrow f(z\ b)\}$ and $\tau = \{x \leftarrow f(y\ b)\}$, then $\rho \leq \sigma[\{x\}]$ and $\sigma \leq \tau[\{x\}]$, but $\rho \nleq \tau[\{x\}]$.

$\equiv$ [W]: $\rho = \{z \leftarrow f(x\ y)\}$, $\sigma = \{z \leftarrow f(u\ v)\}$ and $\tau = \{z \leftarrow f(y\ x)\}$, then $\rho \equiv \sigma[\{z\}]$ and $\sigma \equiv \tau[\{z\}]$, but $\rho \not\equiv \tau[\{z\}]$.

□

Lemma II.6: Let $\sigma,\tau \in \$$. If there exists $\lambda \in \Sigma$ with $\sigma = \lambda\tau$ then there exists $\lambda' \in \$$ with $\sigma = \lambda'\tau$.

Proof: Since $\sigma = \lambda\tau$ and $\tau \in \$$ we have $\sigma = \lambda\tau = \lambda\tau\tau = \sigma\tau$ and $\sigma \in \$$.

□

The proof of this lemma yields a characterization of the subsumption relation on substitutions:

Corollary: Let $\sigma,\tau \in \$$. Then

$$\sigma \leq \tau \quad \text{iff} \quad \sigma = \sigma\tau\ .$$

Lemma II.7: "$\leq$" is a reflexive and transitive relation on substitutions.

Proof: With $\lambda = \varepsilon$ the reflexivity is immediate. For the transitivity, suppose $\rho \leq \sigma$ and $\sigma \leq \tau$, i.e. there exist $\kappa,\lambda \in \$$ with $\rho = \kappa\sigma$ and $\sigma = \lambda\tau$ therefore $\rho = \kappa\lambda\tau$ and $\kappa\lambda \in \Sigma$. By Lemma II.6 there exists $\mu \in \$$ with $\rho = \mu\tau$ and hence $\rho \leq \tau$.

□

Lemma II.8: "$\equiv$" is an equivalence relation on substitutions.

Proof: Reflexivity and symmetry are obvious. For the transitivity let $\rho \equiv \sigma$ and $\sigma \equiv \tau$, i.e. $\rho \leq \sigma$ and $\sigma \leq \rho$ and $\sigma \leq \tau$ and $\tau \leq \sigma$, but then we have $\rho \leq \tau$ and $\tau \leq \rho$ by the transitivity of $\leq$ and therefore $\rho \equiv \tau$.

□

But note that "$\equiv$" is not an equivalence relation on terms (confer Lemma II.2).

The following lemma shows the connection between the subsumption-relation and the subset-relation of substitutions.

Lemma II.9: Let $\sigma,\tau \in \mathcal{S}$. If $\sigma \subset \tau$, then $\tau \leq \sigma$.

Proof: We have to define a λ with $\tau = \lambda\sigma$. We choose $\lambda x = \tau x$ for $x \in DOM(\tau) \setminus DOM(\sigma)$ and $\lambda x = x$ else. Obviously $\lambda \in \mathcal{S}$ and $\tau = \lambda\sigma$.

□

The following example shows that if $\tau \leq \sigma$ then not necessarily $\sigma \subset \tau$: $\sigma = \{x \leftarrow f(a\ y)\}$ and $\tau = \{x \leftarrow f(a\ b),\ y \leftarrow b\}$.

A lemma which will often be used is the following:

Lemma II.10: Let $\sigma,\tau,\theta \in \mathcal{S}$. If $\sigma \leq \tau$ then $\sigma\theta \leq \tau\theta$ but in general $\theta\sigma \nleq \theta\tau$.

Proof: If $\sigma \leq \tau$ then there exists $\lambda \in \mathcal{S}$ with $\sigma = \lambda\tau$ and by Lemma I.10 $\sigma\theta = \lambda\tau\theta$, i.e. $\sigma\theta \leq \tau\theta$.

A counterexample for the second argument is
$\sigma = \{x \leftarrow f(a\ b),\ y \leftarrow b\}$, $\tau = \{x \leftarrow f(a\ y)\}$ and
$\theta = \{y \leftarrow a\}$, then $\theta\sigma = \{x \leftarrow f(a\ b),\ y \leftarrow b\}$ and
$\theta\tau = \{x \leftarrow f(a\ a),\ y \leftarrow a\}$, i.e. $\theta\sigma \nleq \theta\tau$.

□

3. *Matching of Terms*

The problem of finding a substitution $\sigma \in \mathcal{S}$ for two terms s and t such that $s = \sigma t$ is called a *matching problem* denoted as $<s \le t>$. We call σ a *matcher* of s and t and we write $M(s \le t)$ for the set of all matchers of s and t.

<u>Def. II.4:</u> For $s,t \in \mathcal{T}$ σ is a *most general matcher* of $<s \le t>$ (mgm) iff

(i) $\sigma \in M(s \le t)$

(ii) $\forall \tau \in M(s \le t) \quad \tau \le \sigma$

For example let $s = f(a\ b)$ and $t = f(x\ y)$, then $<s \le t>$ has a mgm $\sigma = \{x \leftarrow a,\ y \leftarrow b\}$. The next lemma shows, that every matcher of $<s \le t>$ restriced to var(t) is an mgm.

<u>*Lemma II.11:*</u> If σ is a matcher of $<s \le t>$, then

(i) $DOM(\sigma) \cap var(s) = \emptyset$ and

(ii) for $V = var(t)$ $\bar{\sigma} = \sigma|_V$ is a most general matcher.

Proof: (i) Suppose $x \in DOM(\sigma) \cap var(s)$. If $x \notin var(t)$, then since $x \notin VCOD(\sigma)$ $x \notin var(\sigma t)$ which is a contradiction to $x \in var(s) = var(\sigma t)$. Suppose $x \in var(t)$. Since $x \notin VCOD(\sigma)$ we have $x \notin var(\sigma t)$ which again is a contradiction to $x \in var(s)$.

(ii) Since $\bar{\sigma} t = \sigma t = s$ it is $\bar{\sigma} \in M(s \le t)$. Now let $\tau \in M(s \le t)$ and for evey $x \in DOM(\bar{\sigma})$, i.e. $x \in var(t)$, there exists $\alpha \in SEL^*$ with $\alpha(t)\downarrow$ and $\alpha(t) = x$. Since $s = \bar{\sigma} t$ and $s = \tau t$ we have $\alpha(s) = \alpha(\bar{\sigma} t) = \bar{\sigma} x$ and also $\alpha(s) = \alpha(\tau t) = \tau x$ and therefore $\bar{\sigma} x = \tau x$ for all $x \in DOM(\bar{\sigma})$. Hence it is $\bar{\sigma} \subset \tau$ and with Lemma II.9 we have $\tau \le \bar{\sigma}$.

□

<u>*Lemma II.12:*</u> If a mgm exists for $<s \le t>$ then it is unique.

Proof: Let σ and τ be mgm's for $<s \le t>$, then $DOM(\sigma) \subset V$ and $DOM(\tau) \subset V$ with $V = var(t)$. By the proof of Lemma II.11 we know $\sigma|_V = \tau|_V$ and hence $\sigma = \sigma|_V = \tau|_V = \tau$.

□

This lemma shows that the set of most general matchers is either empty or a singleton, in contrast to set of most general unifiers as shown in section III.

4. *Matching of Substitutions*

Similar to the matching of terms we define the matching of substitutions. The problem of finding a substitution $\lambda \in \mathcal{S}$ for $\sigma,\tau \in \mathcal{S}$ such that $\sigma = \lambda\tau$ ($\sigma = \lambda\tau[V]$ with $V \subset \mathcal{V}$) is called a *matching problem for substitutions*, denoted as $\langle\sigma \le \tau\rangle$ ($\langle\sigma \le \tau,V\rangle$). The solutions are called *matcher of* σ *and* τ (w.r.t. V) and the set of all solutions is $M(\sigma \le \tau)$ ($M(\sigma \le \tau,V)$).

