Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

**Research
Report**
RR-08-01

String Unification is Essentially Infinitary

Michael Hoche

Jörg Siekmann

Peter Szabo

July 2008

**Deutsches Forschungszentrum für Künstliche Intelligenz
GmbH**

Postfach 2080
D-67608 Kaiserslautern
Tel: +49 (631) 205-75 0
Fax: +49 (631) 205-75 503

Stuhlsatzenhausweg 3
D-66123 Saarbrücken
Tel: +49 (681) 302-5151
Fax: +49 (681) 302-5341

Robert-Hooke-Str. 6
D-28359 Bremen
Tel.: +49 (421) 218-64 100
Fax.: +49 (421) 218-64 150

E-Mail: info@dfki.de WWW: http://www.dfki.de

Deutsches Forschungszentrum für Künstliche Intelligenz
DFKI GmbH
German Research Center for Artificial Intelligence

Founded in 1988, DFKI today is one of the largest nonprofit contract research institutes in the field of innovative software technology based on Artificial Intelligence (AI) methods. DFKI is focusing on the complete cycle of innovation - from world-class basic research and technology development through leading-edge demonstrators and prototypes to product functions and commercialization.

Based in Kaiserslautern, Saarbrücken and Bremen, the German Research Center for Artificial Intelligence ranks among the important „Centers of Excellence“ worldwide.

An important element of DFKI's mission is to move innovations as quickly as possible from the lab into the marketplace. Only by maintaining research projects at the forefront of science DFKI has the strength to meet its technology transfer goals.

The key directors of DFKI are Prof. Wolfgang Wahlster (CEO) and Dr. Walter Olthoff (CFO).

DFKI's research departments are directed by internationally recognized research scientists:

- Image Understanding and Pattern Recognition (Prof. T. Breuel)
- Knowledge Management (Prof. A. Dengel)
- Deduction and Multiagent Systems (Prof. J. Siekmann)
- Language Technology (Prof. H. Uszkoreit)
- Intelligent User Interfaces (Prof. W. Wahlster)
- Institute for Information Systems at DFKI (Prof. Dr. P. Loos)
- Robotics (Prof. F. Kirchner.)
- Safe and Secure Cognitive Systems (Prof. B. Krieg-Brückner)

and the associated Center for Human Machine Interaction (Prof. Dr.-Ing. Detlef Zühlke)

In this series, DFKI publishes research reports, technical memos, documents (eg. workshop proceedings), and final project reports. The aim is to make new results, ideas, and software available as quickly as possible.

Prof. Wolfgang Wahlster
Director

String Unification is Essentially Infinitary

Michael Hoche, Jörg Siekmann, Peter Szabo

DFKI-RR-08-01

ISSN 0946-008X

String Unification is Essentially Infinitary

Michael Hoche
EADS Deutschland GmbH, HMSI e.V.
Normannenweg 48, D-88090 Immenstaad

Jörg Siekmann,
DFKI
Stuhlsatzenweg 3, D-66123 Saarbrücken

Peter Szabo
HMSI e.V.
Kurt-Schumacherstr. 13, D-75180 Pforzheim

July 2, 2008

Abstract

A *unifier* of two terms s and t is a substitution σ such that $s\sigma = t\sigma$ and for first-order terms there exists a *most general unifier* σ in the sense that any other unifier δ can be composed from σ with some substitution λ, i.e. $\delta = \sigma \circ \lambda$.

This notion can be generalised to E-unification, where E is an equational theory, $=_E$ is equality under E and σ is an E-unifier if $s\sigma =_E t\sigma$. Depending on the equational theory E, the set of most general unifiers is always a singleton (as above), or it may have more than one, either finitely or infinitely many unifiers and for some theories it may not even exist, in which case we call the theory of type nullary.

String unification (or Löb's problem, Markov's problem, unification of word equations or Makanin's problem as it is often called in the literature) is the E-unification problem, where $E = \{f(x, f(y,z)) = f(f(x,y),z)$, i.e. unification under associativity or string unification once we drop the fs and the brackets. It is well known that this problem is infinitary and decidable.

Essential unifiers, as introduced by Hoche and Szabo, generalise the notion of a most general unifier and have a dramatically pleasant effect on the set of most general unifiers: the set of essential unifiers is often much smaller than the set of most general unifiers. Essential unification may even reduce an infinitary theory to an essentially finitary theory. The most dramatic reduction known so far is obtained for idempotent semigroups or bands as they are called in computer science: bands are of type nullary, i.e. there exist two unifiable terms s and t, but the set of most general unifiers is not enumerable. This is in stark contrast to essential unification: the set of essential unifiers for bands always exists and is finite.

We show in this paper that the early hope for a similar reduction of unification under associativity is not justified: string unification is *essentially infinitary*. But we give an enumeration algorithm for essential unifiers. And beyond, this algorithm terminates when the considered problem is finitary.

Keywords: E-unification, equational theory, essential unifiers, string unification, unification of words, universal algebra, varieties.

1 Introduction

Unification is a well established concept in artificial intelligence, automated theorem proving, computational linguistics, universal algebra, in theoretical and applied computer science, and e.g. semantics of programming languages. Surveys of unification theory can be found in [18, 6, 7]. A survey of the related topic of rewriting systems is presented in [9] and more recently in [12]. A standard textbook is Franz Baader, Tobias Nipkow, *Term Rewriting and All That* [6].

Unification is a general concept to solve equational problems, which is especially embedded in a plurality of deduction and inference mechanisms. For practical applications it is often crucial to have a finite or at least minimal representation of all the solutions, i.e. a minimal complete set of unifiers from which all other solutions (unifiers) can be derived.

For equational problems in the free algebra of terms (also known as syntactic unification), there exists always a unique unifier for solvable unification problems from which all other unifieres can be derived by instantiation. This unique unifier is called the *most general unifier*, [14]. For equational algebras however the situation is completely different: a minimal complete set of unifiers is not always finite and it may not even exist, which was conjectured by Gordon Plotkin in his seminal paper in 1972, [13]. Since then unification problems and equational theories have been classified with respect to the cardinality of their minimal complete set of unifiers. These results led to the development of general approaches and algorithms, which can be applied to a whole class of theories. This is the topic of *universal unification*, see e.g. [18].

More specifically, a unification problem $s =^?_E t$ for two given terms s and t under an equational theory E is the problem to find a minimal and complete set of unifiers $\mu\mathcal{U}\Sigma_E$ for s and t such that for every unifier $\sigma \in \mu\mathcal{U}\Sigma_E$ we have $s\sigma =_E t\sigma$. We say a unification problem is *unitary* if $\mu\mathcal{U}\Sigma_E$ is always a singleton, it is *finitary* if $\mu\mathcal{U}\Sigma_E$ is finite for every s and t and it is *infinitary* if there are terms s and t such that $\mu\mathcal{U}\Sigma_E$ is infinite. Unfortunately there are theories such that two terms are unifiable,but the set $\mu\mathcal{U}\Sigma_E$ is not recursively enumerable. In this case we call the problem *nullary* or of type zero.

It turned out that this well established view of unification theory changes drastically if we redefine the notion of a most general unifier. Recall that a unifier σ is most general if for any other unifier τ there exists a substitution λ

such that

$$\tau = \sigma \circ \lambda$$

We generalise this notion and define an *essential unifier* σ if for any other unifier τ there exist substitutions λ_1 and λ_2 such that

$$\tau = \lambda_1 \circ \sigma \circ \lambda_2$$

where λ_1 has to have certain properties to be defined below.

We say a unification problem is e-unitary (is e-finitary) if the set of essential unifiers is always a singleton (is always finite). A unification problem is e-infinitary (e-nullary) if there are two terms such that the set of essential unifiers is infinite (is not recursive enumerable).

These notions were first introduced by Hoche and Szabo in [5] and it was shown in their paper that the unification problem for idempotent semigroups (bands) is e-finitary. Bands are well known since it was one of the early examples to demonstrate Plotkin's conjecture, that there exist nullary equational theories, which was shown one and a half decades later by Manfred Schmidt-Schauss,[15]. Now the unification problem for bands is nullary in the traditional sense but it is e-finitary in our sense: this is so far the most drastic reduction of the cardinality of the set of most general unifiers to a set of essential unifiers.

The question is: can similar results be obtained for other theories as well and a natural candidate for this kind of investigation is string unification. Why is that?

In the 1950s A. A. Markov was interested in the solvability of word equations in free semigroups: he noted that every word equation over a two constant alphabet can be translated into a set of diophantine equations. Using this translation he hoped to find a proof for the unsolvability of Hilbert's tenth problem by showing that the solvability of word equations is undecidable. This put the problem firmly on the map and others joined in: see the volumes edited by M. Lothaire and others on *Combinatorics on Words* [2]. The problem was finally solved in the affirmative in the seminal work by G. S. Makanin. An excellent exposition of Makanin's algorithm (with several improvements) is presented by Klaus Schulz [3] and by Volker Diekert (Chapter 12 in [2]).

Apart from its theoretical interest, the problem became more widely known, because of its relevance in computer science, artificial intelligence and automated reasoning. As opposed to the above works on decidability which just enumerate all solutions and make the decidability of the existence of a solution their primary focus, we are interested in the latter works, inspired by automated theorem proving, where the set $\mu\mathcal{U}\Sigma$ of the *most general* solutions is the focus of attention.

The most common and simple example to show that string unification is infinitary is the following

$$(1) \qquad xa = ax$$

with the set of most general unifiers

$$\mu\mathcal{U}\Sigma = \{\{x \mapsto a\}, \{x \mapsto aa\}, \{x \mapsto aaa\}, \ldots\}.$$

It is easy to show that indeed this is a solution set and it is not as immediate,but still not too hard to show that there does not exist any other more general set of unifiers $\mu\mathcal{U}\Sigma$ for this problem. Finally $\mu\mathcal{U}\Sigma$ is minimal, which again is obvious, as there are no variables in the a^n and thus they do not yield to instantiation. Hence in general

string unification is infinitary.

As we have said, this is a well known fact since the mid seventies and it is probably the most often quoted example in any lecture or monograph on unification theory.

A similar example

$$(2) \qquad xa = bx$$

is usually chosen to demonstrate that the naive string unification algorithm is not a decision procedure: although it is obvious that the above example is not unifiable, the actual algorithm would run forever.

However, problem (1) has a finite set (in fact an even e-unitary set) of essential unifiers

$$e\mathcal{U}\Sigma = \{\{x \mapsto a\}\} = \{\sigma_1\}$$

and any other unifier can be obtained with $\lambda_1 = \{x \mapsto a^{n-1}x\}, n > 0$ and $\lambda_2 = \varepsilon$. In other words, for any unifier $\sigma_n = \{x \mapsto a^n\}$:

$$\begin{array}{rcl} \sigma_n & = & \lambda_1\sigma_1\lambda_2 \\ & = & \{x \mapsto a^{n-1}x\} \circ \sigma_1 \circ \varepsilon \\ & = & \{x \mapsto a^{n-1}x\} \circ \{x \mapsto a\} \circ \varepsilon \\ & = & \{x \mapsto a^n\} \end{array}$$

where λ_1 obeys a certain structural property, to be defined in the next section.

Once this observation had been made a few years ago, there was an intense struggle to generalise this observation to any string unification problem and to prove the conjecture

string unification is e-finitary.

As we shall show in this paper, this conjecture is false in general, albeit it holds for certain subclasses of strings.

2 Basic Notions and Notation

Notation and basic definitions in unification theory are well known and have found their way into many and diverse research areas, standard survey articles are [18, 6, 7] and the monographs and textbooks on automated theorem proving usually contain sections on unification. Most recent results are presented at the Unification Workshop.[1]

[1] First workshop in Val d'Ajol in 1987 and since then annually. Since 1997, there is a website UNIF'987, UNIF'98, UNIF'99 up to UNIF'05 in Japan and UNIF'06 at the FLOC conference in Seattle.

For the reader's convenience we present some of the standard notation below, followed by the definitions of our novel approach for essential unifiers.

2.1 Unification theory: common definitions

An alphabet $\mathcal{F} = (F_n)_{n \in \mathcal{N}}$ provides a vocabulary, where the function symbols F_i, $i \in \mathcal{N}$ have the *arity* i. Function symbols with arity 0 are called *constants.* A set X gives us a denumerable set of *variable* symbols, usually denoted as x, y, z etc. and $\mathcal{F}$ and X constitute Σ, the *signature* of a term algebra.

The set of (first-order) terms $\mathcal{T}_{F,X}$ over a signature Σ generated by the variables X, is the smallest set containing the variables $x \in X$, and the terms $f(t_1, \ldots, t_n)$, whenever $f \in F_n$ is a function symbol of arity n and $t_1, \ldots, t_n \in \mathcal{T}_{F,X}$ are terms. The set of terms is a *(free) term algebra.*

The set of variable-free terms are called *ground terms.* The set of *variables* occurring in a term t is denoted by $\mathbf{Var}(t)$ and the set of *symbols* of $\mathcal{F}$ occurring in t is denoted by $\mathbf{Sym}(t)$. For a term t the set of *sub-terms* $\mathbf{Sub}(t)$ contains $t \in \mathbf{Sub}(t)$ itself and is closed recursively by containing $t_1, \ldots, t_n \in \mathbf{Sub}(t)$, if $f(t_1, \ldots, t_n) \in \mathbf{Sub}(t)$. For a set of terms $T = \{t_1, t_2, \ldots, t_n\}$ the subterms are defined by $\mathbf{Sub}(T) = \mathbf{Sub}(t_1) \cup \ldots \cup \mathbf{Sub}(t_n)$.

A *substitution* is the (unique) homomorphism in the term algebra generated by a mapping $\sigma : X \longrightarrow \mathcal{T}_{F,X}$ from a finite set of variables to terms. Substitutions are generally denoted by small Greek letters $\alpha, \beta, \gamma, \sigma$ etc. A substitution σ is represented explicitly as a function by a set of *variable bindings* $\sigma = \{x_1 \mapsto s_1, \ldots, x_m \mapsto s_m\}$. The application of the substitution σ to a term t, denoted $t\sigma$, is defined by induction on the structure of terms

$$t\sigma = \begin{cases} s_i & \text{if } t = x_i \\ f(t_1\sigma, \ldots, t_n\sigma) & \text{if } t = f(t_1, \ldots, t_n) \end{cases}$$

The substitution $\varepsilon = \{\}$ with $t\varepsilon = t$ for all terms t in $\mathcal{T}_{F,X}$ is called the *identity.* A substitution $\sigma = \{x_1 \mapsto s_1, \ldots, x_m \mapsto s_m\}$ has the *domain*

$$\mathbf{Dom}(\sigma) := \{x | x\sigma \neq x\} = \{x_1, \ldots, x_m\};$$

and the *range* is the set of terms

$$\mathbf{Ran}(\sigma) := \bigcup_{x \in \mathbf{Dom}(\sigma)} \{x\sigma\} = \{s_1, \ldots, s_m\};$$

The set of variables occurring in the range is $\mathbf{VRan}(\sigma) := \mathbf{Var}(\mathbf{Ran}(\sigma))$ and $\mathbf{Var}(\sigma) = \mathbf{Dom}(\sigma) \cup \mathbf{VRan}(\sigma)$; the *restriction* of a substitution σ to a set of variables $Y \subseteq X$, denoted by $\sigma_{|Y}$, is the substitution which is equal to the identity everywhere except over $Y \cap \mathbf{Dom}(\sigma)$, where it is equal to σ.

Relations such as $=, \geq, \ldots$ between substitutions sometimes hold only if restricted to a certain set of variables V. A relation R which is restricted to V is denoted as R^V, and defined as $\sigma\ R^V\ \tau \iff \sigma_{|V}\ R\ \tau_{|V}$.