Lemma II.13 Let $\sigma,\tau \in \mathcal{S}$

(i) For all $\lambda \in M(\sigma \le \tau)$ it is $DOM(\lambda) \subset DOM(\sigma)$

(ii) If $M(\sigma \le \tau) \neq \emptyset$, then there exists $\lambda \in \mathcal{S}$ such that $\sigma = \lambda\tau$ with $DOM(\lambda) \cap DOM(\tau) = \emptyset$.

Proof: (i) $DOM(\lambda) \subset DOM(\sigma)$ follows from Lemma I.9 (ii).

(ii) Let $\mu \in M(\sigma \le \tau)$ and define λ such that $\lambda x = \mu x$ if $x \in DOM(\mu) \setminus DOM(\tau)$ and $\lambda x = x$ else. Now $\lambda \in \mathcal{S}$ since $\mu \in \mathcal{S}$ and $\lambda \subset \mu$. From the definition follows $DOM(\lambda) \cap DOM(\tau) = \emptyset$. It remains to be shown that $\sigma = \lambda\tau$, where $\sigma = \mu\tau$ by assumption. Proof by structural induction on τx:

Base case: I. $\tau x = c$: $\sigma x = \mu\tau x = \mu c = \lambda c = \lambda\tau x$.

II. $\tau x = y$: We distinguish three cases:

Case 1: $x \neq y$ and $y \in DOM(\mu)$. Then $y \notin DOM(\tau)$ ($\tau \in \mathcal{S}$) and therefore $y \in DOM(\lambda)$ and $\mu y = \lambda y$. Hence $\sigma x = \mu\tau x = \mu y = \lambda y = \lambda\tau x$.

Case 2: $x \neq y$ and $y \notin DOM(\mu)$. Then $y \notin DOM(\lambda)$ and we have $\sigma x = \mu\tau x = \mu y = y = \lambda y = \lambda\tau x$.

Case 3: $x = y$. Then $x \notin DOM(\tau)$ and $\lambda x = \mu x$ and $\sigma x = \mu\tau x = \mu x = \lambda x = \lambda\tau x$.

Induction step: $\tau x = f(t_1 \ldots t_n)$:

$$\begin{aligned}
\sigma x &= \mu\tau x = \\
&= \mu\tau\tau x \\
&= \mu\tau\ f(t_1 \ldots t_n) \\
&= f(\mu\tau t_1 \ldots \mu\tau t_n) \\
&= f(\lambda\tau t_1 \ldots \lambda\tau t_n) \quad \text{(by induction hypothesis)} \\
&= \lambda\tau\ f(t_1 \ldots t_n) \\
&= \lambda\tau\tau x \\
&= \lambda\tau x
\end{aligned}$$

□

We are now looking for the connections between the matching of terms and the matching of substitutions.

Lemma II.14: For every matching problem $<\sigma \leq \tau, W>$ with $\sigma, \tau \in \mathcal{S}$ there exists a matching problem $<s \leq t>$ with $s, t \in \mathfrak{T}$ such that $M(s \leq t) = M(\sigma \leq \tau, W)$.

Proof: Let $\sigma, \tau \in SUB$, $W = \{v_1, \ldots, v_n\}$ $s_i = \sigma v_i$ and $t_i = \tau v_i$, and let h a n-ary functionsymbol not occuring in σ and τ. Then we define

$$s = h(s_1 \ldots s_n) \text{ and } t = h(t_1 \ldots t_n) \quad .$$

$M(s \leq t) \subset M(\sigma \leq \tau, W)$: For $\lambda \in M(s \leq t)$ we have

$$s = h(s_1 \ldots s_n) = \lambda t = \lambda h(t_1 \ldots t_n) = h(\lambda t_1 \ldots \lambda t_n)$$

and therefore

$$\sigma v_i = s_i = \lambda t_i = \lambda \tau v_i \qquad \text{for } i = 1, \ldots, n \quad .$$

Hence $\sigma = \lambda\tau[W]$.

$M(\sigma \leq \tau, W) \subset M(s \leq t)$: For $\lambda \in M(\sigma \leq \tau, W)$, i.e. $\sigma = \lambda\tau[W]$

we have $\sigma v_i = s_i = \lambda t_i = \lambda \tau v_i$ for $i = 1, \ldots, n$ and hence
$s = h(s_1 \ldots s_n) = h(\lambda t_1 \ldots \lambda t_2) = \lambda h(t_1 \ldots t_n) = \lambda t$.

□

Lemma II.15: For every matching problem $<\sigma \leq \tau>$ with $\sigma, \tau \in \mathcal{S}$ there exists a matching-problem $<s \leq t>$ with $s, t \in \mathfrak{T}$ such that

(i) $M(\sigma \leq \tau) \subset M(s \leq t)$

(ii) $\{\lambda \in M(s,t) \mid DOM(\lambda) \subset var(t)\} \subset M(\sigma \leq \tau)$.

Proof: (i) Choose s and t like in the proof of Lemma II.14 with $W = DOM(\sigma) \cup DOM(\tau) \cup VCOD(\tau)$. Now the proof of $M(\sigma \leq \tau) \subset M(s \leq t)$ is as in Lemma II.14.

Part (ii) is proved by cases. We have to show, that $\sigma = \lambda\tau$ for $\lambda \in M(s \leq t)$ with $DOM(\lambda) \subset var(t)$.

Case 1: $x \in DOM(\tau)$, then there exists $i = \{1, \ldots, n\}$ with $x = v_i$ and $\sigma x = \sigma v_i = s_i = \lambda t_i = \lambda \tau v_i = \lambda \tau x$.

Case 2: $x \notin DOM(\tau)$ and $x \in VCOD(\tau)$, then there exists $i \in \{1,\dots,n\}$ with $x = v_i$ and again $\sigma x = \lambda\tau x = \lambda x$.

Case 3: $x \notin DOM(\tau)$ and $x \notin VCOD(\tau)$ and $x \in DOM(\sigma)$ then there exists $i \in \{1\dots n\}$ with $x = v_i$ and $\sigma x = \lambda\tau x$.

Case 4: $x \notin W$, then $x \notin var(t)$ and therefore we have $\sigma x = x = \lambda x = \lambda\tau x$, since $DOM(\lambda) \subset var(t)$. □

Remark. If λ is a mgm of the above problem $\langle s \leq t \rangle$, the condition $DOM(\lambda) \subset var(\{t\})$ is always satisfied by Lemma II.11.

Lemma II.16: Let $\sigma \equiv \tau[W]$ then there exists $\rho \in REN(V)$ with $V \subset var(\tau(W))$ such that $\sigma = \rho\tau$.

Proof: By Lemma II.14 there exist terms s and t with $s \equiv t$. By Lemma II.3 there exists a $\rho \in REN(V)$ such that $s = \rho t$ and therefore $\sigma = \rho\tau[W]$ again by Lemma II.14.

□

Lemma II.17: Let $\sigma \equiv \tau$, then there exists a $\rho \in REN(V)$ with $V \subset var(\tau(W))$ such that $\sigma = \rho\tau$ where $W = DOM(\sigma) \cup DOM(\tau) \cup VCOD(\tau)$.

Proof: If $\sigma \equiv \tau$ then there exist $\lambda, \mu \in \$$ with $\sigma = \lambda\tau$ and $\mu\sigma = \tau$ and $\lambda \subset \sigma$ and $\mu \subset \tau$. By Lemma II.15 we have terms s and t with $s = \lambda t$ and $t = \mu s$ and by Lemma II.3 there exists a $\rho \in REN(V)$ with $V \subset var(t) = var(\tau(W))$ such that $s = \rho t$. By Lemma II.15 again we get the hypothesis $\sigma = \rho\tau$.

□

III. *Unification of Terms*

Let s,t ∈ $\mathfrak{T}$ be two terms. To *unify* s and t is to find a substitution σ ∈ $\mathfrak{S}$ such that σs = σt, in other words to solve the equation s = t. We write <s = t> for such a problem and σ is called a *unifier* of s and t. The set of all unifiers of s and t is denoted as U(s,t).

Def. III.1 A most general unifier (mgu) of s and t is a maximal element of U(s,t), i.e. every substitution σ ∈ $\mathfrak{S}$ with

(1) σ ∈ U(s,t)

(2) τ ≤ σ for all τ ∈ U(s,t)

is a mgu.