The *composition* of two substitutions σ and θ is written $\sigma \circ \theta$ (emphases the composition) or just $\sigma\theta$ and is defined by $t\sigma\theta = (t\sigma)\theta$.

Two substitutions σ and θ are *equal*, denoted $\sigma = \theta$ iff $x\sigma = x\theta$ for every variable x.

A term t is an *instance* of a term s denoted $s \leq t$, if $t = s\sigma$ for some substitution σ, i.e.

$$s \leq t \Leftrightarrow \exists \sigma : s\sigma = t.$$

We also say s is more general or less specific than t. The relation $\leq$ is a quasi-ordering on terms called the *subsumption ordering*, whose associated equivalence relation and strict ordering are called subsumption equivalence and strict subsumption, respectively.

The *encompassment ordering* or *containment ordering* [4] is defined as the subterm ordering composed with the subsumption ordering, i.e. a subterm of t is an instance of s

$$s \sqsubseteq t \iff \exists \sigma : s\sigma \in \mathrm{Sub}(t).$$

Encompassment conveys the notion that s "appears" in t with a context "above" and a substitution "below". We say t *encompasses* s or s *is part of* t.

A substitution θ is called *more general* than σ, denoted $\theta \leq \sigma$, if there exists a λ such that $\sigma = \theta\lambda$, i.e.

$$\theta \leq \sigma \iff \exists \lambda : \theta\lambda = \sigma.$$

The relation $\leq$ is a pre-order, called the *instantiation* ordering for substitutions.

An *equation* or *identity* $s = t$ in a term algebra $\mathcal{T}_{F,X}$ is a pair (s,t) of terms and an algebra A satisfies the equation $s = t$ if for every homomorphism

$$h : \mathcal{T}_{F,X} \longrightarrow A,$$

$h(s) = h(t)$ that is, only if (s,t) is in the kernel of every homomorphism from $\mathcal{T}_{F,X}$ to A.

An *equational theory* is defined by a set of identities $E \subseteq \mathcal{T}_{F,X} \times \mathcal{T}_{F,X}$. It is the least congruence on the term algebra which is closed under substitution and contains E, and will be denoted by $=_E$. If $s =_E t$ we say s and t are *equal modulo E*. The sets $[\mathbf{s}]_{\mathbf{E}} = \{t | t =_E s\}$ are called congruence classes or *equivalence classes* (modulo E).

Let E be an equational theory and Σ the signature of the underlying term algebra. An *E-unification problem* (over Σ) is a finite set of equations

$$\Gamma = \{s_1 \overset{?}{=}_E t_1, \ldots, s_n \overset{?}{=}_E t_n\}$$

between Σ-terms with variables in a (countably infinite) set of variables V.

An *E-unifier of* Γ is a substitution σ, such that

$$s_1\sigma =_E t_1\sigma, \ldots, s_n\sigma =_E t_n\sigma.$$

The set of all E-unifiers of Γ is denoted by $\mathcal{U}\Sigma_E(\Gamma)$ or if the signature Σ is known from the context, we just write $\mathcal{U}_E(\Gamma)$ or even $\mathcal{U}(\Gamma)$.

A *complete set of E-unifiers* of Γ is a set C of substitutions, such that

(1) $C \subseteq \mathcal{U}\Sigma_E(\Gamma)$, i.e. each element of C is an E-unifier of Γ relative to a signature Σ and

(2) for each $\theta \in \mathcal{U}\Sigma_E(\Gamma)$ there exists $\sigma \in C$ with $\sigma \leq_E \theta$.

The set $\mu\mathcal{U}\Sigma_E(\Gamma)$ is a *minimal complete set of E-unifiers* for Γ, if it is a complete set, i.e. $\mu\mathcal{U}\Sigma_E \subseteq C$, and every two distinct elements of $\mu\mathcal{U}\Sigma_E$ are incomparable, i.e., $\sigma \leq_E \sigma'$ implies $\sigma =_E \sigma'$ for all $\sigma, \sigma' \in \mu\mathcal{U}\Sigma_E$. When a minimal complete set of E-unifiers of a unification problem Γ exists, it is unique up to subsumption equivalence.

The empty or unit substitution ε is a unifier in case $s =_E t$ are already equal. Minimal complete sets of unifiers need not always exist, and if they do, they might be singular, finite, or infinite. Since minimal complete sets of E-unifiers are isomorphic whenever they exist, theories can be classified with respect to their corresponding unification problem.

This leads naturally to the concept of a *unification hierarchy* which was first introduced in Siekmann's Ph.D. Thesis in 1975 [16], and further refined and extended by himself and his students, see [18, 6, 7] for surveys.

A *unification problem* Γ is *nullary*, if Γ does not have a minimal complete set of E-unifiers. The unification problem Γ is *unitary*, if it is not nullary and the minimal complete set of E-unifiers is of cardinality less or equal to 1. The unification problem Γ is *finitary*, if it is not nullary and the minimal complete set of E-unifiers is of finite cardinality. The unification problem Γ is *infinitary*, if it is not nullary and the minimal complete set of E-unifiers is of infinite cardinality.

An *equational theory* E is *unitary*, if all unification problems are unitary. An equational theory E is *finitary*, if all unification problems are finitary. An equational theory E is *infinitary*, if there is at least an infinitary unification problem and all unification problems have minimal complete sets of E-unifiers. If there exists a unification problem Γ not having a minimal complete set of E-unifiers, then the equational theory is *nullary* of *type zero*.

2.2 Additional Definitions: Essential Unifiers

Substitutions form a semigroup with respect to their composition. This fact was used to define the instantiation order on unifiers from above, namely

$$\sigma \leq \tau \iff \exists \lambda : \sigma \circ \lambda = \tau,$$

which led to the notion of a most general unifier.
As indicated above this concept does not generalise well on equational theories: the equational theory of associativity $A = \{x(yz) = (xy)z)\}$, i.e. the free semigroup with the unification problem $\{ax =^?_A xa\}$ has the infinite set of most general unifiers $\{\{x \mapsto a^n\} | n \geq 1\}$, as discussed in the introduction.
However, the essentially unifier in this set intuitively seems to be $\{x \mapsto a\}$, because every most general unifier contains this unifier in a certain sense, namely:

$$\{x \mapsto a^n\} = \{x \mapsto a^{n-1}x\} \circ \{x \mapsto a\}.$$

Now having in mind that substitutions form a semigroup, the dual of the instantiation ordering, i.e. left-composition instead of right-composition seems to change the infinitary problem into a finitary one if we redefine the order $\leq$ into $\exists\lambda : \sigma = \lambda\tau$, where $\sigma = \{x \rightarrow a\}$. But this is not compatible with the original notion of generality and it would not quite work in general.

Our solution to this dilemma is based on a lifting of the encompassment order on terms to an encompassment order on substitutions. More specifically we define a tripartition of a substitution i.e. an ordering concept which involves both left composition and right composition:

$$\sigma \trianglelefteq \tau \iff \exists\alpha\exists\beta : \tau = \alpha\sigma\beta.$$

And we say σ *is part of* τ and τ encompasses σ. This ordering concept, called *part ordering* in the following, is the result of lifting the encompassment order on terms and on substitutions, and it can be used to generate all unifiers as well. A unifier like σ above will be called an *essential unifier* if there is no left and right composition for σ and we shall now summarise the formalism to define the concept of an *essential unifier* (see [5] for more details). We say a substitution σ is part of a substitution τ, if and only if the domain of σ is a subset of the domain of τ and there exist α and β that "*build up*" σ into τ by means of composition, i. e. $\tau = \alpha\sigma\beta$, and σ has actually "*contributed* " in this decomposition of τ. The actual "*contribution*" of σ is important, since otherwise we would just end up again with the classical notion of a most general unifier. Technically this can be captured by the requirements, that the domain of σ and α are subsets of the domain of τ and whenever a variable x is in the domains of σ and α, then it is a variable in the range of α.

As usual this part relationship is generalized to equational theories E by considering all relationships modulo E, and we say: σ is a *part-substitution* of τ *modulo* E. This lifts the encompassment relation on basic terms to substitutions.

Part ordering on substitutions $\trianglelefteq_E$ is technically defined as follows:

Definition 2.1 (Part ordering of substitutions) *For substitutions σ and τ with $V = \mathbf{Var}(\tau)$*

(1) σ is **part of** *τ modulo E denoted as $\sigma \trianglelefteq_E \tau$, if there are two substitutions α and β with $\tau =_E^V \alpha\sigma\beta$, where $\mathbf{Dom}(\sigma) \cup \mathbf{Dom}(\alpha) \subseteq \mathbf{Dom}(\tau)$ and $\mathbf{Dom}(\sigma) \cap \mathbf{Dom}(\alpha) \subseteq \mathbf{VRang}(\alpha)$. In other words a substitution σ is part of a substitution τ if there is an instance of σ, namely $(\sigma\beta)$ which is a contributing (right) factor of τ, i. e. $\tau = \alpha(\sigma\beta)$.*[2]

(2) σ is **proper part** *of τ modulo E: $\sigma \triangleleft_E \tau$, if $\sigma \trianglelefteq_E \tau$ where the above two substitutions α and β with $\tau =_E^V \alpha\sigma\beta$ imply that $\alpha\beta \neq_E \iota$, where $\iota = \emptyset$.*

(3) σ and τ are **part equivalent** *modulo E: $\sigma \equiv_E \tau$ by $\sigma \trianglelefteq_E \tau \wedge \tau \trianglelefteq_E \sigma$*

(4) σ is **not part of** *τ modulo E: $\sigma \not\trianglelefteq_E \tau$ by $\neg(\sigma \trianglelefteq_E \tau)$.*

[2] Note the analogy to the encompassment order: "a term s is part of term t if there is an instance of s, namely $(s\sigma)$, which is a subterm of t."

(5) σ and τ are **part extrinsic** *modulo E: $\sigma \bowtie_E \tau$ by $\sigma \not\trianglelefteq_E \tau$ and $\tau \not\trianglelefteq_E \sigma$.*

Note that the contribution constraint assures that each term in the range of σ actually contributes to τ. This makes sure that we do not have unnecessary components like $x \mapsto f(a)$ which are just absorbed, as illustrated by the following example

$$\tau = \{x \mapsto a, y \mapsto b\} = \{x \mapsto a\}\{x \mapsto f(a)\}\{y \mapsto b\},$$

where $\{x \mapsto f(a)\}$ is obviously not a part of τ and absorbed by $\{x \mapsto a\}$. A substitution is a proper part if each part decomposition implies that the framing factors actually contribute something. As an illustration of the above definition consider the following example:

$$\tau = \{x \mapsto f(g(z), h(z)), y \mapsto g(z), z \mapsto j(x), v \mapsto k(y, z)\}$$

has a part $\sigma = \{y \mapsto g(z), z \mapsto j(y)\}$, since there is a left factor $\alpha = \{x \mapsto f(y, u)\}$ and a right factor $\beta = \{u \mapsto h(z), y \mapsto x, v \mapsto k(y, z)\}$. Obviously σ is a part of τ, since it contributes with $y \mapsto g(z)$ to α and it even contributes directly with $z \mapsto j(y)$ to τ, since z is not in the domain of α. Finally β completes the decomposition.

Proposition 2.2 *The part substitution ordering is indeed a pre-order, i.e. reflexive and transitive.*

The concept of an essential unifier can now easily be defined as:

Definition 2.3 *A unifier is* **essential** *if and only if it is minimal with respect to the part ordering, i.e.* **An essential e-unifier has no E-unifying part-substitution**.

The set of essential unifiers is indeed a generating set for all unifiers just as the traditional set of most general unifiers. This can be shown with the existence of a corresponding closure operator. A set of unifiers $C(\Gamma)$ is *e-complete* if for each unifier σ there exists a unifier τ in C which is part of σ. A complete set of unifiers $C(\Gamma)$ is *e-minimal* if any two distinct elements are not part of each other. Such a set is denoted as $e\mathcal{U}\Sigma_E(\Gamma)$ This set exists and is unique, because if there exist two complete sets of essential unifiers $e\mathcal{U}\Sigma_E^1$ and $e\mathcal{U}\Sigma_E^2$ with τ in $e\mathcal{U}\Sigma_E^1 \backslash e\mathcal{U}\Sigma_E^2$ and σ in $e\mathcal{U}\Sigma_E^2 \backslash e\mathcal{U}\Sigma_E^1$ then since $e\mathcal{U}\Sigma_E^1$ is complete, there exist the substitutions α and β for τ such that $\sigma =_E^V \alpha\tau\beta$. Since $e\mathcal{U}\Sigma_E^2$ is a set of essentials it follows $\sigma =_E \tau$, contradicting the assumption.

Lemma 2.4 *Let E be an equational theory and Γ a unification problem. Then the set of essential unifiers $e\mathcal{U}\Sigma_E(\Gamma)$ is a generating set for the set of all unifiers $\mathcal{U}\Sigma_E(\Gamma)$.*

A proof can be found in [5].

Lemma 2.5 $e\mathcal{U}\Sigma_E(\Gamma) \subseteq \mu\mathcal{U}\Sigma_E(\Gamma)$

The interesting observation is that the above subset of essential unifiers can be extremely small in comparison to its superset, as we shall see in the following.

3 Essential String Unification

We are interested now in the A-unification problem, i.e. unification in the free semigroup, where

$$A = \{f(x, f(y, z)) = f(f(x, y), z)\}$$

and the set of terms are built up as usual over constants, variables, but only one function symbol f. In this case, we can just drop the fs and brackets and write strings (or words) over the alphabet of constants and variables. A set of string equations will be denoted as $\Gamma = \{u_1 = v_1, \ldots, u_n = v_n\}$ and $\mathbf{Var}(\Gamma)$ is the set of free variable symbols occurring in u_i and v_i. Let $V = \mathbf{Var}(\Gamma)$, then a *(string-) unifier* $\sigma : V \mapsto \Sigma^*$ is a solution for Γ if $u_i\sigma = v_i\sigma, 1 \leq i \leq n$. The set of all unifiers is denoted as $\mathcal{U}(\Gamma)$. A unifier σ is *ground* if its range contains only constants and no variables. Now let us look at a few motivating examples, which show that indeed an infinite set of most general unifiers $\mu\mathcal{U}\Sigma$ collapses to a finite set of essential unifiers $e\mathcal{U}\Sigma$, supporting the hypothesis that the infinitary string unification problem is essentially finitary (which is false in general, as we shall see below).

Our first example is the well known string unification problem mentioned in the introduction:

$$ax =^? xa \text{ with } \sigma_n = \{x \mapsto a^n\}, n > 0$$

has infinitely many most general unifiers, but there is just *one* e-unifier $\sigma_0 = \{x \mapsto a\}$ because of

$$\sigma_n = \{x \mapsto a^{n-1}x\} \circ \sigma_0.$$

The next example has two variables[3]

$$xy =^? yx$$

and has infinitely many most general unifiers

$$\sigma_{i,j} = \{x \mapsto z^i, y \mapsto z^j\}, i, j > 0, \text{ where } i \text{ and } j \text{ are relative prime,}$$

but it has only one e–unifier $\sigma_0 = \{x \mapsto z, y \mapsto z\}$ because of

$$\sigma_{i,j} = \{x \mapsto z^{i-1}x, y \mapsto z^{j-1}y\} \circ \sigma_0$$

Our next example is taken from J. Karhumäki *Combinatorics of Words.* The system

$$\left\{ \begin{array}{l} xaba =^? baby \\ abax =^? ybab \end{array} \right\}$$

[3] see `http://www.math.uwaterloo.ca/~snburris/htdocs/scav/e_unif/e_unif.html`, example 15

has infinitely many most general unifiers

$$\sigma_n = \{x \mapsto b(ab)^n, y \mapsto (ab)^n a\}, n \geq 0$$

But it has only one e-unifier, namely σ_0 because of

$$\sigma_n = \{x \mapsto x(ab)^n, y \mapsto (ab)^n y\} \circ \sigma_0.$$

Exploiting the analogy between the first and the second example above, we can easily construct the following example (and many more in this spirit): But the unification problem

$$xxyyxx =^? yyxyxyy$$

has only one most general unifiers

$$\sigma_n = \{x \mapsto z^3, y \mapsto z^2\},$$

and this is the only e–unifier.
The fifth example is taken from J. Karhumäki as well:

$$axxby =^? xaybx$$

has infinitely many most general unifiers

$$\sigma_{i,j} = \{x \mapsto a^i, y \mapsto (a^i b)^j a^i\}, i \geq 1, j \geq 0$$

but it has only one e-unifier $\sigma_{1,0} = \{x \mapsto a, y \mapsto a\}$ which is essential because of

$$\sigma_{i,j} = \{x \mapsto ya^{i-1}, y \mapsto (a^i b)^j xa^{i-1}\} \circ \sigma_{1,0}$$