For example if s = f(g(x) h(a x) and t = f(y h(a b)) then σ = {y←g(b), x←b} is a mgu. We like to remark on the above definition:

1. As an immediate consequence of Lemma II.17 a most general unifier is unique modulo a renaming substitution. For instance, let s = x and t = y then σ = {x←y} and τ = {y←x} are both mgu of s and t.
2. Huet [Hu76] shows the existence of a mgu σ ∈ Σ, i.e. σ is not necessarily idempotent, in two ways. In the first method using the algebraic structure of $\mathfrak{T}$ he shows that $\mathfrak{T}/_{\equiv}$ is a join-semilattice and the existence of an mgu is equivalent to the existence of a greatest common instance. His proofs are strictly algebraic in contrast to [PL70] and [Re70].

In order to show the existence of an idempotent mgu we will however follow the second method of Huet using basic properties of certain congruences of terms.

1. Existence of an Idempotent Most General Unifier

First we need some definitions (see [Hu76]). Let $\sim$ be an equivalence relation on the set of terms $\mathfrak{T}$ and $[t]_\sim = \{t' \in \mathfrak{T} \mid t \sim t'\}$ the equivalence class of a term t, then $\sim$ is called *finite* iff $[x]_\sim = \{x\}$ for almost all $x \in \mathfrak{V}$, *simplifiable* (Ω-free in [SS82]) iff $f(t_1 \ldots t_n) \sim f(s_1 \ldots s_n)$ implies $t_i \sim s_i$ for $1 \le i \le n$ and *coherent* iff $f \neq g$ implies $f(t_1 \ldots t_n) \not\sim g(s_1 \ldots s_m)$. We call a finite simplifiable and coherent equivalence relation a *rational equivalence relation*.

Let $\sim$ be a rational equivalence relation then we define a relation $\rightarrow_\sim$ on $\mathfrak{T}/_\sim$ as follows:

$$[f(t_1 \ldots t_n)]_\sim \rightarrow_\sim [t_i]_\sim \quad \text{for} \quad 1 \le i \le n.$$

Let $\chi(t,\sim) \in \mathbb{N} \cup \{\infty\}$ be the length of the longest $\rightarrow_\sim$-chain starting with $[t]_\sim$. Then $\sim$ is said to be *acyclic* iff $\chi(t,\sim) < \infty$ for all $t \in \mathfrak{T}$.

Let $\sim$ be a rational acyclic equivalence relation then for all $x \in \mathfrak{V}$ we take as representative of the class $[x]_\sim$ an element $\tilde{x} \in [x]_\sim$ with the following properties

(i) $\tilde{x} \in [x]_\sim$
(ii) $\tilde{x} \in V$ iff $[x]_\sim \subset \mathfrak{V}$

and define the substitution $\sigma_\sim \in \Sigma$ as

$$\sigma_\sim x = \begin{cases} \tilde{x} & \text{if } \tilde{x} \in \mathfrak{V} \\ \sigma_\sim \tilde{x} & \text{else} \end{cases} .$$

An equivalence relation $\sim$ is said to be a *congruence* iff $s_i \sim t_i$, $1 \le i \le n$, implies $f(s_1 \ldots s_n) \sim f(t_1 \ldots t_n)$. Let $\sigma \in \Sigma$ be a substitution then the *unification congruence* $\sim_\sigma$ of σ defined as

$$s \sim_\sigma t \text{ iff } \sigma s = \sigma t$$

is a rational acyclic congruence.

<u>*Lemma III.1:*</u> Let $\sim$ be a rational acyclic equivalence relation, $\hat{\sim}$ the congruence generated by $\sim$. Then

(i) $\sim_{\sigma_\sim} = \hat{\sim}$

(ii) for all $t \in \mathfrak{T}$ $\quad \sigma_\sim t \mathrel{\hat{\sim}} t$.

For a proof see Lemma 5.25 and Lemma 5.26 in [Hu76].

The following is the Unification Theorem of Huet (cf. Theorem 16 in [Hu76]):

Theorem: Let E be a finite set of terms, $\sim_E$ the smallest equivalence relation containing the pairs of terms of E. Let $\sim$ be the simplifiable closure of $\sim_E$. If $\sim$ is coherent and acyclic then $\sigma_\sim$ is an mgu of E else E is not unifiable.

Lemma III.2: Let s and t be two unifiable terms then there exists an idempotent mgu of s and t.

Proof: Since s and t are unifiable, the equivalence relation $\sim$ in the Unification Theorem is coherent and acyclic and therefore $\sigma_\sim$ is a mgu of s and t. By Lemma III.1 (ii) it is $\sigma_\sim t \hat{\sim} t$ for all $t \in \mathfrak{T}$, by Lemma III.1 (i) $\sigma_\sim t \sim_{\sigma_\sim} t$ and by definition $\sigma_\sim \sigma_\sim t = \sigma_\sim t$. Hence $\sigma_\sim$ is idempotent, i.e. $\sigma_\sim \in \mathfrak{S}$. □

The next lemmas show that the domain and the variables of the codomain of an mgu $\sigma \in \mathfrak{S}$ are a subset of the variables of the terms to be unified.

Lemma III.3: Let $\sigma \in \mathfrak{S}$ be an mgu of s and t then $DOM(\sigma) \subset var(s,t)$.

Proof: Suppose by contradiction that $x \in DOM(\sigma)$ and $x \notin var(s,t)$. Define $\sigma' \in \mathfrak{S}$ by $\sigma' = \sigma|_{var(s,t)}$. Now it is $\sigma's = \sigma s = \sigma t = \sigma't$ and therefore σ' is a unifier of s and t. Since $\sigma' \subset \sigma$ we have $\sigma \leq \sigma'$. But by assumption σ is mgu, i.e. $\sigma' \leq \sigma$ and hence $\sigma \equiv \sigma'$.

If $\sigma x \notin \mathfrak{V}$ it is $x = \sigma'x \neq \lambda\sigma x$ for all $\lambda \in \mathfrak{S}$ which is a contradiction to $\sigma' \leq \sigma$. If $\sigma x \in \mathfrak{V}$ $\sigma x = z$ and $z \neq x$ then $z \notin DOM(\sigma)$ (since σ is idempotent) and $x = \sigma'x = \lambda\sigma x = \lambda z$ and $z = \sigma'z = \lambda\sigma z = \lambda z = x$ which again is a contradiction to $z \neq x$.

□

Lemma III.4 If $\sigma \in SUB$ be an mgu of s and t then $VCOD(\sigma) \subset var(s,t)$.

Proof: Suppose by contradiction that there exists a $z \in VCOD(\sigma)$ and $z \notin var(s,t)$. Define a substitution $\theta \in \mathfrak{S}$ with $\theta x = \{z \leftarrow z'\}\sigma x$

for $x \in DOM(\sigma)$ and $\theta x = x$ else, where z' is a variable not occuring in σ,s and t. Firstly we show that θ unifies s and t. Since $\sigma s = \sigma t$ and $z \notin var(s,t)$ and all occurences of z in $\sigma s, \sigma t$ introduced by σ are replaced by z' in $\theta s, \theta t$, we have $\theta s = \theta t$. Now since σ is an mgu we have $\theta = \theta\sigma$ by the corollary of Lemma II.6. Let $y \notin DOM(\sigma)$ with $z \in var(\sigma y)$, then $z \in var(\theta\sigma y)$ since $z \notin DOM(\sigma)$ (σ is idempotent), but $z \notin var(\theta y)$ by construction, but this contradicts the fact $\theta y = \theta\sigma y$.

□

This property is typical for unification of free terms. In some applications as for example T-unification the contrary is demanded: $VCOD(\sigma) \cap var(s,t) = \emptyset$.

Let D be a set of at least two terms. D is said to be unifiable iff there exists a substitution $\sigma \in SUB$ such that σD is a singeton. We call σ a unifier and we write $U(D)$ for the set of all unifiers of D.