The final example is a bit more elaborate but still in the same spirit.

$$zaxzbzy =^? yyzbzaz$$

has infinitely many most general unifiers

$$\sigma_n = \{x \mapsto b^{2n}a, y \mapsto b^n ab^n, z \mapsto b^n\}, n > 0$$

but it has only one e-unifier, namely $\sigma_1 = \{x \mapsto bba, y \mapsto bab, z \mapsto b\}$ because of

$$\sigma_n = \{x \mapsto b^{2n-2}x, y \mapsto b^{n-1}yb^{n-1}, z \mapsto b^{n-1}z\} \circ \sigma_1$$

3.1 String Unification with at most one variable is e-unitary

So let us assume our unification problem

$$\Gamma = \{u_1 =^? v_1, \ldots, u_n =^? v_n\}$$

over the signature Σ consists of at most one variable, but arbitrary many constants. Without loss of generality, each arbritrary set of string equations is equivalent to a single string equation preserving the solutions. For example Diekert used the following construction

$$\{u_1a \dots u_nau_1b \dots u_nb =^? v_1a \dots v_nav_1b \dots v_nb\}$$

where a and b are distinct constants. The two equational problems have the same solutions.

Let $\Gamma = \{u_0xu_1...xu_n = v_0xv_1...xv_m\}, u_i, v_i$ are ground strings, x in $\Sigma = X \cup F$ and $V = \mathbf{Var}(\Gamma) = \{x\}$. The following facts are well known.

1. The equation in Γ can be reduced to the form $u_0xu_1...xu_n = xv_1...xv_m$, where u_0 is not the empty string and either u_n is nonempty and v_m is empty or vice versa. This form implies also that any unifier is a prefix of the string u_0^k.

2. if $m \neq n$ there is at most one unifier.

3. If $m = n = 1$, i.e. $\Gamma = \{u_0x = xv_1\}$, and the unifiers are of the form: $x \mapsto (pq)^ip, i \geq 0$, where pq is *primitive*. Note: A word is primitive if it is not the power of some other word, i.e. it cannot be represented as uv^nw, for some words u, v, w and $n > 1$.

4. Considering $m = n > 1$ the unifiers are of the form: $x \mapsto (pq)^{i+1}p, i \geq 0$, where pq is *primitive.*

5. For a given Γ there exist at most one infinite solution of the form: $\sigma_i = \{x \mapsto (pq)^{i+1}p\}, i \geq 0$.

6. Unifiers of string equations with at most one variable are ground substitutions. We were not able to find a publication with a proof. We show this result below.

These results are now used to show that *string unification with only one variable* is *e-unitary.* The first step is to prove that all unifiers are ground substitutions. The second step is to prove that all unifiers share an essential unifier.

Proposition 3.1 *Let* $\Gamma = \{u_0xu_1...xu_n = xv_1...xv_n\}$ *be a string equation with at most* one *variable* x *and* $\mathcal{U}(\Gamma) = \{x \mapsto (pq)^{i+1}p\}, i \geq 0$. *Then* $\mathbf{Var}(pq)$ *is empty, i.e. all unifiers are ground substitutions.*

Proof.

1. Suppose p contains a variable z, i.e. $p = p_1zp_2$ where p_1 is ground. Applying the unifier $x \mapsto (pq)^{i+1}p$ yields

   $$u_0(pq)^{i+1}pu_1 \dots = (pq)^{i+1}pv_1 \dots = u_0(p_1zp_2q)^{i+1}pu_1 \dots = (p_1zp_2q)^{i+1}pv_1 \dots$$

   Consider the prefixes $u_0p_1 \dots = p_1z \dots$ Since $|u_0p_1| \geq |p_1z|$ and u_0 is nonempty, z must be a symbol in u_0p_1, which is impossible.

2. Suppose q contains a variable z, i.e. $q = q_1 z q_2$ where q_1 is ground. Applying a unifier $x \mapsto (pq)^{i+1}p$ yields

   $$u_0(pq)^{i+1}pu_1 \ldots = (pq)^{i+1}pv_1 \ldots = u_0(pq_1zq_2)^{i+1}pu_1 \ldots = (pq_1zq_2)^{i+1}pv_1$$

   Consider the prefixes $u_0pq_1 \ldots = pq_1z \ldots$. Since $|u_0pq_1| \geq |pq_1z|$ and u_0 is nonempty, z must be a symbol in q_1 which is impossible.

Hence $\mathbf{Var}(pq)$ is empty. ■

Theorem 3.2 *String unification with one variable is e-unitary.*

Proof. Without loss of generality, let $\Gamma = \{u_0xu_1...xu_n = xv_1...xv_n\}$ be a string unification problem in *one* variable x and

$$\mathcal{U}(\Gamma) = \{\{x \mapsto (pq)^{i+1}p\} : i \geq 0\}.$$

Then there are the following decompositions, where V ={x}

1. In case of $n = 1$ and p is empty then

   $$\{x \mapsto (pq)^i p\} =^V \{x \mapsto (pq)^i x\} \circ \{x \mapsto p\} \circ \varepsilon$$
   if p is nonempty then
   $$\{x \mapsto q^i\} =^V \{x \mapsto q^{i-1}x\} \circ \{x \mapsto q\}$$

   either $\{x \mapsto p\}$ or $\{x \mapsto q\}$ are essential unifiers.

2. In case of $n > 1$ and p is empty then

   $$\{x \mapsto (pq)^{i+1}p\} =^V \{x \mapsto (pq)^i x\} \circ \{x \mapsto pqp\}$$
   or p is nonempty then
   $$\{x \mapsto q^{i+1}\} =^V \{x \mapsto q^i x\} \circ \{x \mapsto q\}$$

   either $\{x \mapsto pqp\}$ or $\{x \mapsto q\}$ are essential unifiers.

Hence the unification problem is e-unitary. ■

3.2 String unification is e-infinitary

String unification with at most one variable in the signature Σ is e-finitary as we have seen above and surely there are many more special cases of signature restrictions, where the set of e-unifiers is always finite. Special cases of this nature have been investigated extensively for the solvability problem of words[4].

Theorem 3.3 *String unification with more than one variable is e-infinitary*

[4] Google scholar finds 70,300 entries in 0.13 sec for "word equation" (not all of which is relevant) and several 100,000 more entries if you are patient enough to continue the search and to filter gold from garbage.

Proof. For $\Gamma = \{xby = ayayb\}$ the set of essential unifiers is

$$e\mathcal{U}(\Gamma) = \{\{x \mapsto ab^n a, y \mapsto b^n\} : n > 0\}$$

Correctness
Any substitution $\sigma_n = \{x \mapsto ab^n a, y \mapsto b^n\}$ is a unifier since $(xby)\sigma_n = ab^n abb^n = ab^n ab^{n+1} = (ayayb)\sigma_n$.

Completeness
We show that any unifier is of the form $\{x \mapsto ab^n a, y \mapsto b^n\}$. Now considering some unifier $\{x \mapsto u, y \mapsto v\}$. Since $\Gamma = \{xby = ayayb\}, u = au'$ and $v = v'b^k$, $k > 1$. Applying the unifier in $xby = ayayb$ yields $au'bv'b^k = av'b^k av'b^k b$. Since v' can not contain any a $v' = b^i$. Hence the unifier is now $\sigma = \{x \mapsto au', y \mapsto b^j\}$, where $j = i + k$. Thus $\Gamma\sigma = \{au'bb^j = ab^j ab^j b\}$, $xbb^j = ab^i ab^i b$, and $x = ab^j a$.

Essential
We show that the set $\{\{x \mapsto ab^n a, y \mapsto b^n\} : n > 0\}$ is e-minimal. So take any pair of different unifiers $\{x \mapsto ab^m a, y \mapsto b^m\}$ and $\{x \mapsto ab^n a, y \mapsto b^n\}$ and we show that they are incomparable with respect to the part ordering. Suppose $m < n$, then $|ab^n a| > |ab^m a| > |b^m|$, therefore $\{x \mapsto ab^m a, y \mapsto b^m\} \neq \alpha\{x \mapsto ab^n\ y \mapsto b^n\}\beta$. Now the other way round; the longer unifier could contain the shorter, but then there exists a decomposition $\{x \mapsto ab^n a, y \mapsto b^n\} = \alpha\{x \mapsto ab^m a, y \mapsto b^m\}\beta$ where w.l.o.g. $\alpha = \{x \mapsto u, y \mapsto v\}$. Since $y \mapsto b^n$ it follows $v \in \{b, y\}^*$, because the longer unifier maps y to b^n. Now let‘s look at $x \mapsto ab^n a$ contains only two times the letter a, $u = u_1 x u_2$ with $u_1, u_2 \in \{a, b\}^*$ or $u \in \{a, b\}^*$. In the latter case x occurs not in the range of α, and x is in the domain of $\{x \mapsto ab^n a, y \mapsto b^n\}$. Thus, in this case $\{x \mapsto ab^m a, y \mapsto b^m\}$ is not a part of $\{x \mapsto ab^n a, y \mapsto b^n\}$. In the case of $u = u_1 x u_2 = ab^n a$ with $x \mapsto ab^m a$, it follows that u_1 and u_2 are empty, contradicting $ab^n a \neq ab^m a$. ■

3.3 A General A-Theorem

Let E be a set of equational axioms containing the associativity axiom of a binary operator $*$, i.e. $A = \{x * (y * z) = (x * y) * z\}$ and $E = A \cup R$, where R is some set of equations. We call the theory modulo E A-separate, if any equation in R can not be applied to a pure string $x_1 * x_2 * \cdots * x_n$, where the brackets are suppressed.

For instance consider distributivity (which is an infinitary unification theory, see [19]

$$D = \{x * (y + z) = (x * y) + (x * z), (x + y) * z = (x * z) + (y * z)\},$$

then the theory of $E = A \cup D$ is A-separate. To see this, note that no equation of D can be applied to a string of $x_1 * x_2 * \cdots * x_n$, simply because there are no sums involving the plus sign $+$, but each equation in D has the sum symbol $+$ on its left and on its right side.

Formally, $E = A \cup R$ is A-separate, if for all elements u of the A-theory $u =_R v$ implies $u = v$.

Theorem 3.4 *All A-separate E-theories are e-infinitary*

Proof. Reconsider the unification problem of section 3.2 above. It has in the associative sub-algebra infinitely many e-unifiers. Each of the elements of the range of the essential unifiers is not affected by the remaining equational axioms in $R = E \backslash A$, since E is A-separate. Hence each A-separate theory is e-infinitary. As noted above the theory $A \cup D$ is A-separate. ■

Corollary 3.5 *The theory $A \cup D$ is e-infinitary.*

4 Idempotent semigroups are e-finitary

The following theory of *idempotent Semigroups or Bands* defined by

$$AI = \{f(x, f(y, z)) = f(f(x, y), z), f(x, x) = x\}$$

demonstrates another interesting applicability of essential unifiers. This theory is not A-separate. This theory is nullary with respect to the instantiation order, since there are solvable AI-unification problems which do not posses a minimal complete set of AI-unifiers with respect to the instantiation ordering [1, 15].

However, with respect to the part ordering $\trianglelefteq_E$ this well-known situation changes completely as this theory is essentially finitary. Associativity and idempotency constitute the algebra of idempotent strings and it was shown in [5] that:

Theorem 4.1 *The theory AI is not nullary with respect to essential unifiers*

Proposition 4.2 *AI is not unitary with respect to essential unifiers.*

And finally the most striking result:

Theorem 4.3 *The theory AI is finitary with respect to essential unifiers.*

5 A derivation system for A–Unification

Let Σ be the set of symbols (alphabet) and let X be the set of variables. Let u, v, w be strings, i.e. elements of the free monoid $(X \cup \Sigma)^*$. Let $\Gamma = \{u =^? v\}$ be a A-unification problem. A solution σ is a substitution, such that the equality $u\sigma = v\sigma$ is valid, denoted by $\sigma \models u = v$.

Let Λ be the homomorphism between the strings of the free monoid $(X \cup \Sigma)^*$ into $\mathcal{P}^{\Sigma}(X)$, where $\mathcal{P}(X)$ are the polynomials in X, that is defined by

$$\Lambda : x \mapsto \begin{cases} x \text{ for } x \in Var(\Gamma) \\ 1 \text{ otherwise} \end{cases} \quad \text{and} \quad \Lambda(uv) = \Lambda(u) + \Lambda(v).$$

Extend the notation for unification problems $\Gamma = \{u =^? v\}$ by

$$\Lambda(\Gamma) = \{\Lambda(u) =^?_P \Lambda(v)\}\}$$

mapping a string unification problem to a system of linear equations, where P is the set of Peano axioms; and for substitutions $\sigma = \{x_1 \mapsto u_1, \ldots, x_n \mapsto u_n\}$ to

$$\Lambda(\sigma) = \sigma\Lambda = \{x_1 \mapsto \Lambda(u_1), \ldots, x_n \mapsto \Lambda(u_n)\}.$$

Since substitutions are in the following assumed to consists only of free variables, the image is a vector of lengths $\sigma\Lambda = \{x_1 \mapsto |u_1|, \ldots, x_n \mapsto |u_n|\}$.

For instance let Σ be the alphabet $\{a, b\}$ and $\Gamma = \{xby =^? ayayb\}$. Then

$$\Lambda(\Gamma) = \{x + 1 + y =^? 1 + y + 1 + y + 1\} = \{x =^? y + 2\}.$$

For the unifier $\sigma = \{x \mapsto abba, y \mapsto bb\}$,

$$\Lambda(\sigma) = \{x \mapsto 4, y \mapsto 2\},$$

which is obviously a solution of $\Lambda(\Gamma)$.

Lemma 5.1 *If σ is an A-unifier for Γ, then the linear diophantine equation $\Lambda(\Gamma)$ has an integer solution $\Lambda(\sigma) : x \mapsto \Lambda(\sigma(x))$.*

Proof. Follows from the homomorphism definition. ■

Let $\sigma = \sigma_\top \sigma_\bot$ be a leaf decomposition for a substitution σ, $\sigma_\bot(x) \in \Sigma \cup X$. Define for a solution $\alpha : X \mapsto \mathcal{N}$ of a linear diophantine equation Δ, i.e. $\alpha \models \Delta$, a substitution δ_α with $\delta_\alpha(x) = x_1 \ldots x_n$, where $n = \alpha(x)$ and all x_i are free variables, i.e. not in $Var(\sigma)$.