Def. III.2 Let D be a set of at least two terms. Then $\sigma \in SUB$ is called a most general unifier (mgu) of D iff

(1) $\sigma \in U(D)$

(2) $\tau \leq \sigma$ for all $\tau \in U(D)$.

The following lemma shows that the unification of sets of terms can be reduced to the unification of two terms.

Lemma III.5 Let $D = \{s_1,\ldots,s_2\}$ be a unifiable set of at least two terms. Then there exist terms s and t with

$$U(D) = U(s,t).$$

Proof: Let h be an n-1-ary functionsymbol not occuring in D and let $s = h(s_1,\ldots,s_1)$ and $t = h(s_2,\ldots,s_n)$. Now it is

$$\begin{aligned} \sigma \in U(D) \quad &\text{iff} \quad \sigma s_1 = \ldots = \sigma s_2 \\ &\text{iff} \quad \sigma s = h(\sigma s_1 \ldots \sigma s_1) = h(\sigma s_2 \ldots \sigma s_n) = \sigma t \\ &\text{iff} \quad \sigma \in U(s,t). \end{aligned}$$

□

The following lemma establishes a connection between most general matching and most general unification.

Lemma III.6 Let s,t be terms. Every most general matcher for $<s \leq t>$ is a most general unifier for $<s = t>$.

Proof: Obviously every mgm is a unifier: Let σ be an mgm for $<s \leq t>$, i.e. $s = \sigma t$. Since $\sigma \in \mathcal{S}$ we have $\sigma s = \sigma\sigma t = \sigma t$.

Now we have to show that σ is a most general unifier, i.e. $\tau \leq \sigma$ for every $\tau \in U(s,t)$.

For any $x \in \mathcal{V}$ if $x \notin DOM(\sigma)$ then $\tau x = \tau\sigma x$.

If $x \in DOM(\sigma)$ then $x \in var(t)$ by Lemma II.11 and therefore there exists $\alpha \in SEL^*$ with $\alpha(t)\downarrow$, $\alpha(s)\downarrow$, $\alpha(t) = x$ and $\alpha(s) = \alpha(\sigma t) = \sigma\alpha(t) = \sigma x$. Since $\tau \in U(s,t)$ we have $\tau t = \tau s$ and therefore $\alpha(\tau t) = \tau\alpha(t) = \tau x$ and $\alpha(\tau s) = \tau\alpha(s) = \tau\sigma x$, i.e. $\tau x = \tau\sigma x$. Summarizing we have $\tau = \tau\sigma$ and hence $\tau \leq \sigma$.

□

Under some restricted conditions on the mgu we state a converse of the last lemma:

Lemma III.7 Let s,t be terms, σ a most general unifier for $<s = t>$. If $\sigma|_W = \rho \in REN(W)$ with $W = DOM(\sigma) \cap var(t)$ such that $COD(\rho) \cap var(t) = \emptyset$, then $\rho^c\sigma$ is a mgm for $<s \leq t>$.

Proof: Let ρ be as above and $\sigma = \{x_1 \leftarrow y_1, \ldots, x_k \leftarrow y_k, x_{k+1} \leftarrow t_{k+1}, \ldots, x_n \leftarrow t_n\}$ and $W = \{x_1, \ldots, x_k\}$, then $\rho^c\sigma = \{y_1 \leftarrow x_1, \ldots, y_k \leftarrow x_k, x_{k+1} \leftarrow \rho^c t_{k+1}, \ldots, x_n \leftarrow \rho^c t_n\}$. But since $COD(\rho) \cap var(t) = \emptyset$, we have $DOM(\rho^c\sigma) \cap var(t) = \emptyset$ and hence $\rho^c\sigma s = \rho^c\sigma t = t$, i.e. $\rho^c\sigma$ is a matcher for $<s \leq t>$. Since $var(\sigma) \cap var(s,t)$ it is $DOM(\rho^c\sigma) \subset var(t_1)$ and hence by Lemma II.11 $\rho^c\sigma$ is a mgm for $<s \leq t>$.

□

The final two lemmata of this chapter are concerned with unification and renaming substitutions.

<u>Def. III.3</u> Let s,t be terms. We say s and t are R-unifiable if there exists a renaming substitution $\rho \in REN(var(s,t))$ such that s and ρt are unifiable.

For instance let $s = x$ and $t = f(x)$ then s and t are not unifiable but s and ρt are unifiable with $\rho = \{x \leftarrow z\}$, i.e. s and t are R-unifiable. If s and t are unifiable then s and t are clearly R-unifiable.

Lemma III.8: Let $s,t \in \mathfrak{T}$ and $var(s) \cap var(t) = \emptyset$, i.e. s and t are variable disjoint. If s and t are R-unifiable then they are unifiable.

Proof: Let $var(s) \cap var(t) = \emptyset$ and $\rho \in REN(var(s,t))$ such that s and ρt are unifiable. Let $\sigma \in \mathfrak{S}$ be a unifier of s and ρt and $\rho' = \rho_{|var(t)}$. Hence we have $\rho' s = s$ and $\rho' t = \rho t$ and therefore $\sigma\rho' s = \sigma s = \sigma\rho t = \sigma\rho' t$, i.e. s and t are unifiable.

□

The following lemma is quite technical.

Lemma III.9: Let s and t be terms and $\sigma \in \mathfrak{S}$. If σs and t are R-unifiable then s and t are R-unifiable.

Proof: By definition there exists $\rho \in REN(V)$ with $V = var(s,t,\sigma s)$ such that σs and σt are unifiable, $var(\sigma s) \cap var(\rho t) = \emptyset$ and w.l.o.g. $DOM(\sigma) \cap var(\rho t) = \emptyset$. Hence $\sigma\rho t = \rho t$ and there exists $\theta \in \mathfrak{S}$ such that $\theta\sigma s = \theta\rho t = \theta\sigma\rho t$, i.e. s and t are R-unifiable.

□

Lemma III.10: Let $s_1, s_2, t_1, t_2 \in \mathfrak{T}$, let σ be an mgu of s_1, s_2 and let τ be an mgu of t_1, t_2. Let the following variable conditions be satisfied:

(i) $var(s_2) \cap var(s_1,t_1) = \emptyset$,
(ii) $var(t_2) \cap var(s_1,t_1) = \emptyset$,
(iii) $var(s_2) \cap var(t_2) = \emptyset$.

Then: $\tau s_1, s_2$ are R-unifiable iff $\sigma t_1, t_2$ are R-unifiable.

Proof: Since σ is an mgu of s_1 and s_2, we have $var(\sigma) \subset var(s_1,s_2)$ and by the same argument $var(\tau) \subset var(t_1,t_2)$. Therefore $\sigma t_2 = t_2$ by (ii) and (iii) $\tau s_2 = s_2$ by (i) and (iii), $var(\sigma t_1) \cap var(t_2) = \emptyset$ by (ii) and (iii) and $var(\tau s_1) \cap var(s_2) = \emptyset$ again by (i) and (iii). Now $\tau s_1, s_2$ are R-unifiable iff $\tau s_1, s_2$ are unifiable (Lemma III.8) iff σ and τ are compatible (by the corollary of Proposition IV.1) iff σt_1 and t_2 are unifiable iff σt_1 and t_2 are R-unifiable (again by Lemma III.8).

□

IV. Unification of Substitutions

In this chapter we present some results for the unification of substitutions.

Let $\sigma,\tau \in \mathcal{S}$ then σ and τ are called *compatible* or *unifiable* (short: σ comp τ) iff there exists $\lambda \in \mathcal{S}$ such that $\lambda\sigma = \lambda\tau$. We write $U(\sigma,\tau) = \{\lambda \in \mathcal{S} \mid \lambda\sigma = \lambda\tau\}$ for the set of all unifiers of σ and τ.

Def. IV.1: Let $\sigma,\tau \in \mathcal{S}$. A substitution θ is called a most general unifier (mgu) of σ and τ iff

(i) $\theta \in U(\sigma,\tau)$

(ii) $\lambda \leq \theta$ for all $\lambda \in U(\sigma,\tau)$.

If θ is an mgu of σ,τ we call the substitution $\theta\sigma = \theta\tau$ a *unifying composition* or a *merge* of σ and τ and write

$$\sigma * \tau = \{\theta \in \mathcal{S} \mid \theta \equiv \lambda\sigma \text{ and } \lambda \text{ is mgu of } \sigma \text{ and } \tau\}$$

for the set of all merges of σ and τ. In the corollary of Lemma IV.2 it is shown that $\sigma * \tau$ is not empty, if σ comp τ. Let $\sigma * \tau$ always denote an arbitrary element of $\sigma * \tau$.

Just as the mgu of two terms is unique modulo a renaming substitution the set of all merges of two substitutions $\sigma * \tau$ contains only elements that differ under renaming, in other words: $\theta_1,\theta_2 \in \sigma * \tau$ iff $\sigma_1 \equiv \sigma_2$.

The $*$-operation is a commutative operation.

As in the case of matching of substitutions, for each unification problem of substitutions there exists an equivalent unification problem of terms.

Lemma IV.1: Let $\sigma,\tau \in \mathcal{S}$. Then there exist terms $s,t \in \mathcal{T}$ with

$$U(\sigma,\tau) = U(s,t) \quad .$$

Proof: Let $W = DOM(\sigma) \cup DOM(\tau) = \{w_1,\ldots,w_n\}$, let h be a "new" n-ary functionsymbol and let $s = h(\sigma w_1\ldots\sigma w_n)$ and $t = h(\tau w_1\ldots\tau w_n)$. Now if σ comp τ then there exists $\lambda \in \mathcal{S}$ such that $\lambda\sigma = \lambda\tau$ and therefore $\lambda s = h(\lambda\sigma w_1\ldots\lambda\sigma w_n) = h(\lambda\tau w_1\ldots\lambda\tau w_n) = \lambda t$, i.e. s and t are unifiable.

If s and t are unifiable then there exists $\lambda \in \$$ such that $\lambda s = \lambda t$. Now if $x \in W$, i.e. $x = w_i$ for $1 \le i \le n$, we have $\lambda\sigma w_i = \lambda\tau w_i$, and if $x \notin W$, i.e. $\sigma x = x = \tau x$ we have $\lambda\sigma x = \lambda\tau x$ and hence $\lambda\sigma = \lambda\sigma$, i.e. σ comp τ.

Now $\lambda \in U(\sigma,\tau)$ iff $\lambda\sigma = \lambda\tau$ iff $\lambda s = \lambda\tau$ iff $\lambda \in U(s,t)$.

□

Corollary: Let $\sigma,\tau \in \$$. If σ comp τ then there exists an mgu of σ and τ.

Proof: By the above lemma $U(\sigma,\tau) = U(s,t)$ and hence by Lemma III.2 there exists an mgu for s and t, which is an mgu for σ and τ.

□

In order to actually compute a unifying composition of two substitutions σ,τ we construct a pair of terms whose mgu is the unifying composition we are looking for. Of course we could just compute a λ on the basis of Lemma IV.1 such that $\lambda\sigma = \lambda\tau$ and then compute $\sigma * \tau$ from $\lambda\sigma$ which is considerably more inefficient then the method of the following lemma. Moreover in [Ch72], [CL73] and [Ni80] a unifying composition is defined as an mgu of this pair of terms. In the second part of the lemma we show that this definition is compatible with ours.

Lemma IV.2: Let $\sigma,\tau \in \$$. Then there exist terms s and t such that

(i) σ,τ are compatible iff s,t are unifiable

(ii) if λ is a mgu of s and t, then λ is a unifying composition of σ and τ, i.e. $\lambda \in \sigma \circledast \tau$.

Proof: Let $\sigma = \{x_1 \leftarrow s_1, \ldots, x_n \leftarrow s_n\}$, $\tau = \{y_1 \leftarrow t_1, \ldots, y_m \leftarrow t_m\}$ and let h a (n+m)-ary functionsymbol not occuring in s_i or t_i. Define $s = h(x_1 \ldots x_n\ y_1 \ldots y_m)$ and $t = h(s_1 \ldots s_2\ t_1 \ldots t_m)$.

(i) If σ comp τ, i.e. if there exists a $\mu \in \$$ with $\mu\sigma = \mu\tau$, then $\lambda := \mu\sigma = \mu\tau$ is an unifier of s and t: With $\lambda x_i = \mu\sigma x_i = \mu\sigma\sigma x_i = \mu\sigma s_i = \lambda s_i$ and $\lambda y_i = \mu\tau y_i = \mu\tau\tau y_i = \mu\tau t_i = \lambda t_i$, we have

$$\begin{aligned}\lambda t &= \lambda\ h(x_1 \ldots x_n\ y_1 \ldots y_m) \\ &= h(\lambda x_1 \ldots \lambda x_n\ \lambda y_1 \ldots \lambda y_m) \\ &= h(\lambda s_1 \ldots \lambda s_n\ \lambda t_1 \ldots \lambda t_m) \\ &= \lambda\ h(s_1 \ldots s_n\ t_1 \ldots t_m) \\ &= \lambda t, \text{ i.e. s and t are unifiable by } \lambda.\end{aligned}$$

Now let s and t be unifiable, i.e. there exists a λ with $\lambda s = \lambda t$. We show by cases that $\lambda\sigma x = \lambda\tau x$ for all $x \in \mathfrak{V}$.

Case 1: $x \in DOM(\sigma) \cap DOM(\tau)$, then there exists i,j with $1 \le i \le n$ and $1 \le j \le m$ and $x = x_i = y_i$ such that $\lambda x = \lambda x_i = \lambda s_i = \lambda\sigma x_i = \lambda\sigma x$ and $\lambda x = \lambda y_i = \lambda t_j = \lambda\tau y_j = \lambda\tau x$.

Case 2: $x \in DOM(\sigma) \setminus DOM(\tau)$, then there exists i with $1 \le i \le n$ and $x = x_i$ and $\tau x = x$. Hence we have $\lambda\tau x = \lambda x = \lambda x_i = \lambda s_i = \lambda\sigma x_i = \lambda\sigma x$.

Case 3: $x \in DOM(\tau) \setminus DOM(\sigma)$ like case two.

Case 4: $x \notin DOM(\tau) \cup DOM(\sigma)$, i.e. $\tau x = x = \sigma x$ and therefore $\lambda\tau x = \lambda x = \lambda\sigma x$.

(ii) For the second part we have to show $\lambda \in \sigma \circledast \tau$. Since λ is a mgu of s and t, we have

(1) $\quad \forall\mu \in U(s,t) \quad \mu \le \lambda$.

Since σ comp τ, there exists a $\theta \in \mathcal{S}$ with $\sigma * \tau \equiv \theta\sigma = \theta\tau$ and

(2) $\quad \forall\nu \in U(\sigma,\tau) \quad \nu \le \theta$.

We have shown in (i) that $\lambda\sigma = \lambda\tau$ and $\sigma * \tau\, s = \sigma * \tau\, t$. By (2) we have $\lambda \le \theta$ and hence with Lemma II.10

(3) $\quad \lambda\sigma \le \theta\sigma$.

Using (1) we have

(4) $\quad \theta\sigma \equiv \sigma * \tau \le \lambda$.

But then if

(5) $\quad \lambda = \lambda\sigma$

we get

$$\underset{(5)}{\lambda = \lambda\sigma} \underset{(3)}{\leq} \theta\sigma \equiv \sigma * \tau \underset{(4)}{\leq} \lambda$$

and therefore $\lambda \equiv \sigma * \tau$, i.e. $\lambda \in \sigma * \tau$.

We have to show (5) $\lambda = \lambda\sigma$. If $x \in DOM(\sigma)$ then there exists i with $1 \leq i \leq n$, $x = x_i$ and $\sigma x_i = s_i$ and hence we have $\lambda x_i = \lambda s_i = \lambda\sigma x_i$. If $x \notin DOM(\sigma)$ then $\lambda\sigma x = \lambda x$, which finishes the proof.

□

Corollary: If σ and τ are compatible there always exists a unifying composition, which is idempotent.

Proof: By Lemma IV.2 a unifying composition is an mgu of two terms and hence by Lemma III.2 there exists an idempotent unifying composition.