Let Ψ be the reduction system of the following reduction rules

$$\begin{array}{ll} \text{Truncation} & \dfrac{[u_l u_r =^? u_l v_r, S]}{[u_r =^? v_r, S]}, \quad \dfrac{[u_l u_r =^? v_l u_r, S]}{[u_l =^? v_l, S]} \\ \text{Generation} & \dfrac{[\Gamma, S], \alpha \models \Lambda(\Gamma), \exists \lambda : X \to \Sigma \cup X : \delta_\alpha \lambda \in e\mathcal{U}(\Gamma)}{[\Gamma, S \cup \{\delta_\alpha \lambda\}]} \end{array}$$

Define $A \Longrightarrow_\Psi B$ if $\frac{A}{B}$ in Ψ, and let $\Longrightarrow_\Psi^*$ be the transitive closure of $\Longrightarrow_\Psi$.

Lemma 5.2 $[\Gamma, \emptyset] \Longrightarrow_\Psi^* [\Gamma', S \cup \{\sigma\}]$ *iff* $\sigma \models \Gamma$.

Proof. "$\Rightarrow$" by definition of the Generation Rule. "$\Leftarrow$" $\sigma \models \Gamma$ implies $\Lambda(\sigma) \models \Lambda(\Gamma)$. Let $\Lambda(\sigma) =: \alpha$. Thus the Generation Rule is applicable with $\sigma = \delta_\alpha \lambda$ for a λ. ■

Define the ordering $\alpha \leq \beta$ by $\sum_{x \in Dom(\alpha)} \alpha(x) \leq \sum_{x \in Dom(\beta)} \beta(x)$.

Lemma 5.3 $\lambda \trianglelefteq \sigma$ *implies* $\Lambda(\lambda) \leq \Lambda(\sigma)$.

Proof. $\lambda \triangleleft \sigma$ implies that there exists α and β such that $\sigma = \alpha\lambda\beta$. Thus $\Lambda(\sigma) = \Lambda(\alpha) + \Lambda(\lambda) + \Lambda(\beta)$. Hence

$$\sum_{x \in Dom(\sigma)} \Lambda(\sigma(x)) = \sum_{x \in Dom(\alpha)} \Lambda(\alpha(x)) + \sum_{x \in Dom(\lambda)} \Lambda(\lambda(x)) + \sum_{x \in Dom(\beta)} \Lambda(\beta(x)).$$

Hence $\Lambda(\lambda) \leq \Lambda(\sigma)$. ■

Lemma 5.4 *Let $\Gamma = \{u =^? v\}$ be an A-unification problem with $\alpha \models \Lambda(\Gamma)$, such that there exists $\lambda : X \rightarrow \Sigma \cup X$ with $\delta_\alpha \lambda \models \Gamma$, then λ is unique.*

Proof. Syntactic unification is unitary. Hence there exists a function *uni* that maps a solution α of $\Lambda(\Gamma)$ to a unifier $uni(\alpha) = \delta_\alpha \lambda$. ■

Lemma 5.5 *Let Γ be an A-unification problem and $\alpha < \beta$ be two solutions of $\Lambda(\Gamma)$. If there exist the two unifiers $uni(\alpha)$ and $uni(\beta)$, then $uni(\beta) \not\trianglelefteq uni(\alpha)$.*

Proof. Suppose the contrary $uni(\beta) \trianglelefteq uni(\alpha)$. Note $\Lambda(uni(\beta)) = \beta$ and $\Lambda(uni(\alpha)) = \alpha$. That implies $\beta \leq \alpha$, a contradiction. ■

A controlled algorithm for enumerating the essentials could look like

```
FOR ALL $i \geq 0$ DO COMPUTE
    $S(i) = \{\alpha \mid \alpha \models \Lambda(\Gamma), \sum_{x \in Dom(\alpha)} \alpha(x) = i\}$ (* diophantine equation *)
    $U(i) = \{uni(\alpha) \mid \alpha \in S(i)\}$                    (* real unifiers *)
    $E(i) = E \cup \{\lambda \in U(i) \mid \neg\exists\sigma \in E : \sigma \trianglelefteq \lambda\}$     (* essential unifiers *)
    EXIT WHEN $P(E(i), \Gamma) = \emptyset$
    WHERE $P(E, \Gamma) = \{\sigma : X \rightarrow \Sigma^* \mid \sigma \models \Gamma \wedge \forall\beta \in E : \beta \not\trianglelefteq_A \sigma\}$
END FOR
```

Lemma 5.6 *For a finite set E of substitutions and a finite equational problem Γ modulo A the above predicate $P(E, \Gamma)$ is decidable.*

Proof.

$$\begin{aligned} P(E,\Gamma) &= \{\sigma : X \rightarrow \Sigma^* \mid \sigma \models \Gamma \wedge \forall\beta \in E \neg\exists\alpha, \gamma : \sigma = \alpha\beta\gamma\} \\ &= \{\sigma : X \rightarrow \Sigma^* \mid \sigma \models \Gamma \wedge \forall\beta \in E : \sigma \neq \alpha\beta\gamma\} \end{aligned}$$

To simplify the notation we assume without loss of generality that the substitutions might even erase variables, i.e., we consider homomorphisms in the free monoid instead of homomorphisms in the free semi group. The case considerations of the proof for semi groups is rather cumbersome but straight forward by introducing equation variations where the variables that might be erased are eliminated from the beginning.

Proposition 5.7 *For a pair of substitutions β and σ there exists a set $\mathbf{\Gamma}$ of equational systems with*

$\sigma = \alpha\beta\gamma$ if and only if at least one equational system of $\mathbf{\Gamma}$ is solvable.

Proof. Consider the following simplifications: If there is a factorzation $\sigma = \alpha\beta\gamma$ then without loss of generality there exists also a factorization $\sigma = \alpha'\beta\gamma'$ where the domain and the set of variables in the range of α' are made disjoint by renamings introducing free variables, i.e.

$$(Dom(\alpha') \cap VRan(\alpha')) \setminus Dom(\beta) = \emptyset$$

and for γ

$$Dom(\gamma') \cap VRan(\gamma') = \emptyset \text{ and } VRan(\gamma') \cap Dom(\alpha') = \emptyset$$

Let further without loss of generality

$$VRan(\alpha') \cap Dom(\gamma') = \emptyset \text{ and } Dom(\gamma') \subseteq VRan(\beta) = \emptyset$$

which can be reached by applying γ on α. Thus for $\sigma = \alpha\beta\gamma = \alpha'\beta\gamma'$ for each x in the domain of σ the following equation is valid

$$\sigma(x) = \begin{cases} \gamma'(\alpha'(x) = \alpha'(x) & x \in Dom(\alpha'), R_\alpha(x) = \emptyset \\ \gamma'(\beta(\alpha'(x))) & x \in Dom(\alpha'), R_\alpha(x) \neq \emptyset \\ \gamma'(\beta(x)) & x \notin Dom(\alpha'), x \in Dom(\beta) \\ \gamma'(x) & x \notin Dom(\alpha'), x \notin Dom(\beta) \end{cases}$$

where $R_\alpha(x) = Var(\alpha'(x)) \cap Dom(\beta)$.

Define a family of equation systems under the hypothesis of a family of $R_\alpha(x) \subseteq Dom(\beta)$, where x varies in $Dom(\sigma)$, and a domain $D_\alpha = Dom(\alpha) \subseteq Dom(\sigma)$. Thus all variations of hypotheses are defined by σ and β. Let the equation system be

$$\begin{array}{llll} x_\beta & =^? & \beta(x_\beta) & x_\beta \in Dom(\beta) & (1) \\ x_\sigma & =^? & u_{x_\sigma x_\beta} x_\beta v_{x_\sigma x_\beta} & x_\sigma \in D_\alpha, x_\beta \in R(x_\sigma) & (2) \\ x_\sigma & =^? & x_\beta & x_\sigma \notin D_\alpha & (3) \\ x_\sigma & =^? & \sigma(x_\sigma) & x_\sigma \in Dom(\sigma) & (4) \end{array}$$

$\Rightarrow$: $\sigma = \alpha\beta\gamma$ implies that there is a $\Gamma \in \mathbf{\Gamma}$ with a solution $\delta \models \Gamma$: Consider Γ with $D_\alpha = Dom(\alpha)$ and for $x \in Dom(\sigma)$ let $R_\alpha(x) = Var(\alpha(x)) \cap Dom(\beta)$. The equation sub-system (1) and (4) are obviously solved by β and σ. Equational sub-system (2) is solvable, since for every $x \in Dom(\sigma)$ that is mapped by α onto a string $\alpha(x)$ that contains a variable out of the domain of β, there are bindings for a prefix and a suffix, that solve each equation in (2). (3) is also solvable, since there exist a γ such that for each $x \in Dom(\sigma) \cap Dom(\beta)$ that does not occur in $Dom(\alpha)$ it is shown by $\sigma(x) = \gamma(\beta(x))$ that the equation is solvable.