□

Lemma IV.3: Let $\sigma, \tau \in \$$. If σ comp τ, then

$$\sigma * \tau \equiv \sigma(\sigma * \tau) \equiv \tau(\sigma * \tau) \quad \text{and}$$
$$\sigma * \tau = (\sigma * \tau)\sigma = (\sigma * \tau)\tau \ .$$

Proof: Since σ comp τ there exists $\lambda \in \$$ such that $\sigma * \tau \equiv \lambda\sigma = \lambda\tau$, i.e. $\sigma * \tau = \rho\lambda\sigma = \rho\lambda\tau$ with a renaming substitution ρ. Now we have $\sigma * \tau = \rho\lambda\sigma = \rho\lambda\sigma\rho\lambda\sigma \leq \sigma\rho\lambda\sigma = \sigma(\sigma * \tau) \leq \sigma * \tau$ and similarily $\sigma * \tau = \rho\lambda\tau = \rho\lambda\tau\rho\lambda\tau \leq \tau\rho\lambda\tau = \tau(\sigma * \tau) \leq \sigma * \tau$.

The second equation is trivial: $\sigma * \tau = \rho\lambda\sigma = \rho\lambda\sigma\sigma = (\sigma * \tau)\sigma$ and $\sigma * \tau = \rho\lambda\tau = \rho\lambda\tau\tau = (\sigma * \tau)\tau$.

□

The motivation for this lemma is the definition of a unifying composition of two substitutions in [Si76] which is as follows:

> a unifying composition $\gamma = \sigma \cdot \tau$ of two substitutions σ and τ is a most general substitution λ such that
>
> (*) $\gamma = \gamma\sigma = \gamma\tau = \sigma\gamma = \tau\gamma$

i.e. for all substitutions δ satisfying (*) $\delta \leq \gamma$. But this definition is different from ours. Take e.g. $\sigma = \{x \leftarrow y\}$ and $\tau = \{y \leftarrow x\}$ then $\sigma \circledast \tau = \{\{x \leftarrow y\}, \{y \leftarrow x\}\}$. But $\gamma = \sigma \cdot \tau = \{x \leftarrow z, y \leftarrow z\}$ where z is a variable different from x and y, since by (*) we get $\gamma x = \gamma \sigma x = \gamma y$ and w.l.o.g. we can assume that $DOM(\gamma) \subset \{x,y\}$. Now suppose $\gamma y = y$ and $\gamma x = y$ then $\tau \gamma x = \tau y = x$ and $\gamma x = y$ which contradicts (*). If $\gamma x = x$ and $\gamma y = x$ then $\sigma \gamma y = \sigma x = y$ and $\gamma y = x$ which again contradicts (*). Hence we have $\sigma \cdot \tau \leq \sigma * \tau$ and $\sigma * \tau \not\leq \sigma \cdot \tau$.

The next lemma will be frequently used in the sequel.

Lemma IV.4: Let $\delta, \sigma, \tau \in \mathcal{S}$.

(i) If $\delta \leq \sigma$ and $\delta \leq \tau$, then σ comp τ and $\delta \leq \sigma * \tau$.

(ii) If σ comp τ, then $\sigma * \tau \leq \sigma$ and $\sigma * \tau \leq \tau$.

(iii) If σ comp τ, and $\sigma \leq \delta$ or $\tau \leq \delta$, then $\sigma * \tau \leq \delta$.

Proof: (i) If $\delta \leq \sigma$ and $\delta \leq \tau$, then we have by the corollary of Lemma II.6 $\delta = \delta\tau = \delta\sigma$ and hence σ comp τ. By definition there exists a $\lambda \in \mathcal{S}$ such that $\lambda\sigma = \lambda\tau \equiv \sigma * \tau$ and $\delta \leq \lambda$. Hence by Lemma II.10 we have $\delta = \delta\sigma \leq \lambda\sigma \equiv \sigma * \tau$, i.e. $\delta \leq \sigma * \tau$ by the transitivity of "$\leq$" of substitutions.

(ii) Since σ comp τ there exists $\lambda \in \mathcal{S}$ such that $\sigma * \tau \equiv \lambda\sigma = \lambda\tau$, hence $\sigma * \tau \leq \sigma$ and $\sigma * \tau \leq \tau$.

(iii) By definition there again exists $\lambda \in \mathcal{S}$ such that $\sigma * \tau \equiv \lambda\sigma = \lambda\tau$. With Lemma II.10 we have $\lambda\sigma \leq \lambda\delta$ or $\lambda\tau \leq \lambda\delta$ and hence using again the transitivity $\sigma * \tau \leq \delta$.

□

Lemma IV.5: Let $\theta, \sigma, \tau \in \mathcal{S}$. If $\sigma \leq \tau$ and θ comp σ, then θ comp τ and $\theta * \sigma \leq \theta * \tau$.

Proof: Since θ comp σ, we have using Lemma IV.4 (ii) $\theta * \sigma \leq \sigma$ and $\theta * \sigma \leq \theta$. Using the transitivity of $\leq$ (Lemma II.7) and $\sigma \leq \tau$ we have $\theta * \sigma \leq \tau$ and hence with Lemma IV.4 (i) θ comp τ and $\theta * \sigma \leq \theta * \tau$.

□

Corollary: Let $\theta, \sigma, \tau \in \$$. If $\sigma \equiv \tau$ and θ comp σ, then θ comp τ and $\theta * \tau \equiv \theta * \tau$.

Next we want to state that merging substitutions is an associative operation:

Lemma IV.6: Let $\theta, \sigma, \tau \in \$$. If σ comp τ, θ comp σ, θ comp $\sigma * \tau$ and $\theta * \sigma$ comp τ then $\Theta * (\sigma * \tau) \equiv (\theta \times \sigma) * \tau$.

Proof: Let $\delta_l = \theta * (\sigma * \tau)$ and $\delta_r = (\theta * \sigma) * \tau$. By Lemma IV.4 (ii) we have $\delta_l \leq \theta$ and $\delta_l \leq \sigma * \tau$. By the transitivity of $\leq$ and using Lemma IV.4 (ii) we get $\delta_l \leq \Theta$ and ($\delta_l \leq \sigma$ and $\delta_l \leq \tau$). Hence we have ($\delta_l \leq \theta$ and $\delta_l \leq \sigma$) and $\delta_l \leq \tau$ and using Lemma IV.4 (i) $\delta_l \leq \theta * \sigma$ and $\delta_l \leq \tau$. Thus $\delta_l \leq (\theta * \sigma) * \tau = \delta_r$. In a similar way we show that $\delta_r \leq \delta_l$ and therefore $\delta_l \equiv \delta_r$.

□

Corollary: Let $\theta, \sigma, \tau \in \$$, then σ comp τ and θ comp $\sigma * \tau$ iff θ comp σ and $\theta * \sigma$ comp τ.

Proof: By the assumption there exists $\delta = \theta * (\sigma * \tau)$ and by Lemma IV.4 (ii) $\delta \leq \theta$, $\delta \leq \sigma$ and $\delta \leq \tau$ as in the proof of Lemma IV.6. Then by Lemma IV.4 (i) θ comp σ and $\delta \leq \theta * \sigma$ and again by Lemma IV.4 (i) $\theta * \sigma$ comp τ. The converse is shown in the same way.

□

Def. IV.2: A set of substitutions $\{\theta_i \mid 1 \leq i \leq n\}$ with $n \geq 2$ is said to be compatible or unifiable iff there exists a $\sigma \in \$$ such that $\sigma\theta_i = \sigma\theta_k$ for $1 \leq i, k \leq n$, i.e. $\{\sigma\theta_i \mid 1 \leq i \leq n\}$ is a singleton. Let $U(\{\theta_i \mid 1 \leq i \leq n\}) = \{\sigma \mid \sigma\theta_i = \sigma\theta_k, 1 \leq i, k \leq n\}$ be the set of all such substitutions.

Lemma IV.7: Let $\{\theta_i \mid 1 \leq i \leq n\}$ be a set of at least two substitutions. Then there exists a set of terms $\{s_i \mid 1 \leq i \leq n\}$ with

$$U(\{\theta_i \mid 1 \leq i \leq n\}) = U(\{s_i \mid 1 \leq i \leq n\}).$$

Proof: Let $W = \bigcup_{i=1}^{n} DOM(\theta_i) = \{x_1, \ldots, x_k\}$ and let h be a new functionsymbol. Then we define

$$s_i = h(\theta_i x_1, \ldots, \theta_i x_k) \quad \text{for} \quad 1 \leq i \leq n .$$

The proof is now the same as for Lemma IV.1.

□

Def. IV.3: A unifier σ of $\{\theta_i \mid 1 \le i \le n\}$ is called a most general unifier (mgu) iff

(i) $\sigma \in U(\{\theta_i \mid 1 \le i \le n\})$

(ii) $\tau \le \sigma$ for all $\tau \in U(\{\theta_i \mid 1 \le i \le n\})$.

If σ is an mgu of $\{\theta_i \mid 1 \le i \le n\}$ then $\sigma\theta_i$ is called a unifying composition or merge of $\{\theta_i \mid 1 \le i \le n\}$. We write

$$\theta_1 \circledast \ldots \circledast \theta_n = \{\lambda \in \$ \mid \lambda \equiv \sigma\theta_1 \text{ and } \sigma \text{ mgu of } \{\theta \mid 1 \le i \le n\}\}$$

for the set of all unifying compositions, which is again not empty if $\{\theta_i \mid 1 \le i \le n\}$ is compatible.

The following lemma is the analogue of Lemma IV.2.

Lemma IV.8: Let $\{\theta_i \mid 1 \le i \le n\}$ be a set of at least two substitutions.

(i) There exists terms s and t such that $\{\theta_i \mid 1 \le i \le n\}$ is compatible iff s and t are unifiable.

(ii) If λ is a mgu of s and t, then λ is a unifying composition of $\{\theta_i \mid 1 \le i \le n\}$, i.e. $\lambda \in \theta_1 \circledast \ldots \circledast \theta_n$.

Proof: (i) Let $\theta_i = \{x_1^i \mid t_1^i, \ldots, x_{m_i}^i \mid t_{m_i}^i\}$ for $1 \le i \le n$, let h be a new functionsymbol and

$$s = h(x_1^1 \ldots x_{m_1}^1 \quad x_1^2 \ldots x_{m_2}^2 \ldots x_1^n \ldots x_{m_n}^n)$$

$$t = h(t_1^1 \ldots t_{m_1}^1 \quad t_1^2 \ldots t_{m_2}^2 \ldots t_1^n \ldots t_{m_n}^n) \quad .$$

If $\{\theta_i \mid 1 \le i \le n\}$ is compatible, i.e. there exists a $\sigma \in \$$ with $\lambda = \sigma\theta_1 = \ldots = \sigma\theta_n$, then λ is a unifier of s and t; since for $1 \le i \le n$ and $1 \le j \le m_i$ we have

$$\lambda x_j^i = \sigma\theta_i x_j^i = \sigma\theta_i\theta_i x_j^i = \sigma\theta_i t_j^i = \lambda t_j^i \quad .$$

Let s and t be unifiable with $\sigma \in SUB$. We have to show

$$\sigma\theta_i = \sigma\theta_{i+1} \quad \text{for} \quad 1 \le i \le n-1$$

which is shown as in Lemma IV.2.

(ii) The second part of the lemma is also proved in the same way as Lemma IV.2.

□

Lemma IV.9: Let $\sigma,\tau \in \mathcal{S}$ and $s,t \in \mathfrak{T}$. If σ comp τ and $\sigma s = \sigma t$ then $\sigma * \tau s = \sigma * \tau t$ and τs and τt are unifiable.

Proof: Since σ comp τ, there exists $\lambda \in \mathcal{S}$ such that $\sigma * \tau \equiv \lambda\sigma = \lambda\tau$. Hence $\sigma * \tau = \rho\lambda\sigma$, where ρ is a renaming substitution, and $\sigma * \tau s = \rho\lambda\sigma s = \rho\lambda\sigma t = \sigma * \tau t$. Moreover it is $\rho\lambda\tau s = \rho\lambda\sigma s = \rho\lambda\sigma t = \rho\lambda\tau t$, i.e. τs and τt are unifiable.

Lemma IV.10: Let $\sigma,\tau \in \mathcal{S}$, then $\sigma \leq \tau$ iff σ comp τ and $\sigma * \tau \equiv \sigma$.

Proof: Since $\sigma \leq \tau$ and $\sigma \leq \sigma$, we get with Lemma IV.4 (i) that σ and τ are compatible and $\sigma \leq \sigma * \tau$. By Lemma IV.4 (ii) we have $\sigma * \tau \leq \sigma$ and therefore $\sigma * \tau \equiv \sigma$.

By Lemma IV.4 (ii) we have $\sigma \equiv \sigma * \tau \leq \tau$, which proves the other direction.

□

Corollary: (i) Let s and t be terms and let $\sigma \in \mathcal{S}$ such that $\sigma s = \sigma t$. If τ is an mgu of s and t, then

$$\sigma \equiv \sigma * \tau .$$

(ii) If $\sigma \subset \tau$ then $\sigma * \tau \equiv \tau$.

Proof: (i) Since τ is an mgu, we have $\sigma \leq \tau$ and by the above lemma $\sigma \equiv \sigma * \tau$.

(ii) Since $\sigma \subset \tau$ we have $\tau \leq \sigma$ and hence by the above lemma $\sigma * \tau \equiv \tau$.

□

Lemma IV.11: Let $\sigma,\tau \in \mathcal{S}$. If σ comp τ and $\sigma * \tau \subset \tau$ then $\sigma * \tau \equiv \tau$.

Proof: We have $\tau \leq \sigma * \tau$ and since we have $\sigma * \tau \leq \tau$ (by Lemma IV.4 (ii)), hence $\sigma * \tau \equiv \tau$.

□

Lemma IV.12: Let $\sigma,\tau \in \mathcal{S}$. If

(i) $DOM(\sigma) \cap DOM(\tau) = \emptyset$ and

(ii) $DOM(\tau) \cap VCOD(\sigma) = \emptyset$,

then σ comp τ and $\sigma * \tau \equiv \sigma\tau$.

Proof: With (ii) and Lemma I.8 we have $\sigma\tau \in \mathcal{S}$. We show that $\sigma\tau$ unifies σ and τ. First it is $\sigma\tau\tau = \sigma\tau$. If $x \in DOM(\sigma)$, then $\sigma\tau\sigma x \underset{(ii)}{=} \sigma\sigma x = \sigma x \underset{(i)}{=} \sigma\tau x$ and if $x \notin DOM(\sigma)$, then $\sigma\tau\sigma x = \sigma\tau x$, i.e. σ comp τ. Hence there exists $\lambda \in \mathcal{S}$ such that $\sigma * \tau \equiv \lambda\sigma = \lambda\tau$ and $\sigma\tau \le \lambda$ and with Lemma II.10 we have $\sigma\tau = \sigma\tau\tau \le \lambda\tau \equiv \sigma * \tau$. But by Lemma IV.4 (ii) we have $\sigma * \tau \le \sigma$ and again with Lemma II.10 $\sigma * \tau = \sigma * \tau\tau \le \sigma\tau$. Summarizing we get $\sigma * \tau \equiv \sigma\tau$.

Lemma IV.13: Let $\sigma,\tau \in \mathcal{S}$ and $s,t \in \mathfrak{T}$. If σ is an mgu of s and t and σ comp τ, then there exists a $\theta \in \mathcal{S}$ such that $\theta\tau \equiv \sigma * \tau$ and θ is an mgu of τs and τt.