$\Leftarrow$: If there is a $\delta \models \Gamma \in \mathbf{\Gamma}$, then there are α and γ with $\sigma = \alpha\beta\gamma$: Suppose Γ with D_α and $R_\alpha(x)$, $x \in Dom(\sigma)$. Define $\alpha(x) = \delta(u_{x_\sigma x_\beta}) x_\beta \delta(v_{x_\sigma x_\beta})$ for all $x \in D_\alpha$. Define $\gamma(x) = \delta(x)$ for all $x \notin D_\alpha$. This corresponds to the normalization considerations in the beginning of the proof of the proposition and shows $\sigma = \alpha\beta\gamma$. ■

Proposition 5.8 *For a finite set E of substitutions β there exists an equational system $\Gamma(E)$ with the property*

$$\sigma \models \Gamma(E) \textit{ if and only if there exists } \beta \in E \textit{ with } \sigma = \alpha\beta\gamma$$

Proof. This follows from the fact that the equational theory of semi groups is boolean closed, i.e. a boolean combination of equation systems can be expressed by a single equation system. ■

From the proposition follows that

$$\begin{aligned} P(E,\Gamma) &= \{\sigma : X \to \Sigma^* \mid \sigma \models \Gamma \wedge \forall \beta \in E : \neg\sigma = \alpha\beta\gamma\} \\ &= \{\sigma : X \to \Sigma^* \mid \sigma \models \Gamma \wedge \bigwedge_{\beta \in E} \neg\sigma \models \Gamma(\beta)\} \end{aligned}$$

For each equational problem Γ there exists a complementary equational problem $\overline{\Gamma}$ with $\neg(\sigma \models \Gamma)$ if and only if $\sigma \models \overline{\Gamma}$. Hence it follows

$$\begin{aligned} P(E,\Gamma) &= \{\sigma : X \to \Sigma^* \mid \sigma \models \Gamma \wedge \bigvee_{\beta \in E} \sigma \models \overline{\Gamma(\beta)}\} \\ &= \bigcup_{\beta \in E} \{\sigma : X \to \Sigma^* \mid \sigma \models (\Gamma \cup \overline{\Gamma(\beta)})\} \end{aligned}$$

Thus the predicate $P(E,\Gamma) = \emptyset$ is reformulated as a finite set of string unification problem which is known as being decidable. ■

As a corollary from the above considerations we state

Theorem 5.9 *The algorithm enumerates all essential unifiers for an A-unification problem Γ and terminates if the set of essential unifiers is completed.*

6 Conclusion

The results reported above come as a disappointment to some extent: while the set of *e*-unifiers is considerably 'smaller' — albeit still infinite in general — than the set of most general unifiers for a string unification problem, the anticipated collapse of the infinitary theory to an *e*-finitary theory did not hold up to scrutiny.

This may not surprise those familiar with this problem, in spite of the simplicity and immediate intuitiveness of the problem formulation (using strings) the solvability as well as the unification problem turned out to be of exceptional difficulty and complexity.

For practical purposes, e.g. as a unification component within an automated theorem proving system, based on resolution or rewriting,there are two problem left over

1. To find a unification algorithm which generates — as efficiently as possible — the set of *e*-unifiers

2. To show how the reasoning machinery can be built upon *e*-unifiers instead of most general unifiers.

We have a solution to both problems, however far from anything practically useful: the unification algorithm is to resolution based theorem proving what the addition-and-multiplication unit is to a general purpose computer — and hence deserves the utmost effort in engineering, measured not in MiPs but in LiPs (logical inferences per sec, i.e. in fact the number of unifications p.sec) which was the hallmark of the fifth generation computer race in the 1980s.

Acknowledgements

We like to thank our first reviewer and the anonymous reviewers of this workshop.

The theoretically inclined reader, familiar with unification theory, may forgive us our lengthy motivational introduction and notational sections but unfortunately, as our first reviewer pointed out rightly, the concept of an essential unifier is still not as widely known as we would like, let alone fully appreciated as a fundamentally new concept. The results and counterexamples in this paper, however, make sense only in this particular context — hence the introduction.

References

[1] F. Baader. *Unification in idempotent semigroups is of type zero.* Journal of Automated Reasoning, 2(3), 1986.

[2] M. Lothaire (ed) *Algebraic Combinatorics on Words* Cambridge University Press, 2001.

[3] K. Schulz *Word Unification and Transformation of Generalized Equations* Journal of Automated Reasoning, vol. 11, pp 149-184, 1993

[4] F. Baader, T. Nipkow *Term Rewriting and all That* Cambridge University Press, 1998.

[5] M. Hoche, P. Szabó. *Essential Unifiers* Journal of Applied Logic, vol. 4, no. 1, 2006 University Press.

[6] F. Baader, J. Siekmann. *Unification theory.* in D. Gabbay, C. Hogger, and J. Robinson, eds, "Handbook of Logic in Artificial Intelligence and Logic Programming", 1994, Oxford University Press.

[7] F. Baader, W. Snyder. *Unification Theory.* In A. Robinson, A. Voronkov, editors, Handbook of Automated Reasoning Volume 1. Elsevier Science Publishers B. V. (North-Holland), 2001.

[8] H.-J. Bürckert, A. Herold, and M. Schmidt-Schauß. *On equational theories, unification and (un)decidability.* J. of Symbolic Computation 8, 3-49, 1989

[9] N. Dershowitz, J.-P. Jouannaud. *Rewrite Systems.* In J. van Leeuwen, editor, Handbook of Theoretical Computer Science, chapter 6, pages 244-320. Elsevier Science Publishers B. V. (North-Holland), 1990.

[10] E. Eder. *Properties of substitutions and unifications.* Journal of Symbolic Computation, 1(1):31-46, 1985.

[11] G. Huet. *A complete proof of correctness of the Knuth and Bendix completion algorithm.* Journal of Computer and System Sciences, 23:11-21, 1981.

[12] C. Kirchner, H. Kirchner. *Rewriting Solving Proving.* http://www.loria.fr/~ckirchne or http://www.loria.fr/~hkirchne

[13] G. Plotkin. *Building-in equational theories.* Machine Intelligence, 7:73-90, 1972.

[14] J.A.Robinson. *A machine-oriented logic based on the resolution principle.* Journal of the ACM, 12(1):23-41, 1965.

[15] M. Schmidt-Schauß. *Unification under Associativity and Idempotence is of type nullary.* Journal of Automated Reasoning, 2(3), 1986.

[16] J. Siekmann. *Unification and matching problems.* Ph.D. thesis, 1975, Essex University.

[17] J. Siekmann, P. Szabó. *A noetherian and confluent rewrite system for idempotent Semigroups.* Semigroup Forum, 25:83-110, 1982.

[18] J. Siekmann. *Unification theory.* Journal of Symbolic Computation, 7(3 & 4): 207-274, 1989. Special issue on unification. Part one.

[19] P. Szabó. *Unifikationstheorie erster Ordnung.* PhD thesis, University Karlsruhe, 1982.

[20] A. C. Varzi, *Parts, wholes, and part-whole relations.* The prospects of mereotopology, Data and Knowledge Engineering (DKE) Journal 20, North-Holland, Elsevier, 1996.

String Unification is Essentially Infinitary

Michael Hoche, Jörg Siekmann, Peter Szabo

RR-08-01

Research Report