Proof: By Lemma IV.9 τs and τt are unifiable and let $\theta \in \mathcal{S}$ be an mgu of τs and τt. Next we show $\theta\tau \in \mathcal{S}$: by the corollary of Lemma I.8 it is sufficient to show that $DOM(\tau) \cap VCOD(\theta) = \emptyset$. Let $x \in DOM(\tau) \cap VCOD(\theta)$ then by Lemma I.4 $x \notin var(\{\tau s,\tau t\})$ and hence, since θ is mgu, $x \notin var(\theta)$ which is a contradiction to $x \in VCOD(\theta)$. Now $\theta\tau \in \mathcal{S}$ is a unifier of s and t and therefore $\theta\tau \le \sigma$ and since $\theta\tau \le \tau$ using Lemma IV.4 (i) we have $\theta\tau \le \sigma * \tau$. Since $\sigma * \tau$ is a unifier of τs and τt it is $\sigma * \tau \le \theta$ and with Lemma II.10 $\sigma * \tau \le \theta\tau$. Summarizing we have $\sigma*\tau \equiv \theta\tau$.

□

Corollary: Let $\sigma,\tau,\theta \in \mathcal{S}$ and $s,t \in \mathfrak{T}$ be as in Lemma IV.13. If τ is a ground substitution, i.e. $VCOD(\tau) = \emptyset$, then $\sigma * \tau = \theta \sqcup \tau$, where θ is an mgu of τs and τt.

Proof: By Lemma IV.13 we know $\sigma * \tau \equiv \theta\tau$. In order to see $\theta\tau = \theta \sqcup \tau$ it is sufficient by Lemma I.17 to show (1) $DOM(\tau) \cap DOM(\theta) = \emptyset$. (2) $DOM(\theta) \cap VCOD(\tau) = \emptyset$ and (3) $DOM(\tau) \cap VCOD(\theta) = \emptyset$.

(1) Suppose by contradiction there exists $x \in DOM(\tau) \cap DOM(\theta)$ then $x \notin var(\tau s,\tau t)$ by Lemma I.4 and hence since θ is mgu of

τs and τt, $x \notin var(\theta)$ which is a contradiction to $x \in DOM(\theta)$.

(2) $DOM(\theta) \cap VCOD(\tau) = \emptyset$, since $VCOD(\tau) = \emptyset$.

(3) $DOM(\tau) \cap VCOD(\theta) = \emptyset$ as in the proof of the above lemma.

□

Lemma IV.14: Let $\sigma, \tau \in \mathcal{S}$ and $s, t \in \mathfrak{T}$. If σ is an mgu of s and t and θ is a mgu of τs and τt then σ comp τ and $\sigma * \tau \equiv \theta\tau$.

Proof: Since $\theta\tau$ unifies s and t, we have $\theta\tau \le \sigma$ and by definition $\theta\tau \le \tau$. With lemma IV.4 (i) σ comp τ and $\theta\tau \le \sigma * \tau$.

Let $\lambda \in \mathcal{S}$ such that $\sigma * \tau \equiv \lambda\sigma \equiv \lambda\tau$. Then λ is a unifier of τs and τt and therefore $\lambda \le \theta$ and by Lemma II.10 $\lambda\tau \le \theta\tau$, i.e. $\sigma * \tau \le \theta\tau$. Hence $\sigma * \tau \equiv \theta\tau$.

□

Summarizing Lemma IV.13 and Lemma IV.14 we get

Proposition IV.1: Let $\sigma, \tau \in \mathcal{S}$, $s, t \in \mathfrak{T}$ and let σ be an mgu of s and t.

(i) σ comp τ iff τs and τt are unifiable.

(ii) If σ comp τ or τs and τt are unifiable, then $\sigma * \tau \equiv \theta\tau$ where θ is an mgu of τs and τt.

The following corollary is a specification of the above proposition and was used in the proof of Lemma III.10.

Corollary: Let $\sigma, \tau \in \mathcal{S}$ and $s, t \in \mathfrak{T}$, let σ be an mgu of s and t and $DOM(\tau) \cap var(t) = \emptyset$. Then τs and t are unifiable iff σ comp τ.

Proof: Since $\tau t = t$ the proof is trivial.

□

V. Conclusion

During the preparation of this report [Ed83] was published, which shows that the set of equivalence classes of idempotent substitutions together with an added greatest element is a complete lattice. The definition of a supremum of two classes of idempotent substitutions is equivalent to our definition of a unifying composition of two substitutions (cf Lemma IV.4). The concept of weak unification introduced there is equivalent to the concept of R-unification.

Finally I would like to emphasize that the purpose of this report is not so much in providing new results but it should serve as a reference which collects some basic notions of first-order unification theory.

References

[BA73] Baxter, L.D., An Efficient Unification Algorithm, University of Waterloo, Techn. Rep. CS-73-23, 1973

[BES81] Bläsius, K.-H., Eisinger, N., Siekmann, J., Smolka, G., Herold, A., and Walther, C., The Markgraf Karl Refutation Procedure, Proc. of the 7th International Joint Conference on Artificial Intelligence (1981)

[Ch72] Chang, C.L., Theorem Proving with Variable-Constrained Resolution, Information Sci. 4(1972), 217-231

[CL73] Chang, C.L. and Lee, R.C.T., Symbolic Logic and Mechanical Theorem Proving (Academic Press, New York, 1973)

[CS79] Chang, C.L. and Slagle, J.R., Using Rewriting Rules for Connection Graphs to Prove Theorems, Artificial Intelligence 12(1979) 159-180

[Ed83] Eder E., Properties of Substitutions and Unifications, Institut für Informatik, Universität München, ATP-17-II-83

[Hu76] Huet, G., Resolution D'Equations dans des Langages D'Ordre 1,2,...,ω (Thèse d'Etat), Université Paris VII, 1976

[Hu80] Huet, G., Confluent Reductions: Abstract Properties and Applications to Term Rewriting Systems, JACM, vol. 27, No. 4, 1980

[KB70] Knuth, D., Bendix, P., Simple Word Problems in Universal Algebras, in: Comp. Problems in Abstract Algebra, J. Leech (ed.) Pergamon Press, 1970

[Lo78] Loveland, D., Automated Theorem Proving, North Holland, 1978

[MM79] Martelli, H., Montaneri, U., An Efficient Unification Algorithm, University of Pisa, Techn. Report, 1979

[Ni80] Nilsson, N., Principles of Artificial Intelligence, Tioga Publ. Comp., Cal., 1980

[Oh82] Ohlbach, H.J., The Markgraf Karl Refutation Procedure: The Logic Engine, Interner Bericht 24/82, Institut für Informatik I, Universität Karlsruhe (1982)

[Pl70] Plotkin, G.D., Lattice-theoretic Properties of Subsumption, Memorandum MIP-R-77, University of Edinburgh

[PS81] Peterson, G., Stickel, M., Complete Sets of Reductions for Equational Theories with Complete Unification Algorithms, JACM, vol. 28, No. 2, 1981

[PW78] Paterson, M., Wegman, M., Linear Unification, Journal of Comp. and Syst., 16, 1978

[Re70] Reynolds, J., Transformational Systems and the Algebraic Structure of Atomic Formulas, Machine Intelligence 5, pp. 135-152, American Elsevier, New York, 1970

[Ro65] Robinson, J.A., A Machine Oriented Logic based on the Resolution Principle, JACM 12, 1965

[Ro71] Robinson, J.A., Computational Logic: The Unification Computation, Machine Intelligence, vol. 6, 1971

[Ros73] Rosen, B.K., Tree-Manipulating Systems and Church-Rosser Theorems, JACM vol. 20, 1973

[Si76] Sickel, S., A Search Technique for Clause Interconnectivity Graphs, IEEE Trans. on Computers C-25(8), 823-835, 1976

[SS81] Siekmann, J., Szabó, P., Universal Unification and Regular ACF-Theories, Proceedings of the 7th International Joint Conference on Artificial Intelligence (1981)

[SS82] Siekmann, J., Szabó, P., Universal Unification and a Classification of Equational Theories, Proceedings of the 6th Conference on Automated Deduction, Springer LNCS 138, 1982

[Va75] van Vaalen, J., An Extension of Unification to Substitutions with an Application to ATP, Proc. of the Fourth International Joint Conference on Artificial Intelligence (1975)

[Wa82] Walther, C., A Many-Sorted Calculus Based on Resolution and Paramodulation, Interner Bericht 34/82, Institut für Informatik I, Universität Karlsruhe (1982)